

RIVISTA DI

POLITICA

I - III

ECONOMICA

CONTRIBUTI SCIENTIFICI:

Gabriele Ballarino

Piero Bini

Fabio Castiglionesi

Daniele Checchi

Giuseppe Garofalo

Adriano Giannola

Marco Leonardi

Paolo Liberati

Antonio Lopes

Francesca Mazzarella

Massimo Pallini

Silvia Pugliesi

Giorgio Szegö

Vincenzo Visco Comandini

Alberto Zazzaro

2013

RIVISTA DI

POLITICA

I-III



CONFINDUSTRIA

ECONOMICA

IN QUESTO NUMERO:

Guerino Ardizzi
Gabriele Ballarino
Giuseppe Bianchi
Piero Bini
Fabio Castiglionesi
Tatiana Cesaroni
Daniele Checchi
Giuseppe Garofalo
Adriano Giannola
Marco Leonardi
Paolo Liberati
Antonio Lopes
Francesca Mazzarella
Massimo Pallini
Carmelo Petraglia
Alessandro Petretto
Massimiliano Piacenza
Silvia Pugliesi
Ottavio Ricchi
Marco Rossi
Giorgio Szegö
Federico Tomassi
Gilberto Turati
Vincenzo Visco Comandini
Alberto Zazzaro

RIVISTA DI POLITICA ECONOMICA

Fondata nel 1911

Direttore

Marcella Panucci

Direttore Responsabile

Gustavo Piga

Comitato Scientifico

Presidente

Mario Baldassarri

Mario Anolli

Michele Bagella

Giorgio Basevi

Sebastiano Bavetta

Leonardo Becchetti

Pierpaolo Benigno

Carlotta Berti Ceroni

Magda Bianco

Franco Bruni

Giacomo Calzolari

Roberto Cellini

Daniele Checchi

Bruno Chiarini

Gabriella Chiesa

Innocenzo Cipolletta

Guido Cozzi

Valentino Dardanoni

Giorgio Di Giorgio

Massimo Egidi

Riccardo Fiorito

Michele Grillo

Luigi Guiso

Elisabetta Iossa

Fiorella Kostoris

Luca Lambertini

Stefano Manzocchi

Riccardo Martina

Alessandro Missale

Giuseppe Moscarini

Giovanna Nicodano

Francesco Nucci

Luigi Paganetto

Luca Paolazzi

Alberto Quadrio Curzio

Annalisa Rosselli

Lucio Sarno

Valeria Termini

Ufficio Editoriale: rpe@confindustria.it

Tel. +39 06 5903793 - Telefax +39 06 5903349

Indirizzo: Viale Pasteur, 6 - 00144 Roma

Coordinatore Editoriale: Adriana Leo - a.leo@confindustria.it

Sito web: <http://www.rivistapoliticaeconomica.it>

©Copyright 2013

Indice

Invited Paper

Financial Intermediation and Liquidity Fabio Castiglionesi	7
---	---

Invited Policy Paper

La Germania può essere un termine di paragone per l'Italia? Istruzione e formazione in un'economia di mercato coordinata Gabriele Ballarino - Daniele Checchi	39
Violare gli equilibri. Gli economisti italiani di sinistra nella crisi degli anni Settanta del Novecento Piero Bini	75
Variabili predittive dell'astensionismo elettorale: una modellizzazione della partecipazione al voto Giuseppe Garofalo - Silvia Pugliesi	113
La convergenza dello sviluppo finanziario tra le regioni italiane dal 1890 ad oggi Adriano Giannola - Antonio Lopes - Alberto Zazzaro	145
Un commento alla riforma Monti - Fornero Marco Leonardi - Massimo Pallini	199

Saggi

L'economia non osservata fra evasione e crimine: una rivisitazione del <i>Currency Demand Approach</i> con una applicazione al contesto italiano Guerino Ardizzi - Carmelo Petraglia - Massimiliano Piacenza - Gilberto Turati	229
Previsioni delle spese del bilancio dello Stato attraverso i flussi di contabilità finanziaria Giuseppe Bianchi - Tatiana Cesaroni - Ottavio Ricchi	271

On the Fuzzy Boundaries between Public
and Private in Health-Care Organization
and Funding Systems 327
Alessandro Petretto

Alcune implicazioni fiscali di scenari alternativi
alla proibizione della cannabis 371
Marco Rossi

Disuguaglianze, beni relazionali ed elezioni
nelle periferie di Roma 403
Federico Tomassi

Parole chiave

Industria postale e corriere espresso 453
Francesca Mazzarella - Vincenzo Visco Comandini

Rassegna bibliografica

Inequality and Instability
by GALBRAITH J.K. 503
a cura di Paolo Liberati

*The Basel Committee on Banking Supervision.
A History of the Early Years 1974 - 1997*
by GOODHART C. 509
a cura di Giorgio Szegö

INVITED PAPER

Financial Intermediation and Liquidity

Fabio Castiglionesi*

CentER, EBC and Tilburg University

This paper first reviews the basic theory of financial intermediation. In particular, the paper presents the main modeling device that rationalizes financial intermediaries as liquidity providers. In the second part of the paper we build on the basic model to analyze recent developments in the theory of financial intermediation. Especially motivated by the 2007/2008 financial crisis, we show how this literature could help to explain phenomena like financial contagion and systemic liquidity crises. Finally, we deal with the possible welfare benefit of liquidity regulation.

[JEL Classification: G21].

Keywords: financial intermediation; liquidity holding; inter-bank markets; contagion; systemic liquidity crisis.

* <fabio.castiglionesi@uvt.nl>, Department of Finance.

I would like to thank Sandro Brusco, Fabio Feriozzi, Guido Lorenzoni and Wolf Wagner for many discussions that shaped my understanding of the topic. I also thank Andrea Attar, Elena Carletti, Gabriella Chiesa and Lucy White for valuable feedbacks on my research. The usual disclaimer applies.

1. - Introduction

The challenge that theoretical models face to rationale the existence of financial intermediaries (or banks) is to highlight the reasons of their advantage over financial markets in allocating the resources optimally. Such advantage is clearly absent in environments where markets are perfect and complete (*i.e.*, in an Arrow-Debreu world). If households and firms have access to perfect and complete markets, a competitive equilibrium guarantees that financial markets channel resources efficiently and banks are redundant.¹

Therefore we have to depart from the Arrow-Debreu paradigm to explain the role of banks. The analysis of financial institutions become possible with the foundation of the economics of information started in the late 1970s. In this paper we build on one of these contributions: the Diamond and Dybvig (1983) model. They show how banks can be seen as a liquidity insurance mechanism. Banks provide a pool of liquidity and insure against individual (idiosyncratic) liquidity shocks. In this sense the bank can be interpreted as a coalition of depositors. As long as the idiosyncratic liquidity shocks are not perfectly correlated, the total liquidity hold by a bank increases less than proportionally of the sum of individual liquidity holding. Hence a bank can finance more profitable but illiquid investments.

Indeed in the economic system there are different liquidity needs. On the one hand, consumers face uncertain short-run expenses. In this case they need to hold liquid assets, which can be sold without liquidation costs (*i.e.*, loss of value in case of premature liquidation). On the other hand, investors need long-term financing to avoid premature and costly liquidation of their assets. Banks therefore can act as intermediary between the two sides of the economy, borrowing short-term from consumers and lending long-term to investors. Banks' essential function is of liquidity provision: they issue liquid assets (deposits) while financing illiquid ones.²

¹ See section 1.7 in FREIXAS J. and ROCHET J.C. (2008) for a formal proof.

² Alternative rationales are obtained focusing on the asset side of banks' balance sheet. In DIAMOND D. (1984) banks perform firms project monitoring on behalf of depositors (delegated monitoring), showing that there are scale economies in monitoring (a bank can finance many projects) and low cost of delegation. In HOLMSTROM B. and TIROLE J. (1998) firms are subject to moral hazard and liquidity risk. The latter could be in principle solved by firms holding diversified portfolios of each other shares. Such redistribution of liquidity is not optimal *ex-post* since it cannot satisfy the liquidity demand of all firms. Banks instead can commit to redistribute liquidity through credit lines (or short-term loans). KASHYAP A.K., RAJAN R. and STEIN J. (2002) study when there are synergies between the two sides of the banks' balance sheet.

Because of the liquidity provision function, banks hold a substantial amount of liquidity risk. The root of this risk lies precisely in the maturity mismatch between the liability and asset sides of their balance sheet. The easiest way banks face this liquidity risk is to hold directly liquid assets (reserves and treasury bonds), moreover banks also use interbank markets (or more generally money markets) where they coinsure the liquidity risk with other banks. In section 2 we analyze how a bank provides efficient liquidity coinsurance to heterogeneous consumers by being able to hold the optimal level of liquidity (contrary to the financial market). In section 3 we show how the liquidity coinsurance between different banks through the interbank market takes place. We then analyze the negative consequences (financial contagion) that the interbank market can be responsible of in case an adverse shock occurs to one bank. In section 4 we provide a theory of systemic liquidity crisis. In section 5 we build on the previous results to analyze prudential liquidity regulation in banking.

2. - The Model

In this section we describe the model by Diamond and Dybvig (1983), which represents the basic explanation of financial intermediaries as providing liquidity insurance to consumers facing uncertain liquidity needs. Consider an economy with three dates, $t = 0, 1, 2$, and a single consumption good that serves as the numeraire.

Consumers. There is a continuum of agents, which are *ex-ante* identical. They have 1 unit of the good at $t = 0$, but they consume either at $t = 1$ (impatient or early consumers) or $t = 2$ (patient or late consumers). Their preferences, can be expressed as

$$U(c_1, c_2) = \begin{cases} u(c_1) & \text{with probability } \omega \\ u(c_2) & \text{with probability } (1 - \omega) \end{cases}$$

They consider an institution that provides liquidity to depositors through demand deposits, and a separate one that provides liquidity to firms through short-term loans. If these two institutions cannot trade funds on the market and liquidity shocks affecting depositors and firms are imperfectly correlated, then they improve efficiency by pooling their resources.

The function $u(\cdot)$ is assumed to be increasing, strictly concave, twice continuously differentiable, and satisfies the Inada conditions. That is, $u'(\cdot) > 0$, $u''(\cdot) < 0$ and $\lim_{c \rightarrow 0} u'(c) = \infty$. Since there is a continuum of consumers, the probability ω represents also the fraction of early consumers. Consumers know their type only at $t = 1$ which is private information. The fraction of early consumers ω is known *ex-ante*, so there is no aggregate uncertainty.

Investments. There are two types of assets, both in perfectly elastic supply. The first asset is a storage technology, also called the short asset or the liquid asset, and it yields one unit of consumption at date $t + 1$ for each unit of consumption invested at date t , for $t = 0, 1$. The second asset, also called the long asset or the illiquid asset, yields $R > 1$ units of consumption at date 2 for each unit of consumption invested at date 0. However, if liquidated in $t = 1$, the long asset returns $L < 1$. Notice that investing in the liquid asset (*i.e.*, holding liquidity) is costly since the more profitable return of the illiquid asset is given up, but it is also costly to liquidate early the illiquid asset.

2.1 The Autarky Allocation

Let's analyze the simple situation in which each (atomist) consumer is isolated, so there is no possibility of trade (either through markets or banks). The consumers in this case has to invest his endowment directly in the assets available in the economy (*i.e.*, $x + y = 1$). The problem of the consumer is:

$$(I) \quad \max_{x, c_1, c_2} [\omega u(c_1) + (1 - \omega)u(c_2)]$$

subject to

$$c_1 = y + xL = (1 - x) + xL = 1 - x(1 - L)$$

$$c_2 = y + xR = (1 - x) + xR = 1 - x(R - 1)$$

The objective function reflects the uncertainty on when to consume. In each period the consumer faces a feasibility constraint. If the consumer turns out to be an early type, the consumption c_1 is equal to the amount invested in the liquid asset plus the return from the illiquid asset (which he has to liquidate with the low return L). If the consumer is a late type, the consumption c_2 is equal to the

liquid asset plus the return of the illiquid asset (with full return R). The feasibility constraints highlight the liquidity risk: if $x > 0$, then consumer's consumption will be higher or lower than 1 (his endowment). To avoid this risk the investment in the long asset x should be 0, but the consumer would consume 1 even if he turns out to be late consumer.

The solution to problem (1) is given by the following first-order condition (FOC):

$$\frac{u'(c_1)}{u'(c_2)} = \frac{(1-\omega)(R-1)}{\omega(1-L)}$$

To have an interior solution the RHS has to be greater than 1, which implies $\omega L + (1-\omega)R > 1$ (*i.e.*, the expected return of the illiquid asset is greater than the return on liquidity, or ω sufficiently low). This implies that $L < c_1^A < 1 < c_2^A < R$ and $x^A > 0$ (and consumer holds liquidity risk).

Notice $x^A = 1$ is never optimal for sufficiently high risk aversion. Otherwise, if RHS is less or equal than 1, we have the corner solution $x^A = 0$ and $c_1^A = c_2^A = 1$. To sum up: either the consumer partially covers his liquidity risk (with the inefficient early liquidation in case of being early consumer), or he gives up the investment in the illiquid asset (holding only liquidity).

2.2 The Market Allocation

Is there a possibility to improve upon the autarky allocation? The answer is yes. Consider there is a financial market in $t = 1$ in which the illiquid asset can be traded against the liquid one. This allows to sell the illiquid asset without incurring in costly liquidation. Early consumers sell a claim on x , while late consumers buy it. Let us indicate with p the price of the claim on one unit of x .

The individual constraints in each period become (again, $x + y = 1$)

$$c_1 = y + pxR = (1-x) + pxR$$

$$c_2 = y/p + xR = (1-x)/p + xR$$

since early consumers can issue claims for an amount xR at the price p , while late consumers can buy, with their liquidity, an amount of claims equal to y/p .

What is the equilibrium price \hat{p} of these claims? First notice that a price $\hat{p} > 1$ cannot be an equilibrium. Nobody would buy the illiquid asset, being more convenient to hold liquidity. A simple arbitrage argument shows that $\hat{p} = 1/R$.³ Considering the equilibrium value of \hat{p} , from the previous feasibility constraints we get the consumption achieved by the financial market $c_1^M = 1$ and $c_2^M = R$. The market allocation guarantees clearly a higher expected utility to the consumer than the autarky allocation (where either $c_1^A < 1$ and $c_1^A < R$ or $c_1^A = c_2^A = 1$).

Finally, the clearing condition in the market for claims on the illiquid asset at $t = 1$ is $\omega xR = (1 - \omega)(1 - x)R$ (that is, the quantity offered by early consumers is equal to the quantity demanded by late consumers). This implies that $x^M = 1 - \omega$. Therefore the financial market allows to avoid the inefficient early liquidation. Indeed the total amount of liquidity is equal to the total consumption of early consumers, that is $y^M = 1 - x^M = \omega = \omega c_1^M$. Even if the financial market avoids the inefficient early liquidation of the long asset, is there any mechanism that would allow an even higher expected utility for depositors?

2.3 *The Efficient Allocation and the Banking Solution*

To characterize the efficient allocation, we have to establish what the optimal liquidity insurance between different consumers type is. Assume there is a social planner able to transfer resources between agents. In particular, the planner transfers resources from late to early consumers (avoiding the inefficient liquidation) in $t = 1$, and from early to late consumers in $t = 2$. This implies that the social planner has only one aggregate resource constraint to satisfy (instead of the two individual constraints like in autarky or with the financial market). The aggregate resource constraint determines the aggregate level of consumption. The planner maximizes the objective function as in (1) subject to the aggregate constraint (again, $x + y = 1$)

$$(1 - x) + xR = \omega c_1 + (1 - \omega)c_2$$

where on the LHS there is the total amount of resources invested on the assets and on the RHS we have the total consumption. The planner's ability to transfer resources between the different consumers represents the liquidity coinsurance.

³ Assume $p > 1 = R$, by selling R claims and buying 1 unit of the illiquid asset there is a riskless profit of $pR - 1$ in $t = 1$. Assume $p < 1 = R$, by selling 1 unit of the illiquid asset (giving up R) and buying $1/p$ claims there is a riskless profit of $(1/p) - R$ in $t = 2$.

Given that there is no aggregate uncertainty, the optimal liquidity holding is to have just enough for consumption in $t = 1$. The planner in this way minimizes the use of the low-return liquid asset avoiding as well the costly early liquidation of the illiquid asset. That is, the planner chooses $y = (1 - x) = \omega c_1$. Considering the aggregate constraint we have that consumption in $t = 2$ is equal to

$$c_2 = \frac{(1 - \omega c_1)R}{1 - \omega}$$

Inserting the last expression in (1), the planner problem becomes unconstrained

$$\max_{c_1} \left[\omega u(c_1) + (1 - \omega) u\left(\frac{(1 - \omega c_1)R}{1 - \omega}\right) \right]$$

and the planner's FOC is

$$\frac{u'(c_1)}{u'(c_2)} = R$$

that does not depend on ω and L . Let (c_1^*, c_2^*) the efficient allocation, *i.e.* the one that satisfies the planner's FOC.

Assume the consumer is sufficiently risk-averse. In particular, the coefficient of relative risk aversion is greater than 1, that is $-cu''(c)/u'(c) > 1$.

This implies that $cu'(c)$ is decreasing in c , and therefore $u'(1) > Ru'(R)$. This means that the market allocation $(c_1^M, c_2^M) = (1, R)$ does not satisfy the planner's FOC and hence it is not efficient. By increasing c_1 and decreasing c_2 the planner can offer the optimal allocation $c_1^* > 1$ and $c_2^* < R$, such that the FOC is satisfied and an higher expected utility is obtained.

Early consumers are getting insurance in the optimal allocation (they have an higher consumption than in autarky and in the market allocation). Since $R > 1$, we have $c_1^* < c_2^*$ and there is only partial insurance from liquidity risk (consumption is not equalized across types). This implies that the optimal allocation is incentive compatible (*i.e.*, everybody reports his own type to the planner). The social planner is a theoretical device, but which institution can implement the

optimal allocation? We show that the bank can do it with the drawback of generating a second bad equilibrium.

The banking allocation. Assume there is a bank that collects the endowments of consumers and offers them demand deposit contracts (which allow consumers to choose the timing of withdrawals). The deposit contract guarantees the optimal consumption (c_1^*, c_2^*) . The bank uses the law of large numbers to economize on the use of liquidity. The bank does not know the identity of each early consumer, but it does know that a fraction ω of them will need to withdraw at $t = 1$. Then the bank anticipates that the total withdrawals in $t = 1$ are ωc_1^* , therefore it will invest $y^* = 1 - x^* = \omega c_1^*$ like in the efficient allocation.

Notice: the bank does not observe the type of depositors (which is private information) but since optimal allocation is incentive compatible (*i.e.*, $c_1^* < c_2^*$) it does not matter as long as depositors have the correct beliefs. That is, if depositors believe that only early consumers withdraw in $t = 1$, the bank can satisfy early withdrawals. It also satisfies late withdrawals since $x^*R = (1 - \omega c_1^*)R = (1 - \omega)c_2^*$. Then the bank reproduces the efficient allocation.

However, if depositors hold different beliefs things turn out sour. If depositors believe that everybody else will withdraw at $t = 1$, then we have a bank run (the dark side of banks is that they are fragile). This is also an equilibrium because if everybody withdraws at $t = 1$ the bank is bankrupt since it offers $(1 - x^*) + x^*L = 1 - x^*(1 - L) < 1 < c_1^*$. The bank cannot guarantee the promised level of early consumption c_1^* so everybody has incentive to run the bank and the run becomes self-fulfilling. Late consumers would pretend to be early consumers. The bank run allocation is clearly dominated also by the autarky allocation. In a run all resources invested in the illiquid asset have to be liquidated, while in autarky only those invested by early consumers have to be liquidated.⁴

⁴ The possibility of bank run has generated a huge literature (not covered here). Empirical evidence suggests that to trigger a bank run more than just a change in expectations is needed. Something related to fundamental is required as well (see GORTON G., 1998). The theoretical possibility of run has motivated also policy intervention that would eliminate the run. In particular, depositors need to believe that banks are safe. To guarantee the stability of the banking system deposit insurance has been advocated. In this case the government guarantees that c_1^* will be paid in $t = 1$ no matter what. If the government can credibly hold on its promise, a late consumer will wait until $t = 2$ (knowing that all the other late consumers have the same incentive).

2.4 The Jacklin Critique

Beside the “dark side” of the bank run equilibrium, the banking allocation is fragile for a second (and more relevant) reason. In the previous section we analyzed the case in which the bank is able to offer the optimal allocation without the presence of a financial market. However, this result is jeopardized if bank and financial market are in place at the same time. Jacklin (1987) argues that the financial market’s ability to arbitrage the implicit cross-subsidy in favor of the early consumers undermines the insurance provided by the bank.

To see his argument formally, assume a bank offers the optimal allocation with a deposit contract (c_1^*, c_2^*) and at the same time a financial market is open at $t = 1$ to trade claims on the illiquid asset at the price \hat{p} . Let’s see what the incentives of a late consumer are. If he withdraws at $t = 1$ (pretending to be an early consumer) he would get $c_1^* > 1$ from the bank. If he invests this amount in the financial market (buying claims on the illiquid asset), he would obtain in $t = 2$

$$\frac{c_1^*}{\hat{p}} = c_1^* R > R > c_2^*$$

That is, late consumers get a higher level of consumption withdrawing in $t = 1$ and reinvesting in the financial market than withdrawing in $t = 2$ from the bank. Late consumers free ride on the bank’s costly provision of liquidity. The bank is undermined by the presence of a financial market, and the best feasible allocation is the market allocation $(1, R)$. The conflicting role between the bank and the market is a conceptual problem for banking theory. A perfectly liquid financial market (*i.e.*, claims sold and bought at no transaction costs in $t = 1$) makes liquidity provision (*i.e.*, transfer of wealth across periods) impossible. The literature has developed in several directions in response to this issue.

The first reaction was to assume limited participation in financial markets, *i.e.* some consumers do not have access to the financial market. Limited participation is exogenous, but a plausible interpretation could be that participation is limited because of private information on asset values. Under the limited participation assumption, Diamond (1997) shows that a financial market coexists with the bank as long as the fraction of late consumers with no access to financial markets is large enough. In that case the cross-subsidies provided by the bank’s deposit contract survive the presence of a financial market. A second explanation of the coexistence of markets and banks is based on the *ex-ante* heterogeneity of consumers. Gorton

and Pennacchi (1990) emphasize that trading on financial markets can impose costs on agents because of asymmetric information (trading against somebody better informed causes loss of money). Their model is similar to Diamond - Dybvig (1983) with the main difference that the return R of the illiquid asset is uncertain and consumers have different information about its realization. Given this set up market prices fail to be fully revealing and social optimality is not achieved. There is then room for a bank to offer deposit (or debt-like) contracts to avoid the cost of trading in the market. Some consumers prefer to hold a deposit contract, which is less information sensitive (since its payoff is independent of the realization of the return of the illiquid asset) and therefore has lower risk. Other consumers prefer to hold tradable claims and incur the informational risk related to the asymmetric information in the financial market.

The final, and likely more convincing, way to overcome the Jacklin Critique is to consider a dynamic environment. Qi (1994) embeds the basic Diamond - Dybvig model in an OLG framework. In each period t a new cohort enters the economy, and each cohort lives for three periods with a fraction of depositors consuming in $t + 1$ (early consumers) and the remaining in $t + 2$ (late consumers) as in the Diamond - Dybvig set up. It turns out that in the OLG model higher steady state payoffs can be realized than in the 3-period model, that is $c_{t+2} > R > c_{t+1} > 1$. The reason is that the investment/endowment brought by the new generation of consumers reduces the need to hold liquidity to face the liquidity needs of previous generations, then an higher amount of resources can be invested in the illiquid asset. At the steady-state no side trading among depositors in financial market can upset the banking contract. Since consumption for late consumers c_{t+2} is higher than R , the reward of staying with the bank until old is high enough to make it attractive to avoid trading in the financial market.

3. - Interbank Markets and Contagion

So far the liquidity provision of a bank was studied in isolation, without considering the interaction among different banks. Banks can indeed be exposed to different liquidity uncertainty because they can be localized in different regions or countries (demand side) or because they specialize in lending to a particular industry (asset side). When we have more than one bank the role of an interbank market (IBM) can be studied. The IBM (or more generally, the money market) is indeed the place where banks are active to lend and borrow liquidity, with the

scope of insuring each other of the liquidity risk. The IBM has the role to distribute the local/regional/industry-specific liquidity shock over the all banking system.⁵

TABLE 1

REGIONAL LIQUIDITY SHOCKS			
	A	B	Probability
S_1	ω_H	ω_L	1/2
S_2	ω_L	ω_H	1/2

3.1 Interbank Markets

To deal analytically with the IBM we have to introduce two changes in the model shown in Section 2. First, we assume $L = 0$, *i.e.* the illiquid asset gives 0 if liquidated in $t = 1$. This assures that bank run cannot occur (so the focus is on the liquidity insurance mechanism). Second, the fraction of early consumers ω is stochastic. For simplicity it can take two values $\omega_H > \omega_L$. Assume there are two banks A and B, then the possible states of nature are given in Table 1. *Ex ante* each bank does not know which fraction of early consumers will withdraw in $t = 1$ so there is aggregate uncertainty at the level of the single bank (but not in the whole system). We indicate with $\omega = E(\tilde{\omega}) = (\omega_H + \omega_L)/2$.

As said each single bank faces aggregate uncertainty preventing them to get the efficient allocation (c_1^*, c_2^*) that was possible without aggregate uncertainty in the Diamond - Dybvig model. The highest payoff the bank can offer is by a contingent deposit contract $(c_1^L, c_1^H, c_2^L, c_2^H)$. The depositors are then exposed to the realization of the liquidity shock (like the single consumer in the Diamond - Dybvig model). Since both banks face this problem they can set up an IBM to co-insure this liquidity risk.

Assume the banks have an IBM at $t = 1$. The bank with a shortage of liquidity in $t = 1$ (*i.e.*, with the high shock ω_H) borrows from the one with excess of liquidity (*i.e.*, with the low shock ω_L). At $t = 2$ the funds move in the opposite direction.

⁵ The interbank market has two functions. First, it has a micro nature since it allows banks to manage their idiosyncratic liquidity risk (this is the focus in this paper). Second, it has a macro nature since it is the place where central banks implement their monetary policy (not covered in this paper but it is a very active field of research, see FREIXAS X., MARTIN A. and SKEIE D., 2011).

The problem is

$$\max_{x, c_1^L, c_1^H, c_2^L, c_2^H} \frac{1}{2} [\omega_H u(c_1) + (1 - \omega_H) u(c_2)] + \frac{1}{2} [\omega_L u(c_1) + (1 - \omega_L) u(c_2)]$$

subject to

$$y = 1 - x = \frac{1}{2} \omega_L c_1^L + \frac{1}{2} \omega_H c_1^H$$

$$xR = \frac{1}{2} (1 - \omega_L) c_2^L + \frac{1}{2} (1 - \omega_H) c_2^H$$

Thanks to the IBM the only constraints that matter are the two aggregate constraints, and not the four autarky feasibility constraints (two constraints for each bank). In $t = 1$ the total liquidity resources have to satisfy the demand of all early consumers. In $t = 2$ the return of illiquid technology has to meet the demand of all late consumers. From the FOC it is easy to get that $c_1^H = c_1^L = c_1$ and $c_2^H = c_2^L = c_2$. The IBM then guarantees consumers with a fix repayment, and not dependent on the bank-specific shock. Moreover, given the FOC and considering the definition of ω , the problem with the presence of the IBM coincides with the social planner problem in Section 2. Therefore the allocation offered by the IBM is the efficient one (y^*, x^*, c_1^*, c_2^*) .

We have to finally check if the banks are able to effectively redistribute the liquidity in the IBM. In other words, we have to check if the IBM clears in both periods. In $t = 1$ the bank with the high shock has a liquidity need of $\omega_H c_1^* - y^* = \omega_H c_1^* - \omega c_1^* = (\omega_H - \omega) c_1^*$. The bank with the low shock has a liquidity surplus of $y^* - \omega_L c_1^* = \omega c_1^* - \omega_L c_1^* = (\omega - \omega_L) c_1^*$. Given that $(\omega - \omega_L) = (\omega_H - \omega)$ the IBM is in equilibrium in $t = 1$. A similar argument shows that the IBM also clears in $t = 2$.

It is important to notice that IBM provides the efficient allocation since the liquidity shocks in each bank are observable and therefore contractible. Bhattacharya and Gale (1987) provide one of the seminal paper on the IBM analysis. They analyze the case in which banks have incomplete information: each bank observe its own fraction of early consumers, but cannot observe those of the other banks. With this informational friction the IBM is not able to provide the efficient allocation anymore. Indeed, an *ex-post* IBM creates a free-rider problem, where each bank will rely on the *ex-post* IBM to provide them with liquidity and has an incentive to underinvest in the liquid asset ($y < y^*$). In aggregate, the liquidity in the IBM is limited by the sum of investments of the individual banks

in the liquid asset. The IBM cannot provide the efficient liquidity risk insurance. We abstract from this problem until Section 5.

3.2 Financial Contagion

Banks hence have incentive to hold overlapping claims in the IBM to insure each other from regional/local/sector-specific/idiosyncratic liquidity shocks. This *ex-ante* optimal arrangement can however cause distress *ex-post*. The failure of one bank can spread to the entire system through the established links in the IBM. When a bank suffers a liquidity crisis, the neighboring banks can suffer a loss because their claims on the troubled bank fall in value. If this effect is strong enough, it can cause a crisis in the adjacent banks. In extreme cases, the crisis passes from bank to bank in a contagious fashion.

The financial contagion outcome is then not based on bank runs (or sunspot) but on real linkages among banks. The disadvantage of treating contagion as a sunspot phenomenon is that any pattern of correlations is possible. Sunspot theories are equally consistent with contagion and the absence of contagion. Allen and Gale (2000) propose one of the first model on contagion in which they tried to establish a stronger result: every equilibrium of the model is characterized by contagion. This form of contagion must be driven by real linkages between banks (like those established in the IBM).⁶

TABLE 2

REGIONAL LIQUIDITY SHOCKS IN ALLEN AND GALE					
	A	B	C	D	Probability
S_1	ω_H	ω_L	ω_H	ω_L	1/2
S_2	ω_L	ω_H	ω_L	ω_H	1/2
\bar{S}	$\omega + \varepsilon$	ω	ω	ω	0

The Allen and Gale (2000) model. Like in the Diamond - Dybvig model, they consider a positive liquidation value L at $t = 1$ of the illiquid asset. Moreover, there are four banks which interact in the IBM market as explained in section 3.1. The liquidity shocks are state dependent according to Table 2.

⁶ FREIXAS X., PARIGI B. and ROCHET J.C. (2000) provide a model of financial contagion, not covered here, similar to ALLEN F. and GALE D. (2000).

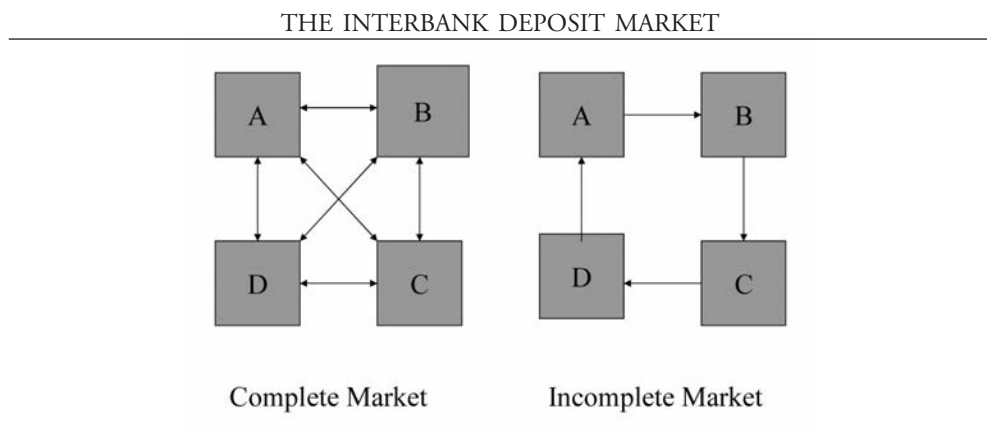
Notice that:

- i) there is a zero-probability state \bar{S} in which the total demand of liquidity in $t = 1$ is higher than the expected average shock ω ;
- ii) *ex-ante* each bank has the same probability of having a high (or low) liquidity shock;
- iii) the aggregate liquidity demand is the same in each state with positive probability (half the regions have the high shock ω_H and half have the low shock ω_L).

The optimal allocation (the one chosen by a social planner) is given as in Section 2 by (c_1^*, c_2^*) with $u'(c_1^*) = Ru'(c_2^*)$. The planner does not take into account the zero-probability state \bar{S} , moving resources in the other two states.

Like in section 3.1, the optimal allocation can be decentralized by an IBM since also banks do not consider the zero-probability state. Allen and Gale (2000) consider two IBM structures. First, a *complete* IBM structure: all banks are connected with each other. Second, an *incomplete* IBM structure: each bank holds deposits only with one adjacent bank. In particular, bank A holds deposits in bank B, bank B holds deposits in bank C, bank C in bank D and bank D in bank A. See Graph 1 for a graphical representation of the two structures.

GRAPH 1



With a complete IBM, each bank is negatively correlated with two other banks. Assume every bank holds $z = (\omega_H - \omega)/2$ deposits in each other bank. With this amount of interbank deposits the IBM is able to reproduce the optimal allocation. In $t = 1$ the bank with the high shock has a budget constraint equal to $\omega_H c_1 + zc_1 = y + 3zc_1$. In $t = 1$ the bank with the low shock has a budget constraint

equal to $\omega_L c_1 + 2z c_1 = y$. Both constraints reduce to the planner (aggregate) constraint $\omega c_1 = y$. A similar argument shows that also in $t = 2$ the IBM guarantees the same planner (aggregate) constraint. The complete structure of IBM delivers therefore the allocation (c_1^*, c_2^*) .

With an incomplete IBM, each bank is negatively correlated with its adjacent bank (every bank with a high liquidity shock in $t = 1$ has deposits in a bank with a low liquidity shock in $t = 1$, and *vice versa*). Assume every bank holds $z = (\omega_H - \omega)$ deposits in the adjacent bank. Let us check the feasibility constraints in the decentralized incomplete IBM. In $t = 1$ the bank with the high shock has a budget constraint $\omega_H c_1 = y + z c_1$. In $t = 1$ the bank with the low shock has a budget constraint $\omega_L c_1 + z c_1 = y$. Both constraints reduce to the planner (aggregate) constraints $\omega c_1 = y$. A similar argument shows that also in $t = 2$ the incomplete IBM constraints reduce to the planner aggregate constraint $(1 - \omega)c_2 = R x$. Also the incomplete structure of IBM delivers (c_1^*, c_2^*) . Although it is possible to achieve the first best with either complete or incomplete IBM markets, the implications for financial contagion are different.

Let us perturb the model to allow for the occurrence of the zero-probability state \bar{S} . All banks have a demand equal to the average ω , except bank A that has higher needs $\omega + \varepsilon$. This implies that the aggregate demand for liquidity is greater than the aggregate supply of liquidity. Assume that the banks offer the optimal deposit contract (c_1^*, c_2^*) , holding the amount z in the IBM which is necessary to implement the optimal contract.

Banks have to meet their promise to pay c_1^* to each depositor who demands withdrawal at $t = 1$. Otherwise, banks have to liquidate their assets. It is assumed the following pecking order of asset liquidation $1 < c_1^*/c_2^* < R/L$, which implies that banks first liquidate the storage technology (cost 1), then the IBM deposits (cost c_1^*/c_2^*) and then the illiquid asset (cost R/L). The first inequality is satisfied by the planner's FOC, while the second is satisfied by a sufficiently low L (which does not appear in the planner's FOC).

The value of a deposit in $t = 1$ is equal to the liquidation value of the bank's assets. Let q^i denote the value of the deposits in bank i at $t = 1$. All depositors must be treated equally, that is, every depositor gets q^i from bank i whether the depositor is a consumer or another bank. Then q^i must be determined simultaneously. Notice that if $q^i < c_1^*$, then depositors (either consumers or other banks) in bank i withdraw as much as they can.

Consider the incomplete IBM structure. If *all* depositors withdraw, the total demands for bank A is $1 + z$ (1 unit of deposits from consumers and z interbank

deposits from bank D). The liabilities of bank A are valued at $(1+z)q^A$. The assets consist of y units of the liquid asset, x units of the illiquid asset, and z interbank deposits in bank B. The assets are valued at $y + Lx + zq^B$. Assuming $q^B = c_1^*$ (i.e., bank B is solvent), the upper bound of the liquidation value in bank A at $t = 1$ is

$$\bar{q}^A = \frac{y + Lx + zc_1^*}{1+z}$$

As said the bank has to give to late consumers at least c_1^* to avoid a run. Therefore a bank with a fraction α of early consumers must keep at least $(1-\alpha)c_1^*/R$ of the illiquid asset to satisfy late consumers in $t = 2$. The amount of consumption that can be obtained by liquidating the illiquid asset without causing a run is then

$$b(\alpha) = L \left[x - \frac{(1-\alpha)c_1^*}{R} \right]$$

where $b(\alpha)$ can be considered the bank's buffer. When state \bar{S} occurs bank A will then suffer a run if the unexpected shock ε is sufficiently high, that is if $\varepsilon c_1^* > b(\omega + \varepsilon)$. Bank A is bankrupt, even if it has IBM deposits in bank B since these deposits are of no use. All banks liquidate their deposits to avoid the cost of early liquidation L . The mutual withdrawals simply cancel out.

Once bank A is bankrupt, there is a contagion to bank D. A deposit in region D is worth more than a deposit in bank A since $\bar{q}^A < q^D = c_1^*$, then bank D suffers a loss. The contagious loss to bank D has to be compared with bank D's buffer. If the loss is larger than the buffer then bank D is bankrupt too. The sufficient condition for the contagious loss to exceed the buffer in bank D is

$$(2) \quad z(c_1^* - \bar{q}^A) > b(\omega)$$

since zc_1^* is the amount promised to bank C, and $z\bar{q}^A$ is the amount obtained from bank A. The contagion can go forward. Bank D's bankruptcy causes a loss to bank C, and the accumulated contagion effect is large enough that bank C too is bankrupt. From region to region contagion gets larger and larger, because more banks are in bankruptcy and more losses have accumulated from liquidation cost L . It is possible to prove that once bank D goes bankrupt, all banks go bankrupt in a contagious "chain-reaction" fashion.

Consider now the complete IBM. What would be the contagion in this case? In this structure to deliver the first-best allocation each bank holds $z/2$ deposits in each of the other banks. The claim on any single bank is smaller than in the incomplete IBM structure, but the total claim $3z/2$ is larger. This implies that the maximum liquidation \hat{q}^A value in bank A is

$$\hat{q}^A = \frac{y + Lx + (3z/2)c_1^*}{1 + (3z/2)} > \bar{q}^A$$

Bankruptcy in bank A will be contagious to the neighboring banks if

$$z/2(c_1^* - \hat{q}^A) > b(\omega)$$

which is less likely to be satisfied than condition (2). Therefore the complete IBM is more resilient to contagion than the incomplete IBM.

The Brusco and Castiglionesi (2007) model. Allen and Gale (2000) rely on the unexpected aggregate shock to rationalize financial contagion. Such explanation can be considered unsatisfactory since it is equivalent to assume that either agents are irrational or that contracts are very costly to write upon states of nature that are very rare. Brusco and Castiglionesi (2007) provide an alternative explanation based on the observation that banks are protected by limited liability so they are subject to moral hazard. The only difference with respect the previous models is the presence of a “gambling” asset, which takes one unit of the good at $t = 0$ and transforms it at $t = 2$ into λR units (with $\lambda > 1$) with probability η , or 0 units with probability $1 - \eta$. The assumption is that $\eta\lambda < 1$, so that risk-averse and risk-neutral agents strictly prefer the safe illiquid asset that delivers R to the gambling asset. The opportunity of investing in the gambling technology appears with probability p .

When the gambling asset is successful the return is λRx , but only the portion Rx is observable and can be used to pay the depositors. The remaining fraction $(\lambda - 1)Rx$ is not observable and can be appropriated by the bank owners (consider them as on-the-job perks or simply money illegally diverted). When the gambling asset fails and its return is 0, the depositors receive nothing at $t = 2$. These assumptions imply that no contract can be made contingent on the realization λR of the gambling asset.

Brusco and Castiglionesi (2007) assume banks are run by investors/shareholders (who choose the type of investment) with risk-neutral preferences. At $t = 0$ investors are endowed with e units of the consumption good and nothing thereafter. Investors either consume or buy shares of the banks. If they buy shares, they are entitled to receive dividends d_t at $t = 1$ and $t = 2$. Investors' utility function is:

$$u(d_0, d_1, d_2) = R d_0 + d_1 + d_2$$

where $d_t \geq 0$ for $t = 0, 1, 2$. Since investors obtain a utility of Re by consuming their endowment in $t = 0$, they have to be rewarded at least R for each unit of consumption they give up in $t = 0$. If they buy bank shares for an amount e_0 , then $d_0 = e - e_0$. Thus investors utility when they buy bank shares is $R(e - e_0) + d_1 + d_2$. The participation constraint of the investors can be written as $R(e - e_0) + d_1 + d_2 \geq Re$, which implies $d_1 + d_2 \geq Re_0$.

The liquidity shocks are state dependent according to Table 1. So there is no unexpected (aggregate) liquidity shock. In this case the optimal allocation is still (c_1^*, c_2^*) since the social planner does not invest in the gambling asset and avoids the aggregate uncertainty at the level of the single bank. In the decentralized economy, the optimal allocation is not achievable.

To fix ideas, let us consider a bank in autarky (no IBM) and no moral hazard ($p = 0$). In this case aggregate uncertainty prevents to reach the first best allocation, but bank capital can increase welfare. Indeed, assuming bank capital is sufficiently high $e > \underline{e}$, the optimal allocation is $c_1^H < c_1^L \leq c_2^L = c_2^H$. We have that it is optimal to set $d_1^H = d_1^L = 0$. By paying no dividends in $t = 1$ the capital can be invested in the (more profitable) safe illiquid asset rather than in the liquid asset. Dividends can be made state-contingent in order to achieve identical consumption across states at $t = 2$ (recall this was not possible without bank capital).

When the bank in autarky has the possibility to invest in the gambling asset ($p > 0$), depositors have to ensure that the investors prefer the safe illiquid asset to the gambling one. Intuitively, this is the case when the amount of bank capital is large with respect to the amount invested in the illiquid asset x , since investors will be more reluctant to gamble with their own money. It turns out that, for a given investment in the illiquid asset x , the bank will invest in the safe asset if and only if the bank's capital is sufficiently high $e \geq \xi x$. Where $\xi = \eta(\lambda - 1)/(1 - \eta\lambda)$ is the lowest value of the ratio e/x such that the bank does not have an incentive to select

the gambling asset. If $e < \xi x$ then it becomes common knowledge that the bank invests in the gambling asset if available (*i.e.*, with probability p).

When the IBM is established we know that the banks achieve the first-best (c_1^*, c_2^*) , as long as the moral hazard is absent, since the IBM eliminates aggregate uncertainty. In this case the presence of bank capital is unnecessary, and its level would be indeterminate in both banks. What happens when we allow for both liquidity coinsurance and moral hazard? As before, the condition under which banks invest in the safe illiquid asset is $e \geq \xi x$. Therefore if $e \geq \xi x^*$ in *both* banks then the first best is attainable, bankruptcy and contagion are avoided. When capital is scarce (*i.e.*, $e < \xi x^*$), moral hazard cannot be prevented and depositors may accept the risk that the gambling asset is used. Depositors may prefer to bear the burden of financial instability rather than restricting illiquid asset x^* , provided that the probability of financial instability is limited, that is, p is low. Thus, if financial instability is accepted as a consequence of the opening of financial markets, it must be the case that instability is a rare event.

What is the intuition of this result? Let $U^g(p)$ the expected utility reached with the presence of IBM and banks investing in the gambling asset. The expected utility $U^g(p)$ is continuous and decreasing in p , and $\lim_{p \rightarrow 0} U^g(p) = U^*$. That is, the allocation obtained investing in the gambling asset can be arbitrarily close to the first-best expected utility when p gets close to zero. Let $U^{ng}(e)$ be the expected utility with IBM and moral hazard is prevented. The function $U^{ng}(e)$ does not depend on p , and since $e < \xi x^*$, the highest utility is achieved by setting $x^{ng} = e/\xi < x^*$. Clearly, $U^* > U^{ng}(e)$ since the optimal investment x^* in the safe illiquid asset has to be reduced to the level equal to x^{ng} . It follows that for p low enough, we have $U^* > U^g(p) > U^{ng}(e)$.

To complete the argument Brusco and Castiglionesi (2007) establish a link between the opening the IBM and financial instability. There are indeed parameter values for which depositors prefer to prevent investment in the gambling asset under autarky, but allow it when IBM open. This happens if the opening of the IBM, by bringing opportunities for coinsurance, increases the utility of the illiquid asset. It turns out that if R is sufficiently close to 1 (so holding liquidity is not too costly) and $e > \underline{e}$ then $x^A < x^*$. That is, when R is sufficiently low, the IBM allows the bank to reduce their investment in liquidity and increase the investment in the illiquid asset.

To sum up assume $x^A < x^*$, $e \in [\max(\xi x^A, \underline{e}); \xi x^*)$, and $p < \bar{p}$, then the two banks invest in the safe illiquid asset under autarky and in the gambling illiquid asset when the IBM is open. Under this allocation there is a (small) positive probability

of bankruptcy and contagion. Since $e \geq \xi x^A$ in autarky there is no gambling and bankruptcy. When financial linkages are established we have $x^A < x^*$ so depositors would like to increase the investment in the illiquid asset. Since $e < \xi x^*$ the capital is not sufficient to prevent investment in the gambling asset. Depositors have to choose between curbing the investment in the illiquid asset or accepting banks to gamble if possible. When the probability that the gambling asset will appear is sufficiently small, the second alternative is more attractive.

What about contagion? Consider four banks and the complete and incomplete IBM structures as in Graph 1. In Brusco and Castiglionesi (2007) bankruptcy occurs in $t = 2$ while in Allen and Gale (2000) it occurs in $t = 1$. Contagion will occur if the gambling asset of a bank fails and another bank is creditor to that bank in the IBM at $t = 2$. Assume for simplicity state S_1 realizes and that bank A is the gambling bank which defaults in $t = 2$. In the incomplete IBM structure, this implies that bank D defaults as well. But the chain stops there since bank C withdraws from D in $t = 1$. In the complete IBM structure, this implies instead that banks D and B default as well. The implication is that the complete IBM structure is less resilient than the incomplete structure (contrary to the Allen and Gale result). Finally notice that, while Allen and Gale (2000) assume exogenously a rare event that causes contagion, Brusco and Castiglionesi (2007) obtain that contagion is a rare phenomenon in equilibrium. Indeed only if the probability of bankruptcy is low is it optimal to create financial linkages and invest in the gambling asset.

TABLE 3

REGIONAL LIQUIDITY SHOCKS IN CASTIGLIONESI, FERIOZZI AND LORENZONI

States	A	B	Probability
<i>HH</i>	ω_H	ω_H	$(1 - p)/2$
<i>HL</i>	ω_H	ω_L	$p/2$
<i>LH</i>	ω_L	ω_H	$p/2$
<i>LL</i>	ω_L	ω_L	$(1 - p)/2$

4. - Interbank Integration and Liquidity Crisis

In this section we analyze an alternative explanation of liquidity crisis provided by Castiglionesi, Feriozzi and Lorenzoni (2012). They highlight that changes in the structure of financial markets enhance their ability to handle risk in normal

times but the same innovations that contribute to risk spreading can become sources of instability following an economy-wide shock. They focus on a specific form of structural change: financial integration in the interbank markets. In particular, they analyze how financial integration affects incentives of the banks to hold liquidity. Indeed more integrated interbank markets allow banks to borrow from other banks when they are in need of funds (and this has a stabilizing effect). However, the *ex-ante* incentives to hold liquidity can be reduced. If a correlated liquidity shock hits all the banks its consequences are more severe with integrated markets since liquidity holding is now reduced (and this is the destabilizing effect).

The set up is similar to Allen and Gale (2000) with only 2 banks. The main departure is that now the aggregate uncertainty is fully anticipated by the agents. The realization of the liquidity shocks are represented in Table 3. The parameter p represents the probability to find coinsurance in $t = 1$, and $(1 - p)$ measures the residual aggregate uncertainty. The higher p and the deeper is the integration in the interbank markets.

Consider the allocation problem in autarky (*i.e.*, without IBM). *Ex-ante* a bank in isolation has a probability 1/2 to have the high shock ω_H . The equilibrium allocation under autarky therefore solves

$$\begin{aligned} \max_{y, c_1^H, c_2^H, c_1^L, c_2^L} & \frac{1}{2} \left[\omega_H u(c_1^H) + (1 - \omega_H) u(c_2^H) \right] + \\ & + \frac{1}{2} \left[\omega_L u(c_1^L) + (1 - \omega_L) u(c_2^L) \right] \end{aligned}$$

subject to

$$\omega_s c_1^s \leq y$$

$$(1 - \omega_s) c_2^s \leq (1 - y)R + y - \omega_s c_1^s$$

$$s = L, H$$

Notice now liquidity can be rolled over in $t = 1$. If $y - \omega_s c_1^s > 0$ part of the liquid asset is rolled over to increase consumption in $t = 2$. When in state s there is positive rollover the consumption is equated across time (*i.e.*, $c_1^s = c_2^s$). Assume banks offer contingent deposit contracts, the optimal allocation is: $c_1^H < c_1^L \leq c_2^L < c_2^H$. Remarks:

- i) there can be positive rollover only in state L , never in state H ;
- ii) it is welfare improving to reduce c_1^L and increase c_1^H , but this is not feasible in autarky.

Under integration, the banks located in the two regions trade contingent credit lines on the IBM. In states HL and LH , the bank with high liquidity needs withdraw at $t = 1$ from the bank with low liquidity needs, and repay at $t = 2$. In states LL and HH the interbank market is useless. Each bank now offers a contingent deposit contract: $(c_t^{LL}, c_t^M, c_t^{HH})_{t=1,2}$. Recall from Section 3.1 that the IBM allows to perfectly coinsure the asymmetric liquidity shocks. Hence $c_t^{LH} = c_t^{HL} = c_t^M$, and states HL and LH will be referred as state M . The equilibrium allocation under integration (notice now $E(\vec{\omega}) = \omega_M$) solves

$$\begin{aligned} \max_{y, (c_t^{LL}, c_t^M, c_t^{HH})_{t=1,2}} & p \left[\omega_M u(c_1^M) + (1 - \omega_M) u(c_2^M) \right] + \\ & + (1 - p) \sum_{s \in \{L, H\}} \frac{1}{2} \left[\omega_s u(c_1^{ss}) + (1 - \omega_s) u(c_2^{ss}) \right] \end{aligned}$$

subject to

$$\begin{aligned} \omega_M c_1^M &\leq y; (1 - \omega_M) c_2^M \leq (1 - y)R + y - \omega_M c_1^M \\ \omega_s c_1^{ss} &\leq y; (1 - \omega_s) c_2^{ss} \leq (1 - y)R + y - \omega_s c_1^{ss} \\ &s = L, H \end{aligned}$$

The optimal allocation is $c_1^{HH} < c_1^M \leq c_1^{LL} \leq c_2^{LL} \leq c_2^M < c_2^{HH}$. Also in this case there can be rollover either in state LL only, or in states LL and M . There is never rollover in state HH . The effect of financial integration on liquidity holding is not obvious in this environment with minimal frictions. In principle the IBM has two opposite effects on the incentives to hold liquidity. On the one hand, banks in need of funds can borrow in the IBM and this give them an incentive to reduce liquidity holding. On the other hand, banks with excess liquidity can lend it and this increase the incentive to hold liquidity. Recall that in this set up banks can equally end up being lender or borrower in the IBM.

To show these contrasting effects formally, define the value function

$$\begin{aligned} V(y, \omega) \equiv \max \{ & \omega u(c_1) + (1 - \omega) u(c_2) \text{ s.t. } \omega c_1 \leq y \text{ and} \\ & (1 - \omega) c_2 \leq R(1 - y) + y - \omega c_1 \} \end{aligned}$$

The value function $V(y, \omega)$ is strictly concave in y (so the problem is well defined) and $\partial V(y, \omega)/\partial y$ is non-decreasing in ω (so the marginal value of liquidity is non-decreasing in the liquidity shock, that is the higher the liquidity shock and more valuable is the liquidity).

We can rewrite the problems under autarky and integration using the value function $V(y, \omega)$ to compare the marginal value of liquidity. The equilibrium liquidity in autarky is given by

$$\max_y \frac{1}{2} V(y, \omega_H) + \frac{1}{2} V(y, \omega_L)$$

while under integration solves

$$\max_y p V(y, \omega_M) + (1-p) \left[\frac{1}{2} V(y, \omega_L) \right]$$

Taking the FOCs of the two problems and taking the difference we obtain the difference between the marginal value of liquidity under integration and in autarky. We get

$$\frac{p}{2} \left(\frac{\partial V(y, \omega_M)}{\partial y} - \frac{\partial V(y, \omega_H)}{\partial y} \right) + \frac{p}{2} \left(\frac{\partial V(y, \omega_M)}{\partial y} - \frac{\partial V(y, \omega_L)}{\partial y} \right)$$

Given that $\partial V(y, \omega)/\partial y$ is non-decreasing in ω , the first term is negative and the second term is positive. The first term then represents the incentive to reduce the liquidity holding when the IBM is open (marginal utility of liquidity is lower with IBM than under autarky). This is the case of the borrowing bank with the high idiosyncratic liquidity shock that can dampen in the IBM. The second term instead captures the incentive to increase the liquidity holding, in this case marginal utility of liquidity is higher with IBM. This is the case of the lender bank that has the low idiosyncratic liquidity shock.

Similarly to Brusco and Castiglionesi (2007), a condition that guarantees the second term is equal to zero is R sufficiently close to 1. If this is the case, then there is positive rollover in all states except HH (in which rollover is never optimal). Since there is rollover in state LL and M , the marginal value of liquidity is the same in these two states (there is abundant liquidity) and so $\partial V(y, \omega_M)/\partial y = \partial V(y, \omega_L)/\partial y$. Therefore the equilibrium investment in the liquid asset is lower

under integration than in autarky (that is, $y^* < y^a$) when R is sufficiently low. Once the second term is equal to zero, y is strictly decreasing in p , indeed the first negative term increases linearly on p . More general sufficient conditions are difficult to obtain analytically, but Castiglionesi, Feriozzi and Lorenzoni (2012) show that, even if R is large, a coefficient of risk aversion higher than 2 is sufficient to induce the same illiquidity effect.

Once the link between IBM integration and its incentive to reduce liquidity holding has been identified, Castiglionesi, Feriozzi and Lorenzoni (2012) study the effect of the occurrence of state HH (we refer to the article for a more detailed analysis). When this state appears the banking system is holding a relatively low amount of liquid asset. This implies various side effects like sudden spikes and increased volatility of the IBM interest rates, severe drops in the level of consumption and negative skewness in consumption.

A detailed analysis of the effects of the liquidity crises once it unfolds is provided in their article.

5. - Regulation

The theory of financial intermediation that rationalizes banks as pool of liquidity gives an explanation for banks liquidity holding. With the exception of the Bhattacharya and Gale (1987) paper, so far we did not address explicitly the issue if this liquidity holding is indeed optimal. In other words, when is there any room for liquidity regulation? The approach is to identify clearly the micro-economic foundation of externality in liquidity provision that justifies regulation. Otherwise, any call for prudential regulation cannot be considered well justified (at least from the theoretical point of view).

This section presents two types of analysis on prudential liquidity regulation. First, based on the Allen and Gale (2000) model analyzed in Section 3, Castiglionesi (2007) shows how a central bank could force the banking system to hold the efficient amount of liquidity holding avoiding crisis and contagion. Second, building on the benchmark model by Castiglionesi, Feriozzi and Lorenzoni (2012) analyzed in Section 4 the regulatory implication of the relationship between IBM integration and illiquidity can be studied.⁷

⁷ For an approach to banking regulation alternative to the one reviewed in this paper see CASTIGLIONESI F. and WAGNER W. (2013).

5.1 Unexpected Aggregate Shock

Castiglionesi (2007) assumes the same liquidity shocks structure as in the Allen and Gale (2000) model in Table 2 with one difference: the state \bar{S} occurs with a very small but positive probability p . The states S_1 and S_2 occur each then with probability $(1 - p)/2$. Commercial banks and depositors attach a probability zero to the state \bar{S} , so there is no aggregate uncertainty for them. Castiglionesi (2007) considers the presence of a central bank (CB) that takes into account the actual (even if very small) probability of the state \bar{S} . The social planner takes now into account the small probability p , solving

$$\begin{aligned} \max_{x, c_1, c_2, \bar{c}_1, \bar{c}_2} & (1-p)[\omega u(c_1) + (1-\omega)u(c_2)] + \\ & + p \left[\left(\omega + \frac{\varepsilon}{4} \right) u(\bar{c}_1) + \left(1 - \omega - \frac{\varepsilon}{4} \right) u(\bar{c}_2) \right] \end{aligned}$$

subject to

$$\begin{aligned} y & \geq \omega c_1 \\ xR & \geq (1-\omega)c_2 \\ y & \geq \left(\omega + \frac{\varepsilon}{4} \right) \bar{c}_1 \\ xR & \geq \left(1 - \omega - \frac{\varepsilon}{4} \right) \bar{c}_2 \end{aligned}$$

The social planner does not implement any early liquidation of the illiquid asset (L sufficiently low). If p is sufficiently low, the optimal allocation is incentive compatible

$$\bar{c}_2 = \frac{Rx^*}{1-\omega-\varepsilon/4} > c_2^* = \frac{Rx^*}{1-\omega} > c_1^* = \frac{y^*}{\omega} > \bar{c}_1 = \frac{y^*}{\omega+\varepsilon/4}$$

Notice that in $t = 1$ it is efficient to give a lower consumption to early consumers when the aggregate liquidity shock occurs avoiding the costly liquidation of the illiquid asset. In $t = 2$ instead an higher consumption is obtained by late consumers when the aggregate shock occurs since the same amount of resources is available for less late consumers.

In the normal states S_1 and S_2 , the optimal allocation can be decentralized by banks through the IBM and the presence of the CB. Each bank offers a deposit

contract (c_1, c_2) , and it invests the deposits in the portfolio (y, x) . The CB imposes a reserve requirement ρ which is the minimum amount that banks have to invest in the storage technology. The timing of the contract in $t = 0$ is as follows.

- 1) Depositors and banks sign the deposit contract, and banks decide the investment portfolio. Both the consumption level and the investment are contingent on ρ .
- 2) The CB determines ρ .
- 3) Deposit contracts and investments are implemented before the realization of the liquidity shock. The rationale is that when depositors and banks sign the deposit contract they take as a parameter the condition of the economy (here represented by ρ).

The reserve requirements imposed by the CB affect the investment of the banks in the liquid asset. Banks anticipate that the CB is a benevolent institution that maximizes social welfare, and they rationally expect to use in $t = 1$ the amount of reserves to face early consumers. The competitive banks solve

$$\max_{x, c_1(\rho), c_2(\rho)} \omega u[c_1(\rho)] + (1 - \omega) u[c_2(\rho)]$$

subject to

$$y(\rho) + x(\rho) \leq 1 - \rho$$

$$\rho \leq \omega c_1 \leq y(\rho) + \rho$$

$$(1 - \omega)c_2(\rho) \geq Rx(\rho)$$

If $\rho = 0$, banks invest $y(0) = \omega c_1(0)$ in the liquid asset. If $\rho < \omega c_1(0)$, banks invest $y(\rho) = \omega c_1(0) - \rho$ in the liquid asset. If $\rho > \omega c_1(0)$, banks invest $y(\rho) = 0$ in the liquid asset and less resources are available to invest in illiquid asset. Clearly if the CB wants to affect the allocation chosen by the banks, it has to fix a reserve $\rho > \omega c_1(0)$. In this way the CB raises $c_1(\rho)$ and reduces $c_2(\rho)$. The allocation when $\rho > \omega c_1(0)$ is

$$y(\rho) = 0 \quad x(\rho) = 1 - \rho \quad c_1(\rho) = \frac{\rho}{\omega}; \quad c_2(\rho) = \frac{R(1 - \rho)}{1 - \omega}$$

However, there is an implicit cost of imposing a too high ρ , since a too high reserve could reduce too much consumption for late consumers therefore violating the incentive compatibility constraint. In order to have $c_1(\rho) \leq c_2(\rho)$ it has to be

$$\rho \leq \frac{R\omega}{1 + \omega(R-1)} < 1$$

Then the decentralized allocation with the CB is incentive efficient. It is also possible to show that the allocation is also feasible, *i.e.* the constraints faced by the banks in $t = 1$ and $t = 2$ are always satisfied in states S_1 and S_2 . Like in Allen and Gale (2000), the same decentralized allocation can be achieved by both IBM structures.

What is the value of ρ that implement the efficient allocation? The optimal ρ is $\rho^{CB} = y^*$, and the decentralized economy reaches the first-best allocation in states S_1 and S_2 . If $\rho < \omega c_1(0)$ then $y(\rho) < y^*$. In fact, when $\rho < \omega c_1(0)$, we have $y(\rho) = \omega c_1(0) < \omega c_1^* = y^*$ since the first-best solution allocates more consumption to early consumers. To get efficiency the CB has to fix a $\rho^{CB} = y^* > y(\rho)$, which guarantees the allocation $c_1(\rho) = c_1^*$ and $c_2(\rho) = c_2^*$. Notice that also in this case the CB and the decentralized banking system reach the first-best allocation, independently of the structure of the IBM.

We analyze now the occurrence of the crisis state \bar{S} . In Allen and Gale (2000) ρ is zero, then banks have to liquidate the illiquid asset since they do not have enough liquidity (including IBM liquidity) and contagion can occur. The CB can prevent financial contagion, allowing the economy to reach the efficient allocation also in state \bar{S} . The CB using the reserves avoid the inefficient liquidation of the illiquid technology and ensures that a financial crisis due to early liquidation in one bank does not spill over other banks.

The CB operates as follows. Suppose that ρ is sufficiently high, so that banks promise an amount of early consumption equal to ρ/ω in $t = 1$. When state \bar{S} occurs however the total demand for early consumption exceeds the quantity of the resources invested in the liquid asset since $(4\omega + \varepsilon)(\rho/\omega) > 4\rho$. The CB then declares the insolvency of the banking sector (since it cannot pay the promised level of consumption) and fixes the amount that can be withdrawn at $t = 1$ equal to $4\rho/(4\omega + \varepsilon)$. In this way the CB can shift part of the reserves to bank A hit by the liquidity shock. Fixing $\rho^{CB} = y^*$ then early consumers withdraw the efficient amount $\bar{c}_1 = 4y^*/(4\omega + \varepsilon)$.

Some comments are in order. First, the declaration of insolvency does not cause any turmoil. Late consumers do not run the banks because $\bar{c}_2 > \bar{c}_1$. Second, the declaration of insolvency is strictly related to the notion of incentive-efficiency. If the more realistic notion of constrained-efficiency is taken into account, then declaration of insolvency is not needed in order to avoid contagion (see Section 6 in Castiglionesi, 2007). Third, the role of the CB in preventing contagion does not depend on the structure of the interbank deposit market. The ultimate reason of contagion in Allen and Gale (2000) is that banks and depositors sign deposit contracts that are not contingent on all the states of nature. The CB becomes a way to complete the contracts, independently of the structure of the IBM.

5.2 *Interbank Integration*

The analysis in Section 4 assumes a frictionless Diamond-Dybvig-Allen-Gale environment with symmetric information, anticipated aggregate shock, and fully contingent deposit (also interbank) contracts. Indeed Castiglionesi, Feriozzi and Lorenzoni (2012) obtain the liquidity crisis in a Pareto-improving environment. The *ex-ante* expected utility of consumers is higher under integration than in autarky for all $p \in (0, 1]$ and it is strictly increasing in p . Hence, to rationalize prudential liquidity regulation at least one of the assumptions in the benchmark model has to be removed. There are several market imperfections that have been analyzed in the literature to advocate prudential regulation. For example, Allen and Gale (2004) focus on incomplete markets and Farhi, Golosov and Tsyvinski (2009) on consumers hidden trades.

Castiglionesi, Feriozzi and Lorenzoni (2012) assume the following friction: the allocation of liquidity in $t = 1$ takes place through a spot interbank market (instead of contingent credit lines). Banks can only borrow and lend *ex-post*, after the realization of the liquidity shocks. With spot markets the equilibrium is not generally constrained efficient (see Battacharya and Gale, 1987; Allen and Gale, 2004), which implies a role for prudential regulation. However, Castiglionesi, Feriozzi and Lorenzoni (2012) focus on how interbank integration affects the need of regulation (in the form of a minimum level of liquidity holding): does IBM integration increase or decrease the need of prudential regulation? Castiglionesi, Feriozzi and Lorenzoni (2012) show by means of examples that the need of prudential regulation can be higher when the banking system undertakes a process of integration especially in the initial stages of the process. The issue clearly deserves more study, but the message of the paper is that the implicit risk that integration poses for banking stability can (and should) be regulated.

BIBLIOGRAPHY

- ALLEN F. - GALE D., «Financial Contagion», *Journal of Political Economy*, no. 108, 2000, pages 1-33.
- .-, «Financial Intermediaries and Markets», *Econometrica*, no. 72, 2004, pages 1023-1061.
- BHATTACHARYA S. - GALE D., «Preference Shocks, Liquidity and Central Bank Policy», in BARNETT W. - SINGLETON K. (eds.), *New Approaches to Monetary Economics*, Cambridge University Press, 1987.
- BRUSCO S. - CASTIGLIONESI F., «Liquidity Coinsurance, Moral Hazard, and Financial Contagion», *Journal of Finance*, no. 62(5), 2007, pages 2275-2302.
- CASTIGLIONESI F., «Financial Contagion and the Role of the Central Bank», *Journal of Banking and Finance*, no. 31(1), 2007, pages 81-101.
- CASTIGLIONESI F. - FERIOZZI F. - LORENZONI G., *Financial Integration and Liquidity Crises*, manuscript, Tilburg University and Northwestern University, 2012.
- CASTIGLIONESI F. - WAGNER W., «On the Efficiency of Bilateral Interbank Insurance», *Journal of Financial Intermediation*, forthcoming, 2013.
- DIAMOND D., «Financial Intermediation and Delegated Monitoring», *Review of Economic Studies*, no. 51, 1984, pages 393-414.
- .-, «Liquidity, Banks, and Markets», *Journal of Political Economy*, no. 105, 1997, pages 928-956, reproduced as Chapter 5, in BHATTACHARYA S. - BOOT A. - THAKOR A. (eds.), *Credit, Intermediation, and the Macroeconomy*, Oxford University Press, 2004.
- DIAMOND D. - DYBVIK P., «Bunk Runs, Deposit Insurance and Liquidity», *Journal of Political Economy*, no. 91, 1983, pages 401-419.
- FARHI E. - GOLOSOV M. - TSYVINSKI A., «A Theory of Liquidity and Regulation of Financial Intermediation», *Review of Economic Studies*, no. 76(3), 2009, pages 972-992.
- FREIXAS X. - MARTIN A. - SKEIE D., «Bank Liquidity, Interbank Markets, and Monetary Policy», *Review Financial Studies*, forthcoming, 2011.
- FREIXAS X. - PARIGI B. - ROCHET J.C., «Systemic Risk, Interbank Relations and Liquidity Provision by the Central Bank», *Journal of Money, Credit and Banking*, no. 32, 2000, pages 611-638.
- FREIXAS X. - ROCHET J.C., *Microeconomics of Banking*, MIT Press, 2nd edition, 2008.
- GORTON G., «Banking Panics and the Business Cycle», *Oxford Economic Papers*, no. 40, 1988, pages 751-781.
- GORTON G. - PENNACCHI G., «Financial Intermediation and Liquidity Creation», *Journal of Finance*, no. 45, 1990, pages 49-71.

- HOLMSTROM B. - TIROLE J., «Private and Public Supply of Liquidity», *Journal of Political Economy*, no. 106(1), 1998, pages 1-40, reproduced as Chapter 6, in BHATTACHARYA S. - BOOT A. - THAKOR A. (eds.), *Credit, Intermediation, and the Macroeconomy*, Oxford University Press, 2004.
- JACKLIN C.J., «Demand Deposits, Trading Restrictions, and Risk Sharing», in PRESCOTT E.D. - WALLACE N. (eds.), *Contractual Arrangements for Intertemporal Trade*, University of Minnesota Press, Minneapolis, 1987.
- KASHYAP A.K. - RAJAN R. - STEIN J.C., «Banks as Liquidity Providers: An Explanation for the CoExistence of Lending and Deposit Taking», *Journal of Finance*, no. 57, 2002, pages 33-73.
- QI J., «Bank Liquidity and Stability in an Overlapping Generations Model», *Review of Financial Studies*, no. 7(2), 1994, pages 389-417.

INVITED POLICY PAPER

La Germania può essere un termine di paragone per l'Italia? Istruzione e formazione in un'economia di mercato coordinata

Gabriele Ballarino* Daniele Checchi**
Università di Milano

Germany can be taken as a reference case for reforming activities called for by the recent economic crisis. In this article we revise recent reforms in the labour market and in the educational sector. These reforms were effective in containing the unemployment, especially of young people, without altering the essence of the socio-economic compact in Germany.
[JEL Classification: I21; J24; Z13].

Keywords: labour market regulation; vocational training; skills.

* <gabriele.ballarino@unimi.it>, Dipartimento di Economia, Management e Metodi Quantitativi;

** <daniele.checchi@unimi.it>, Dipartimento di Scienze Sociali e Politiche.

Questo contributo è nato all'interno di un progetto di ricerca sponsorizzato dalla Fondazione Trelle e dalla Fondazione Rocca sulla organizzazione dei sistemi formativi. Gli autori ringraziano Gianfelice Rocca e Attilio Oliva, e Nazareno Panichella che ha collaborato per il reperimento e l'elaborazione dei dati.

1. - Presentazione

Per i mercati finanziari, la Germania oggi è un punto di riferimento: il suo debito pubblico è ritenuto così sicuro ed è così desiderato dagli investitori, che gli interessi che lo stato tedesco deve pagare per finanziarsi sono estremamente bassi (quando non addirittura negativi). Al di là della volatilità dei mercati finanziari, del peso eccessivo che oggi vi gioca la speculazione a breve termine, e degli squilibri che essa riproduce, non ci sono dubbi sul fatto che la valutazione del debito tedesco misuri qualcosa di strutturale, ovvero la solidità della società e dell'economia tedesche. Il nostro Paese ha molto da guadagnare nel prendere la Germania come un punto di riferimento per immaginare le necessarie riforme economiche e sociali che la grave crisi in corso rende meno eludibili che in passato.

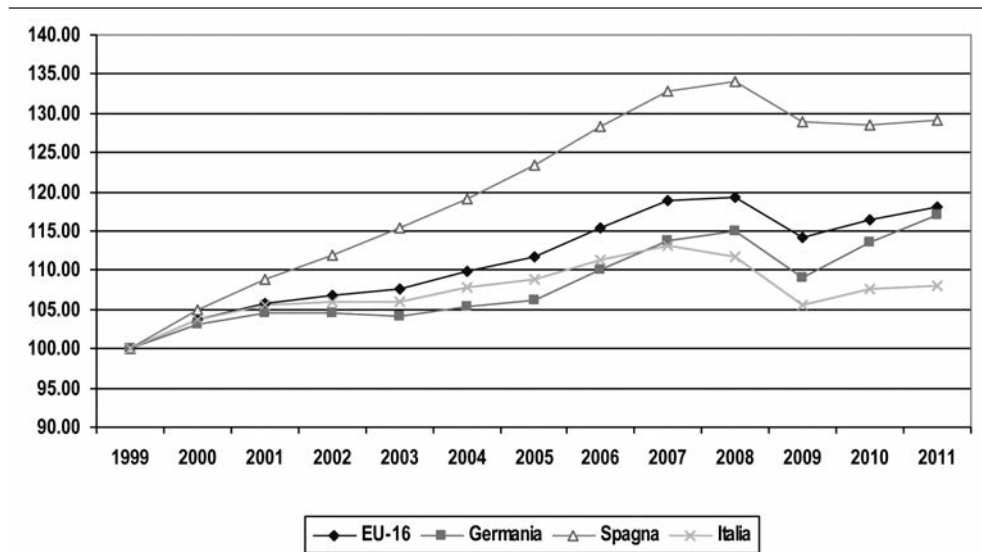
Questo contributo si concentra in particolare sui temi della scuola e del rapporto tra istruzione e mercato del lavoro, uno dei fattori di lungo periodo alla base della recente performance economica tedesca. Preliminarmente, però, guarderemo a quest'ultimo fenomeno, e alle riforme, non solo del mercato del lavoro, che l'hanno reso possibile. Mostriamo che queste ultime sono state incisive, ma che al tempo stesso esse non hanno modificato le specificità di lungo periodo del modello socio-economico tedesco. Passeremo quindi ad approfondire alcune delle specificità tedesche sul terreno dell'organizzazione del sistema scolastico (formazione duale) e del processo allocativo degli studenti tra diversi indirizzi. Concluderemo con il confronto tra sistemi universitari nei due paesi.

2. - Il primato tedesco: congiuntura e struttura

Il Grafico 1 mostra come a partire dal 2005, dopo la conclusione del ciclo di riforme di cui diremo più avanti, la crescita in Germania abbia cambiato ritmo rispetto agli anni precedenti, anche se il suo ritmo non sembra diverso da quello degli altri paesi, a parte il nostro che procede a più lentamente. Nel 2008, quando arriva la crisi, la crescita viene meno per tutti, Germania compresa, ma quando la crescita ricomincia, la Germania parte con un ritmo decisamente diverso, e migliore, di quello degli altri paesi europei. Ancora più interessante è il Grafico 2.

GRAF. 1

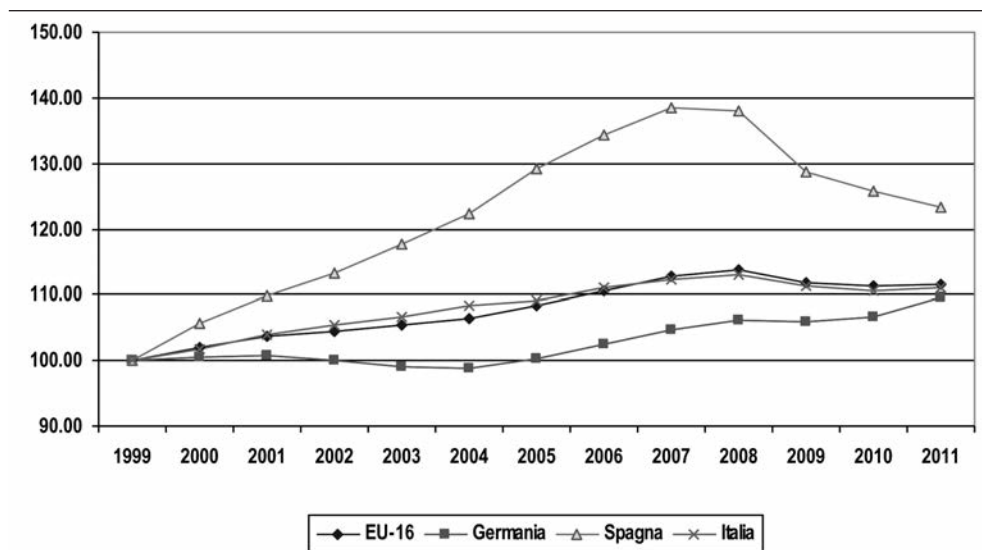
CRESCITA DEL PIL REALE (1999=100), IN EUROPA, GERMANIA, ITALIA E SPAGNA



Fonte: EUROSTAT.

GRAF. 2

VARIAZIONE DELL'OCCUPAZIONE (INDIVIDUI, 1999=100), IN EUROPA, GERMANIA, ITALIA E SPAGNA



Fonte: EUROSTAT.

Qui la particolarità della Germania è un'altra, più sorprendente: diversamente che negli altri paesi, in Germania la crisi non ha praticamente ridotto i livelli di occupazione, neanche al momento del crollo della crescita. Secondo molti osservatori, sia di orientamento liberista che anti-liberista, il fattore decisivo per la crescita sono state le riforme, che avrebbero smantellato il tradizionale modello tedesco dell'"economia sociale di mercato" con un'ampia liberalizzazione di un modello di regolazione dell'economia ad alta intensità di intervento pubblico (sia dello stato che delle associazioni degli interessi). A nostro parere, invece, il discorso è più complesso: le riforme sono state buone politiche nella misura in cui hanno saputo ri-calibrare le linee portanti del modello di regolazione ma senza assolutamente abbandonarlo, nella consapevolezza che questa è *la* buona politica, fonte dell'efficienza economica e della coesione sociale del paese. Il modello tedesco per qualche tempo è stato messo in ombra da un processo di globalizzazione dominato dalle più flessibili economie anglosassoni: alla fine degli anni '90, l'*Economist* poteva chiamare la Germania "il malato d'Europa", e formulare una preoccupata diagnosi della malattia: l'"Eurosclerosi". La malattia consiste nella rigidità delle istituzioni e delle forme regolative europee, ovvero nella scarsa capacità dei governi del vecchio continente di rinnovare le proprie risorse economiche per adeguarle ai nuovi scenari globali successivi alla caduta del muro di Berlino. Vedremo nel prossimo paragrafo quali sono i punti qualificanti del modello tedesco e in che cosa è consistita la ricalibratura della prima metà degli anni 2000.

3. - Il modello tedesco e la riforma degli anni 2000

Il modello socio-economico tedesco si basa sull'esportazione di prodotti ad alta intensità tecnologica e di alta qualità, la cui produzione si basa a sua volta su una forza lavoro qualificata e non conflittuale, ben retribuita e generosamente assistita in caso di disoccupazione. Gli studiosi parlano di un sistema di questo tipo come della "via maestra" allo sviluppo industriale, contrapponendola alla "scorciatoia" basata sulla compressione dei costi del lavoro, e per questo più esposta alle fluttuazioni del mercato e, soprattutto, alla concorrenza dei paesi di nuova industrializzazione, dove salari e costo del lavoro sono inferiori. Nel caso tedesco, l'alta qualificazione è garantita dal sistema duale di formazione professionale, di cui parleremo ampiamente più avanti. La bassa conflittualità è garantita, oltre che dallo stesso sistema duale, da relazioni industriali cooperative, che si fondano sull'ampia estensione della contrattazione collettiva e sul coinvolgimento delle

rappresentanze dei lavoratori nella gestione di alcuni aspetti della vita aziendale, in particolare per quanto riguarda la gestione delle risorse umane interne. Inoltre, le tutele contro il licenziamento sono elevate, così come l'ammontare e la durata dei sussidi di disoccupazione (nonostante le recenti riforme – vedi di seguito).

A livello di relazioni di lavoro, il sistema si basa quindi sulla stabilità del legame tra aziende e lavoratori, istituzionalmente garantita dalla cooperazione tra attori collettivi. Si parla a questo proposito di economia di mercato “coordinata”, contrapposta all'economia di mercato “liberale” tipica dei paesi anglosassoni, gli Stati Uniti in primo luogo. L'espressione “economia sociale di mercato”, tradizionalmente usata dalla politica tedesca e da qui importata in Italia, ha un significato molto simile. L'idea è quella di un intervento istituzionale capace di moderare gli effetti indesiderabili del libero mercato, mantenendone però le positive funzioni di sviluppo ed efficienza. L'esito macro più evidente è una minore disuguaglianza sociale. La Tavola 1 riporta due misure di disuguaglianza: l'indice Gini della dispersione dei redditi e il tasso di immobilità intergenerazionale dei livelli d'istruzione. I due indici vanno da 0 a 1, dove 0 indica la minima e 1 la massima disuguaglianza. In entrambi i casi la Germania si posiziona ai valori più bassi, insieme alla Francia, un'altra economia di mercato coordinata (sia pur con strumenti istituzionali molto diversi). È molto diversa la posizione dell'Italia, che pur essendo di norma classificato tra le economie di mercato coordinate risulta avere livelli di disuguaglianza dei redditi e di immobilità sociale più vicini a quelli dei paesi anglosassoni che a quelli degli altri paesi europei, e una tendenza alla diminuzione della disuguaglianza inferiore a quella riscontrata in questi ultimi.

TAV. 1

MISURE DI DISUGUAGLIANZA SOCIALE:
ITALIA, GERMANIA, FRANCIA, SPAGNA E GRAN BRETAGNA

	immobilità intergenerazionale dell'istruzione, per coorte di nascita						indice di Gini		
	1920	1930	1940	1950	1960	1970	metà 1980	metà 1990	metà 2000
Italia	0,73	0,74	0,62	0,48	0,33	0,43	0,337	0,350	0,352
Germania	0,52	0,37	0,33	0,29	0,32	0,33	0,286	0,290	0,298
Francia	0,52	0,48	0,41	0,31	0,22	0,24	0,290	0,281	0,281
Spagna	0,58	0,71	0,68	0,50	0,38	0,35	0,328	0,319	0,319
Gran Bretagna	0,35	0,31	0,42	0,36	0,33	0,35	0,332	0,341	0,335

Fonte: elaborazioni su dati EUROPEAN SOCIAL SURVEY (istruzione) e OECD (indice di Gini).

La natura sociale del sistema è evidente anche nei rapporti tra attori economici, dove le banche hanno stretti rapporti con le aziende ed esercitano, tramite l'erogazione di credito a buon mercato, una sorta di controllo generale sul sistema, evitando comportamenti aggressivi e troppo competitivi tra aziende, nonché invasioni di campo da parte di aziende straniere in caso di difficoltà di qualche operatore nazionale. Il termine *Deutschland AG* (alla lettera "Germania SPA"), caro ai giornalisti economici tedeschi, esprime bene le due caratteristiche del sistema: da un lato i rapporti tra gli attori economici nazionali sono più cooperativi che competitivi, ma in ultima analisi si tratta pur sempre di rapporti contrattuali; dall'altro siamo in un mondo individualista e liberale, lontano da qualsiasi modello sociale comunitario e organicista. I traumi del passato esercitano una vigilanza tranquilla ma inflessibile su tentazioni di questo genere. La stabilità di governo macro assicura la stabilità del sistema nelle aziende, ma ovviamente lo rende molto rigido rispetto a shocks macroeconomici esterni.

Le riforme

Negli anni successivi alla riunificazione tedesca, si sviluppa un dibattito interno sulla situazione e sulle prospettive del paese (spesso indicato come *Standort Deutschland*, alla lettera "dove si trova la Germania"), che rappresenta il background culturale da cui si sviluppa un importante ciclo di riforme. Alla fine degli anni '90 viene avviata nel quadro del "processo di Bologna" pan-europeo una riforma dell'istruzione superiore senza precedenti, su cui torneremo più avanti. Si ristrutturava il settore bancario, e nella prima metà degli anni 2000 arriva infine il momento della riforma del mercato del lavoro, articolata in una serie di quattro pacchetti di interventi, i cosiddetti "pacchetti Hartz" dal nome dell'amministratore delegato della Volkswagen Peter Hartz, presidente della commissione che nel 2002 mise a punto le proposte poi realizzate, con notevole impegno personale, dal cancelliere socialdemocratico Gerhard Schroeder. La classe dirigente del paese, tanto economica quanto politica, si impegna nella riforma con spirito *bipartisan*, formatosi con lo sforzo della riunificazione e la discussione sullo *Standort Deutschland*, la cui importanza è difficilmente sottovalutabile. Ovviamente non mancano difficoltà e resistenze: vale la pena di citare una frase pronunciata all'epoca dal cancelliere Schroeder, che ci sembra esprimere bene lo spirito della riforma: «se la Germania non modernizza le proprie istituzioni del mercato del lavoro, allora queste saranno modernizzate dalla forza bruta dei mercati globali, che lascerà ben poco spazio per una rete di protezione sociale». L'idea di fondo della riforma

è di cambiare atteggiamento verso i disoccupati, spingendoli al lavoro: lo stato continuerà a sostenerli, ma solo se essi si daranno da fare. In concreto, diminuiscono le garanzie contro la disoccupazione e la durata massima dell'accesso ai sussidi. Sono liberalizzate le agenzie di lavoro interinale, e vengono allentate una serie di misure di tutela dei salari, con l'effetto di rendere più facilmente disponibile alle aziende, soprattutto piccole, lavoratori a bassi salari. L'ultimo pacchetto di riforme, indicato come *Hartz IV*, il più controverso, trasforma la struttura dell'assistenza sociale, spostandola dalle politiche passive di sostegno incondizionato al reddito, verso le politiche attive, finalizzate al reinserimento occupazionale di tutti coloro che vengono giudicati in grado di lavorare. L'assistito che rifiuti, in determinate condizioni, una nuova opportunità di lavoro, perde anche il diritto all'assistenza, e i servizi hanno accesso alla verifica dei mezzi degli assistiti, per evitare casi di opportunismo.

È importante aggiungere che, nonostante una serie di imponenti manifestazioni e il mancato rinnovo di circa 100.000 tessere del partito socialdemocratico, in effetti la riforma raccoglie, se non il consenso esplicito, almeno la collaborazione fattiva del sindacato, un punto (se non *il* punto) decisivo per la sua implementazione efficace. Di fronte al pericolo di una crisi sistemica, le organizzazioni di rappresentanza di entrambe le parti, ugualmente in crisi di adesioni e di motivazione, reagiscono rilanciando la collaborazione a livello aziendale. Si osserva quindi quello che gli esperti chiamano un processo di "decentramento coordinato", con uno spostamento graduale e quanto meno traumatico possibile del baricentro delle relazioni industriali dal livello macro della contrattazione di settore a quello micro della contrattazione aziendale (i cui soggetti, i consigli dei lavoratori, sono più autonomi dalle centrali sindacali, di categoria e non, di quanto non accada in Italia). Per ridurre i licenziamenti al minimo, la contrattazione collettiva a livello aziendale utilizza la flessibilità interna nella gestione dei carichi e degli orari di lavoro individuali, uno strumento tradizionale della gestione del personale nelle aziende tedesche che viene ampiamente rinnovato. L'equivalente tedesco della nostra cassa integrazione, il *Kurzarbeit*, viene utilizzato massicciamente per ridurre le ore di lavoro, ma cercando di evitare nei limiti del possibile l'allontanamento dal posto di lavoro. Inoltre viene sfruttato intensivamente uno strumento introdotto negli anni '90, il conto del tempo di lavoro (*Arbeitszeitkonto*): quando un dipendente effettua ore di straordinario, anziché essergli pagate immediatamente queste vengono messe a suo credito in una sorta di cassa personale di ore lavorate e relativa retribuzione: se tutto va bene, il corrispettivo arriva qualche tempo dopo (l'arco di riferimento temporale può variare da un anno all'intera durata del rap-

porto di lavoro), ma se le cose per l'azienda vanno male, il credito disponibile nella cassa va a compensare la riduzione dell'orario di lavoro, consentendo all'azienda di riaggiustare i propri costi di fronte a una fase critica senza oneri per lo stato, come accade invece con la cassa integrazione italiana o il *Kurzarbeit*.

In Italia, invece, le contemporanee riforme del mercato del lavoro (note come riforme "Treu" e "Biagi") hanno incentivato le aziende a ricorrere alla flessibilità *esterna*, ricorrendo massicciamente ai rapporti di lavoro flessibili di nuova introduzione. Ma dato che la riforma ha flessibilizzato solo "al margine", nei nuovi contratti si sono trovati i giovani (flessibilmente) neoassunti: questo ha disincentivato l'investimento delle aziende nei giovani e dei giovani nelle aziende, senza consentire alcun recupero di produttività, anzi indebolendola ulteriormente. Nei termini proposti sopra, la riforma tedesca ha contribuito a mantenere l'economia sulla "via maestra" allo sviluppo, mentre le modalità con cui ha avuto luogo quella italiana rischiano di spingere il paese verso la "scorciatoia" della compressione dei costi e della competizione al ribasso. Gli esiti in termini di occupazione giovanile sono visibili nel Grafico 3 (a pag. 48).

Insomma, il "miracolo tedesco" tuttora in corso può forse essere visto anche come frutto della deregolazione e della flessibilizzazione del mercato del lavoro. Tuttavia, la riforma del mercato del lavoro è stata congegnata in modo tale da salvaguardare gli elementi essenziali del sistema, in particolare la collaborazione tra datori e lavoratori e tra le loro rappresentanze, facendone il punto centrale per la buona riuscita della ristrutturazione.

4. - Il sistema duale di formazione professionale

A questo punto vale la pena di descrivere in dettaglio il sistema di formazione professionale tedesco. Il sistema viene chiamato "duale" perché unisce apprendistato in azienda (*Ausbildung*) e formazione scolastica. Le sue origini risalgono alle regolazioni delle corporazioni artigiane delle città medievali, che consentivano l'esercizio del mestiere solo a coloro che avessero prestato servizio per un determinato periodo di tempo presso un artigiano. In Germania esse sono sopravvissute fino all'avvento dell'industria moderna, nella quale si sono riqualificate e ben integrate: non solo sono sopravvissute ai traumi della storia tedesca recente, ma sono diventate uno dei pilastri dell' "economia sociale di mercato" e dell'equilibrio socio-economico da essa garantito. Cominceremo con il descrivere l'alternanza scuola-lavoro, l'aspetto più noto del sistema. Passeremo quindi a illustrarne le

modalità di gestione e gli esiti occupazionali: potremo quindi capire perché il sistema è a tutt'oggi uno dei pilastri del modello socio-economico tedesco, non solo dal punto di vista dell'efficienza tecnica e produttiva ma anche da quello della creazione del consenso e della coesione sociali.

L'alternanza scuola-lavoro

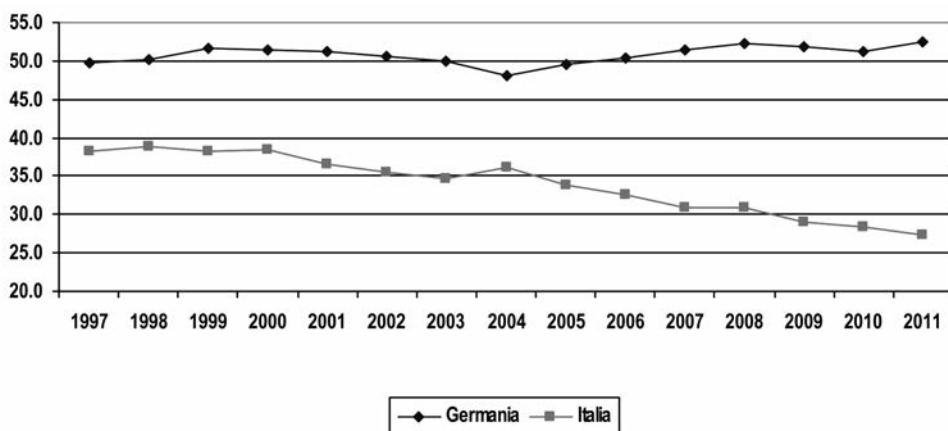
Il nucleo del sistema duale si trova oggi al livello scolastico equivalente alla nostra scuola secondaria superiore. Diversamente che nella maggior parte degli altri paesi europei e non, compreso il nostro, in Germania solo la scuola elementare è *comprehensive*, cioè prevede lo stesso curriculum per tutti gli alunni: a 10 anni questi vengono suddivisi in tre percorsi, solo uno dei quali (il *Gymnasium*), analogo al nostro liceo, consente l'accesso all'università. La maggioranza degli studenti che non seguono il percorso ginnasiale alla fine dell'obbligo scolastico, che dura fino a 16 anni, diventano apprendisti (*Auszubildende*): in complesso, circa il 60% di ogni coorte consegue in questo modo una qualifica professionale (cfr. il Graf. 8 *infra*).

Il contratto di apprendistato è un contratto di lavoro a tutti gli effetti, con un salario inferiore a quello normale, perché l'apprendista trascorre in azienda solo circa metà del proprio tempo di lavoro. Nel tempo rimanente, egli frequenta presso una *Berufsschule* o una *Beruffachschule* (scuola professionale; scuola professionale specializzata) un corso di studi i cui contenuti sono strettamente collegati alla sua occupazione. La durata dei corsi è variabile: lo standard per conseguire il diploma è normalmente di 3-4 anni, che negli anni 2000 sono stati in diversi casi suddivisi in due bienni. I corsi si concludono con il conseguimento di una qualifica occupazionale, diversamente denominata a seconda del settore e della durata dell'apprendistato. Gli esami sono sostenuti presso le camere di commercio locali, e i commissari sono nominati dalle camere stesse e dalle organizzazioni sindacali. Partecipano anche, ma con un ruolo marginale, gli insegnanti delle scuole. Le qualifiche sono molto dettagliate (nell'anno scolastico 2009-2010 erano 340) e sono aggiornate di continuo, per tenere il passo con l'evoluzione della tecnologia e dell'organizzazione del lavoro. È importante sottolineare che la gran parte degli apprendisti, ma non tutti (tra il 60% e il 70%), rimangono nell'azienda dove hanno svolto l'apprendistato, ma il conseguimento della qualifica professionale consente un rapido ingresso in azienda in quasi tutti i casi: il tasso di disoccupazione giovanile tedesco è tra i più bassi in Europa proprio per via di questo nesso istituzionale tra scuola e mercato del lavoro.

Il Grafico 3 mette a confronto il tasso di attività giovanile (15-24 anni) in Germania e in Italia. La differenza si aggira intorno ai 20 punti percentuali a favore della Germania. Per i 15-20enni il divario può dipendere dal fatto che gli apprendisti tedeschi sono classificati come occupati, diversamente dai giovani italiani che frequentano gli istituti tecnici e professionali e la formazione professionale regionale, ma tra la fascia d'età superiore il divario è interamente reale.

GRAF. 3

TASSO DI ATTIVITÀ GIOVANILE (% 15-24ENNI), IN GERMANIA E ITALIA



Fonte: EUROSTAT.

Cogestione e credenzialismo

La capacità del sistema tedesco di contenere la disoccupazione giovanile, che invece rappresenta un grave problema nel nostro Paese, ha attirato da tempo l'attenzione di studiosi e *policy-makers* italiani. Gli ammiratori italiani del sistema duale mettono al centro della loro analisi del sistema l'alternanza tra scuola e lavoro. In particolare, la regione Lombardia negli anni 2000 ha avviato diverse sperimentazioni con questa denominazione, esplicitamente ispirate dal sistema tedesco e dalla volontà di avvicinare scuola tecnico-professionale e mercato del lavoro. Sicuramente condivisibili negli intenti, queste politiche non sembrano però avere prodotto effetti di rilievo. Questo probabilmente perché si fondano su un'analisi che si ferma alla superficie del sistema tedesco e non ne considera le determinanti strutturali. Il motore del sistema duale, in effetti, non sta nell'alter-

nanza tra scuola e lavoro *per se*, ma in altri due elementi: si tratta della cogestione del sistema duale tra le parti sociali e del valore occupazionale delle qualifiche (“credenzialismo” nei termini della sociologia dell’istruzione). Sono questi due elementi, tra loro strettamente connessi a trasformare l’alternanza scuola-lavoro da strumento pedagogico in una vera e propria modalità di regolazione dell’economia. Il termine stesso che viene utilizzato in italiano è fuorviante: *alternare* scuola e lavoro non è difficile, basta organizzare tirocini. Quello che accade in Germania è invece una forma di *integrazione* tra scuola e lavoro di ben maggiore complessità, i cui fondamenti istituzionali sono appunto la cogestione e il credenzialismo.

La cogestione del sistema duale da parte delle associazioni di rappresentanza ha luogo ai diversi livelli in cui esso si articola. Abbiamo già visto a livello di azienda che esse gestiscono gli esami, di cui hanno anche la responsabilità finanziaria, nella misura in cui per gli esaminatori non sono previsti compensi: essi svolgono questa funzione durante il proprio orario di lavoro, quindi a spese dell’azienda (o del sindacato). In azienda i rappresentanti delle parti sociali controllano l’effettiva implementazione dei profili occupazionali previsti dal sistema, garantendo in questo modo il valore occupazionale dei titoli. A livello di settore, le parti sociali condividono con le autorità scolastiche la progettazione dei programmi di formazione e delle qualifiche, per assicurarne la congruenza con le reali esigenze delle aziende. A livello nazionale, infine, esse partecipano al processo di *policy making* che regola e gestisce il sistema nel suo complesso. Le riforme *Hartz* hanno indebolito il ruolo delle parti sociali, e in particolare delle organizzazioni sindacali, nelle politiche sociali e del lavoro, ma il loro ruolo nella formazione professionale rimane intatto.

Le modalità organizzative della partecipazione delle parti sociali possono variare a livello regionale, perché come tutto il sistema scolastico tedesco anche il sistema duale è decentrato a livello regionale: la costituzione, infatti, assegna la competenza sulla scuola ai *Länder*, le regioni che costituiscono la repubblica federale. Lo stato trasferisce i finanziamenti ai *Länder*, che ne dispongono in autonomia, e progettano altrettanto autonomamente la struttura del sistema formativo, a tutti i suoi livelli (incluso quello universitario – vedi oltre). Questo determina una variazione non trascurabile della struttura del sistema scolastico tra i diversi *Länder*.

Al di là delle variazioni organizzative regionali, è estremamente importante sottolineare che la validità dei titoli di studio è nazionale, senza alcuna possibilità di deroga per i *Länder*: questo vale per tutti i titoli scolastici, in particolare per le

qualifiche professionali rilasciate dal sistema duale. Le parti sociali e lo stato sono quindi i garanti del valore occupazionale del titolo di apprendista. In generale, l'importanza occupazionale del titolo di studio, il credenzialismo occupazionale, è una delle caratteristiche più importanti del modello socio-economico tedesco.

Con la riforma credenzialista dell'accesso alle occupazioni il "principio di competenza" sostituisce il "principio di conoscenza", secondo cui invece il reclutamento dei lavoratori avviene, invece, tramite contatti personali tra loro e il datore: questo sistema sopravvive nelle piccole organizzazioni perché scegliere persone conosciute garantisce informazioni affidabili a costo praticamente nullo. Per questo motivo le piccole aziende, che faticano a sopportare i costi informativi e non si possono permettere rischi, tendono tuttora a reclutare in base alle reti di conoscenza. Il titolo di studio può però ridurre questi costi, funzionando come "segnale", nel caso in cui esso contenga informazioni dettagliate sul tipo di competenze possedute dal lavoratore e se queste sono ritenute attendibili dal datore o dal selezionatore. È evidente come il sistema duale tedesco fornisca titoli di studio con proprio queste caratteristiche, il che ne spiega l'utilità anche per il sistema della piccola impresa, che costituisce la spina dorsale della struttura industriale tedesca.

Formazione continua, qualifiche e carriere: il Meister

L'efficacia del sistema duale tedesco si estende oltre la formazione iniziale e l'inserimento in azienda, per arrivare a riguardare anche la formazione continua e lo sviluppo delle carriere, e quindi l'organizzazione aziendale in generale. Gli studiosi parlano di quello tedesco come di un mercato del lavoro "occupazionale", perché le carriere individuali hanno tipicamente luogo non in una sola azienda (come nel caso giapponese) né in diverse aziende (come nei paesi anglosassoni), ma all'interno di un'occupazione: i lavoratori qualificati possono facilmente cambiare azienda, grazie al valore dei titoli conseguiti nel sistema duale, riconosciuto da tutte le aziende del Paese.

L'importanza del sistema duale per la regolazione del mercato del lavoro nel suo complesso, e non solo per la transizione scuola-lavoro, è un punto estremamente importante, tipicamente trascurato dagli osservatori italiani del sistema duale, che però in questo modo perdono di vista la sua capacità di produrre tra i lavoratori non solo competenze elevate, ma anche consenso e cooperazione nei confronti delle aziende. Nella maggior parte dei settori industriali e commerciali la legge tedesca richiede (in continuità con gli usi corporativi) che per poter essere

titolare di un'azienda una persona debba possedere la qualifica di *Meister*, e le aziende usano nella maggior parte dei casi la medesima qualifica per selezionare i capi intermedi, anche all'interno del proprio personale. Il percorso per diventare *Meister* costituisce il segmento superiore del sistema duale: per ottenere il titolo occorre avere superato l'esame finale da un periodo variabile tra i 3 e i 5 anni, a seconda del settore, e quindi seguire presso una *Berufsschule* (o, dove esiste, una *Berufsakademie*) un corso di specializzazione teorico-pratico che può durare, sempre a seconda del settore, tra i 4 e i 6 anni, i cui orari sono compatibili con quelli di un lavoro a tempo pieno. L'esame conclusivo è piuttosto impegnativo e può comprendere anche una prova pratica, e la commissione è costituita in modo analogo a quelle dell'esame conclusivo dell'apprendistato.

L'importanza delle qualifiche nella selezione e nelle progressioni di carriera del personale rappresenta per le aziende un vincolo esogeno, che sottrae loro parte della flessibilità in termini di libertà di assumere o portare a ruoli di responsabilità il personale. Si tratta però, per usare la terminologia della sociologia economica, di un "vincolo benefico", nella misura in cui garantisce alle aziende il consenso e l'impegno dei lavoratori. Dato che le scelte dell'azienda sono vincolate alle qualifiche, che vengono assegnate da un sistema di formazione professionale alla cui gestione i lavoratori partecipano attraverso le proprie rappresentanze, il livello di consenso dei lavoratori verso l'azienda aumenta notevolmente, perché le scelte aziendali in termini di personale non si presentano come arbitrarie, ma come sottoposte a vincoli riconosciuti e condivisi. In altri termini, la regolazione istituzionale del mercato del lavoro da un lato limita la libertà delle aziende, ma dall'altro fornisce loro il vantaggio competitivo non da poco costituito da una forza lavoro non solo tecnicamente qualificata, ma anche disponibile e motivata. Diverse ricerche che hanno messo a confronto l'organizzazione e le relazioni di lavoro delle aziende industriali tedesche con quelle francesi e inglesi hanno fornito evidenza empirica a sostegno di questo punto: i lavoratori tedeschi presentano in media tassi di assenteismo e di sciopero nettamente inferiori a quelli dei paesi in cui non esiste il sistema duale, e il loro atteggiamento nei confronti del management è molto meno conflittuale e più cooperativo.

Il lato negativo del sistema è, come ci si può immaginare, che coloro che si fermano alla scuola dell'obbligo e non proseguono con un apprendistato si trovano in difficoltà: il rischio di disoccupazione per questi giovani è molto elevato. Nel mercato del lavoro tedesco la mancanza di competenze formalmente certificate (le *skill barriers*) rappresenta un problema maggiore che altrove, e potenzialmente potrebbe dare luogo a un mercato del lavoro molto segmentato tra un

gruppo di *insider* qualificati e un gruppo di *outsider* privi di qualificazione, con forti difficoltà per i secondi a transitare tra i primi. Tuttavia le ricerche mostrano che questo effetto, che possiamo chiamare di segmentazione, è meno forte dell'opposto effetto di inclusione garantito dal modo in cui il sistema è gestito e dal valore occupazionale delle qualifiche che esso rilascia. Il sistema duale, possiamo quindi concludere, produce competenze e qualificazione tecnica, ma anche e soprattutto coesione sociale e cooperazione tra lavoratori e aziende.

5. - Un confronto con la situazione italiana

In Italia non esiste nulla di simile al sistema duale. Il modello tedesco può quindi sicuramente rappresentare un *benchmark* utile, visto che la struttura produttiva italiana è abbastanza simile a quella tedesca, in particolare per quanto riguarda il peso delle esportazioni manifatturiere. La grande differenza sta, come è noto, nella diversa dimensione media delle aziende, molto inferiore nel nostro Paese. La Tavola 2 confronta su questa variabile Italia, Germania, Francia e Spagna. La distanza tra la Germania e gli altri paesi è impressionante, e il nostro Paese è quello più lontano. Questa differenza è solitamente ritenuta una delle ragioni della debolezza della formazione tecnico-professionale italiana, ma la relazione tra i due fenomeni potrebbe essere più complessa: è probabile che la dimensione media relativamente elevata delle aziende tedesche sia favorita dal sistema duale tedesco, che rende più agevole la gestione delle grandi imprese.

TAV. 2

DIMENSIONI DELLE AZIENDE MANIFATTURIERE
IN ITALIA, GERMANIA, FRANCIA E SPAGNA

	2000			2003			2007		
	piccole	medie	grandi	piccole	medie	grandi	piccole	medie	grandi
Italia	97,9	1,8	0,3	97,8	2,0	0,3	97,8	2,0	0,3
Germania	90,9	7,2	1,9	89,9	8,0	2,1	89,5	8,4	2,1
Francia	95,7	3,4	0,9	95,7	3,2	0,8	96,4	2,9	0,8
Spagna	96,7	2,4	0,5	96,7	2,8	0,5	96,7	2,8	0,5

Fonte: elaborazione su dati EUROSTAT.

Accanto a questi fattori strutturali, senza dubbio anche fattori culturali di lungo periodo possono contribuire a spiegare la debolezza della formazione tecnica e professionale italiana in confronto a quella tedesca: nonostante il comune re-

taggio idealista, il primato della cultura umanistica su quella tecnico-scientifica è molto più forte in Italia che in Germania. Vediamo ora come si possono caratterizzare, brevemente, la formazione professionale e la transizione dalla scuola al lavoro nel nostro Paese, tenendo presente il caso tedesco come punto di riferimento.

I sistemi della formazione tecnica e professionale in Italia

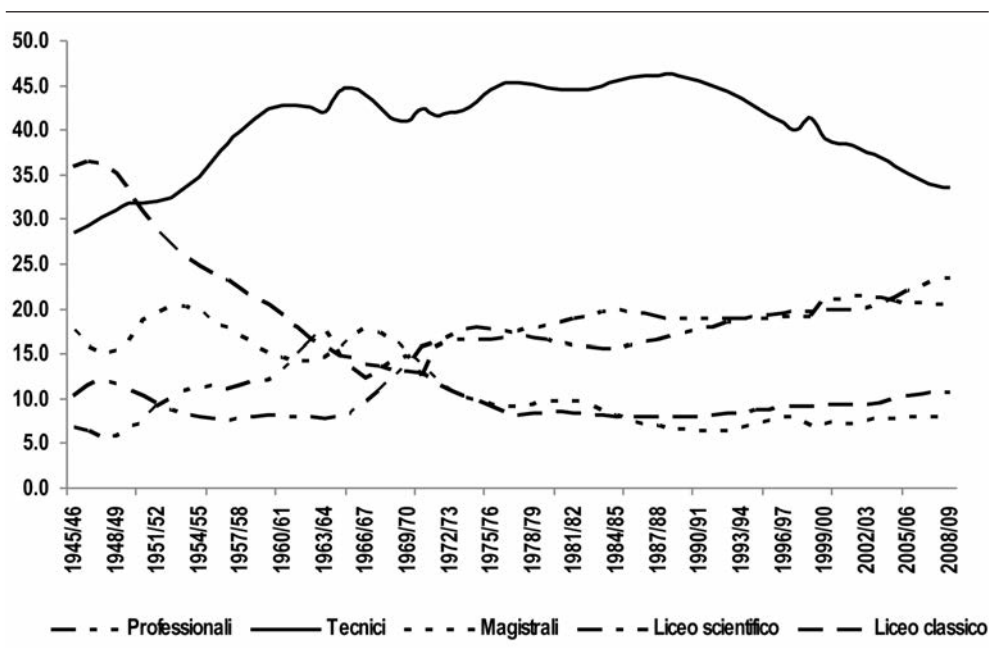
Nel nostro Paese non esiste un sistema di formazione professionale in senso proprio, ma ne esistono tre, tra loro ampiamente sovrapposti: il sistema degli istituti tecnici, quello degli istituti professionali e quello della formazione professionale regionale. I tre sistemi sono cresciuti in modo incrementale e disordinato, senza che il loro sviluppo sia stato pianificato o comunque razionalizzato a livello sistemico. Il sistema degli istituti tecnici è il più antico, nato dall'impegno degli imprenditori e dei riformatori che a partire dalla fine dell'800 hanno creato, nelle zone del paese in cui l'industria si stava più sviluppando, una serie di scuole tecniche e professionali indipendenti. A partire dall'inizio del 900 si avvia un processo di statalizzazione, intensificatosi con il fascismo, che gradualmente integra queste iniziative nel sistema scolastico nazionale.

Gli istituti professionali creati negli anni '50 erano scuole triennali rivolte a ragazzi e ragazze di origine popolare provenienti dall'indirizzo inferiore ("avviamento professionale") della scuola media precedente alla riforma che nel 1962 istituì la scuola media unitaria. Si trattava di corsi "a vicolo cieco", senza possibilità di prosecuzione, dove si poteva però conseguire un diploma molto specifico che dava accesso a ruoli tecnici di qualificazione intermedia. Solo successivamente alle riforme degli anni '60 venne abolito l'accesso differenziale all'università, condizionandolo esclusivamente alla frequenza di cinque anni di scuola post-obbligo. Gli istituti professionali, sempre nel 1969, vennero per molti versi parificati ai tecnici, in particolare con l'istituzione di un biennio aggiuntivo, che dava la possibilità di sostenere l'esame di maturità e quindi di iscriversi all'università. Gli istituti tecnici in anni più recenti sono andati incontro a un processo di graduale "liceizzazione", anche con la creazione di nuovi indirizzi: le materie accademiche acquistarono spazio nei programmi a discapito di quelle tecniche, e i rapporti con le aziende e con le professioni si indebolirono. Questo processo è visibile nel Grafico 4, che riporta le iscrizioni alle scuole secondarie superiori, in percentuale per tipo di istituto, dal secondo dopoguerra ad oggi. Nel periodo che va dagli anni '50 alla metà degli anni '70 si osserva una grande crescita delle iscrizioni agli istituti tecnici e professionali, (dove si concentra la grande espansione della scolarità

post-obbligo) che danno un contributo importante allo sviluppo industriale del paese, prima nelle grandi imprese del Triangolo industriale e quindi nelle piccole imprese dei distretti. Tra gli anni '70 e gli '80, come vediamo nel grafico, la tendenza si inverte: il peso delle iscrizioni ai licei aumenta, e quello degli istituti tecnici inizia a declinare.

GRAF. 4

ISCRIZIONI ALLA SCUOLA SECONDARIA SUPERIORE, ITALIA 1945-2009



Fonte: elaborazione su dati ISTAT (1945-2000) e MIUR (2001-2009).

Il terzo sistema è quello della formazione professionale (FP) regionale, su cui purtroppo non sono disponibili serie storiche. Questo si basa sull'articolo 117 della Costituzione, che attribuisce alle regioni la responsabilità per la FP. Con la nascita operativa delle regioni, negli anni '70, si resero disponibili le risorse con cui nacque il sistema della FP regionale che esiste tuttora, basato sui CFP (centri di formazione professionale) direttamente gestiti dalla regione (o delegati alle province) e su istituti privati finanziati dall'ente pubblico. Come accaduto in precedenza con gli istituti professionali, le nuove scuole sono state giustapposte a quelle esistenti, senza una progettazione sistemica e razionale o una chiara divisione dei compiti. Esistono eccezioni, ma la gran parte della FP regionale può essere definita

una “scuola di serie B”: il sistema opera prevalentemente con i *dropout* della scuola secondaria, il che favorisce una selezione negativa degli insegnanti, e non sono mai state svolte valutazioni sistematiche dell'efficacia occupazionale della formazione impartita.

All'inizio dello scorso decennio il ministro Moratti ha messo a punto un progetto di riforma che andava nella direzione della necessaria razionalizzazione, introducendo una chiara divisione del lavoro che riduceva a due i tre sistemi: la gran parte degli istituti tecnici e parte dei professionali sarebbero confluiti nel sistema dei licei, mentre la parte rimanente sarebbe stata conferita alle regioni, per costruire un sistema di formazione professionale che avrebbe dovuto, nelle intenzioni del ministro, approssimare il modello tedesco. Le basi regolative di questa nuova filiera sono state definite da un accordo tra ministeri e regioni (19 giugno 2003), che prevedeva l'istituzione di percorsi integrati tra scuola e FP. Tuttavia, la resistenza del mondo della scuola e le divisioni di origine politica tra governo e regioni hanno impedito la completa implementazione del progetto. Questo è però sopravvissuto a due cambi di maggioranza e di ministro, trovando una realizzazione, sia pur parziale, nella successiva riforma del 2010 che ha fatto dei passi verso la razionalizzazione del sistema: è stata formalizzata la liceizzazione di parte del sistema degli istituti tecnici, ed è stato eliminato il diploma triennale degli istituti professionali, eliminando quindi parte della sovrapposizione tra questi e la FP regionale. La riorganizzazione di quest'ultima è affidata alle regioni, che negli ultimi anni hanno scelto strade molto diverse: paradossalmente, nell'organizzazione della FP sembra esserci più variazione tra regioni in Italia, dove la repubblica è “una e indivisibile”, che in Germania, dove la repubblica è federale. Questo potrebbe comportare dei problemi dal punto di vista che qui più interessa, ovvero il valore occupazionale del titolo di studio.

La transizione scuola-lavoro

La storia della FP italiana mostra due caratteristiche tipiche del nostro Paese: una forte vitalità e iniziativa della società, e una scarsa capacità dei governi di portare a sintesi e razionalizzare quanto creato dalla società stessa. Un'osservazione simile può essere fatta guardando alla transizione dalla scuola al mercato del lavoro. In Italia il rapporto tra scuola e mondo del lavoro, e in particolare la transizione dei giovani verso le aziende alla conclusione del loro percorso scolastico, è da sempre affidato da entrambi i lati all'iniziativa individuale a livello locale, “dal basso”. Questo modello ha funzionato bene in passato, come abbiamo visto,

ma ha perso efficienza dopo gli anni '70, per via da una parte dell'indebolimento dell'attore esterno (il declino della grande industria, laddove la piccola ha molta meno capacità di investire nella scuola), e dall'altra parte del crescente orientamento delle famiglie verso i corsi di tipo accademico-liceale.

Di fronte a questa situazione, a partire dagli anni '90 diversi attori istituzionali, in particolare le associazioni datoriali, le camere di commercio e diversi enti locali, hanno cercato di rilanciare i rapporti tra istruzione tecnico-professionale e aziende, avendo molto spesso come punto di riferimento il modello tedesco. È stata promossa la creazione di reti tra istituzioni, scuole e aziende, con programmi di orientamento, stage e collaborazioni di vario genere. Non esistono ricerche nazionali sugli esiti di questi sforzi, ma la situazione della Lombardia, una delle regioni in cui questo movimento è stato più intenso, è stata studiata qualche anno fa in modo sistematico con un questionario a cui hanno risposto i dirigenti di tutte le scuole della regione.

La Tavola 3 presenta una sintesi dei risultati di questa ricerca. La gran parte delle scuole ha rapporti con aziende, e molte sono inserite in reti di scuole e *partnership* istituzionali, tra i cui obiettivi è centrale l'agevolazione della transizione scuola-lavoro. Ovunque si svolgono attività di orientamento, anche se l'orientamento all'inserimento occupazionale è meno diffuso di quello alla prosecuzione degli studi, un segno di liceizzazione. Circa tre scuole su quattro hanno rapporti con aziende, che nella gran parte dei casi sono finalizzati agli stage, la cui durata è però in media piuttosto bassa, di poco superiore alle tre settimane. Siamo ben lontani, quindi, dal modello tedesco. Tuttavia, in circa un caso su cinque lo stage si trasforma in un rapporto di lavoro, un dato molto interessante: questa trasformazione è più frequente negli istituti professionali situati in comuni con un'alta incidenza di occupazione industriale, e il cui dirigente ha pregresse esperienze lavorative al di fuori della scuola. Caratteristiche che ci sembrano rimandare al tradizionale modello distrettuale, di rapporti di tipo locale, basati sulle reti personali di insegnanti e dirigenti scolastici che fungono da *trait d'union* tra i due mondi. Tuttavia anche tale fotografia è lontana dal fornirci un quadro d'insieme, in quanto la formazione professionale resta una delle aree meno conosciute tra le attività sotto il controllo regionale, specialmente nelle regioni meridionali.

TAV. 3

RAPPORTI TRA SCUOLE TECNICO-PROFESSIONALI E AZIENDE, PER LE
CARATTERISTICHE DELLE SCUOLE, LOMBARDIA 2005

	% di scuole che fanno parte di reti di scuole	% di scuole che fanno parte di partnership istituzionali	% scuole che svolgono attività di orientamento per: studio lavoro	% scuole dove hanno luogo stage	durata media dello stage (settimane)	% alunni in stage che sono stati assunti dall'azienda	% di scuole che hanno rapporti diretti con aziende	
totale	68,6	39,4	98,0	80,4	92,5	3,4	20,1	74,1
Scuola d'arte	59,1	36,4	95,5	81,8	72,7	4,4	9,5	59,1
Magistrali	73,1	34,6	100,0	73,1	92,3	3,1	10,8	69,2
Professionali	76,3	47,4	97,4	85,1	97,4	3,6	25,1	85,1
Tecnici	64,3	35,7	98,4	78,4	92,0	3,3	18,5	69,7
Statali	73,9	44,5	98,5	84,6	96,3	3,4	20,1	77,2
Private	49,3	21,3	96,0	65,3	77,3	3,5	20,1	62,7
Piccole	56,5	25,2	95,6	70,4	80,9	3,5	21,2	68,7
Medie	74,1	41,4	99,1	84,5	97,4	3,3	18,4	74,1
Grandi	75,0	51,7	99,1	86,2	99,1	3,5	20,9	79,3

Fonte: rilevazione diretta dalle scuole, disponibile in BALLARINO G. (2008).

6. - La scuola e l'università

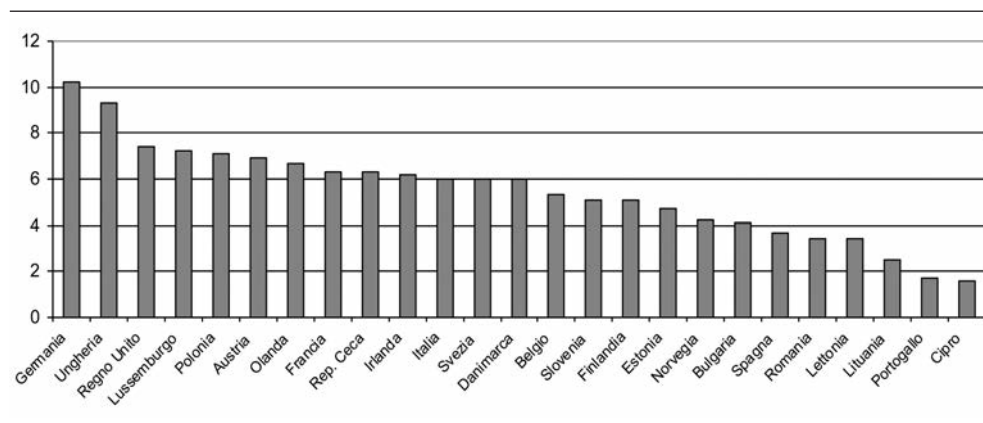
Il sistema duale è considerato in Germania parte integrante del sistema scolastico, mentre in Italia la contrapposizione tra dimensione generalista-accademica e dimensione tecnico-professionale dell'istruzione è invece molto forte. Questa contrapposizione è alle origini della mancanza, nel nostro Paese, di un'istruzione tecnica di livello post-secondario, diversa dall'università. L'istruzione terziaria tedesca prevede invece un canale tecnico-professionale (le *Fachhochschulen*), che deriva dal sistema duale e ne riproduce, sia pur con modalità organizzative e istituzionali molto diverse, la stretta integrazione con il mondo delle aziende.

Questa situazione ha evidenti effetti tanto dal lato dell'economia quanto da quello dell'istruzione. Dal lato dell'economia, il Grafico 5 mostra come il settore manifatturiero tedesco sia il più tecnologicamente avanzato d'Europa, con un'incidenza di imprese high tech di oltre il 10% sul totale, mentre l'Italia si posiziona a metà della graduatoria, con il 6%. Nell'insieme della graduatoria, si nota che a parte il Regno Unito e la Francia, gli altri paesi al vertice sono paesi la cui industria è strettamente collegata con quella tedesca (Ungheria, Polonia, Repubblica Ceca) o che presentano sistemi universitari con un canale professionale analogo alle *Fachhochschulen* tedesche (Austria, Olanda). Dal lato dell'istruzione, l'alta probabilità

di inserirsi in buone posizioni occupazionali incentiva l'iscrizione ai corsi di laurea scientifici, che infatti in Germania (e in paesi vicini geograficamente e culturalmente) sono molto più frequentati che altrove, come mostra il Grafico 6. Si nota, di contro, che il nostro Paese si posiziona al di sotto della media Ocse.

GRAF. 5

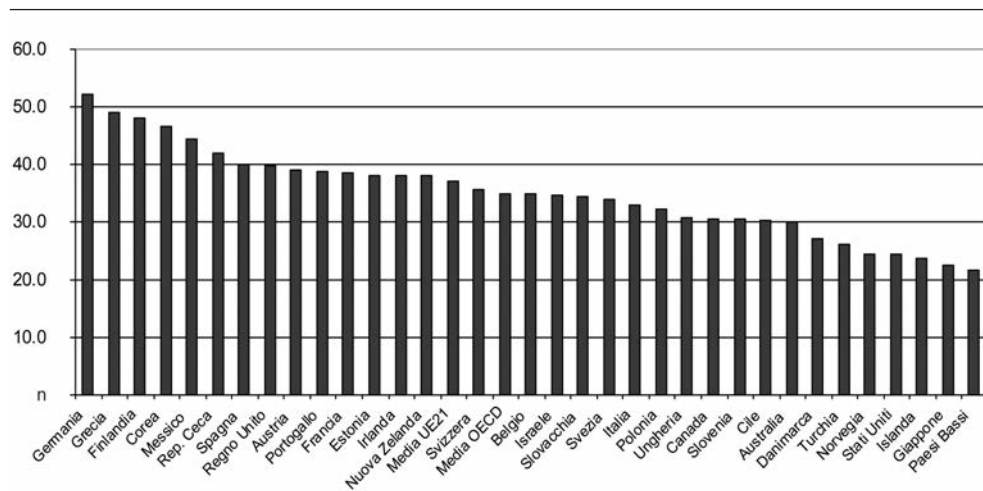
IMPRESE MANIFATTURIERE *HIGH TECH*:
PERCENTUALE SUL TOTALE IMPRESE MANIFATTURIERE - 2006



Fonte: elaborazioni CONFINDUSTRIA su dati EUROSTAT.

GRAF. 6

LAUREATI IN MATERIE TECNICHE E SCIENTIFICHE:
PERCENTUALE SUL TOTALE - 2007



Fonte: elaborazioni CONFINDUSTRIA su dati OECD.

Per prima cosa descriveremo brevemente la struttura e le modalità di selezione nella scuola secondaria tedesca, completando il quadro delineato parlando del sistema duale. Quindi passeremo al livello universitario, mettendo per prima cosa a confronto la struttura del sistema tedesco con quella italiana. Spostandoci verso l'attualità, prenderemo quindi in considerazione le diverse modalità con cui i due sistemi sono stati riformati nello scorso decennio, nel quadro comune del processo sovranazionale di riforma dell'istruzione terziaria europea, di solito definito "processo di Bologna".

La struttura della scuola secondaria

Si è detto sopra che in Germania solo la scuola elementare è unitaria (*comprehensive*), tale cioè da prevedere il medesimo curriculum per tutti. A 10 anni bambine e bambini vengono allocati su tre diversi percorsi, uno dei quali conduce all'università, mentre gli altri due conducono al sistema duale. Esiste la possibilità di passare da un percorso all'altro, seguendo le cosiddette "passerelle", ma il passaggio tradizionalmente non è molto frequente, anche se negli anni 2000 tali passerelle sono state molto rinforzate. Si tratta di una struttura simile a quella della scuola media italiana prodotta dalla riforma Gentile agli inizi del regime fascista, i cui obiettivi erano esplicitamente quelli di stabilizzare e rinforzare la stratificazione di una società intesa come organismo e non come insieme di individui liberi e uguali.

In praticamente tutti i paesi europei, a cominciare dal nostro, la scuola secondaria era originariamente organizzata in questo modo, ma negli anni '60-'70 il sistema è stato "destratificato" e reso *comprehensive*, spesso in modo più radicale che in Italia, limitando la differenziazione dei percorsi agli ultimi due anni di scuola secondaria, in funzione della transizione all'università. Negli Stati Uniti e in Canada, peraltro, la differenziazione della scuola secondaria in percorsi non esiste, ma il sistema è destratificato solo formalmente, perché le scuole secondarie sono organizzate con curricula più flessibili che in Europa, con maggiore spazio per opzioni e scelte del curriculum da parte degli studenti (una situazione che in Europa di solito si verifica solo all'università), e questa organizzazione prevede percorsi (*tracks*) differenziati, di nuovo in funzione dell'iscrizione all'università.

La selezione precoce di bambine e bambini ha importanti conseguenze per il loro futuro scolastico e lavorativo, a maggior ragione in una società credenzialistica come quella tedesca. Essa viene criticata perché a 10 anni le reali capacità cognitive di bambini e bambine sono difficilmente inferibili dalle loro prestazioni al mo-

mento, e questo rende molto difficile una selezione basata sul loro reale potenziale, diversamente dal caso in cui la selezione avesse luogo qualche anno dopo, quando le capacità cognitive sono più facilmente osservabili. Quando ha luogo precocemente, la selezione verrebbe quindi a dipendere da altri fattori, come la capacità di socializzarsi ai valori della scuola o di mantenere la disciplina, che sono ritenuti normalmente più correlati alle origini familiari che l'intelligenza e le competenze cognitive. Dunque, per riprendere la terminologia utilizzata sopra, la selezione non avverrebbe in base al principio di competenza, ma in base ad altri principi, con esiti subottimali in termini sia di efficienza economica che di equità sociale. Si tratta di un argomento piuttosto solido dal punto di vista della pedagogia e della psicologia dello sviluppo, e su questo argomento si sono basati il movimento per la scuola *comprehensive* e le politiche di destratificazione di cui si è detto. L'argomento è oggi alla base delle critiche che vengono rivolte al sistema tedesco anche da autorevoli agenzie internazionali come l'Ocse, nonché delle proposte per un'ulteriore destratificazione della scuola media superiore italiana, frequentemente avanzate anche da molti e autorevoli studiosi, sul lavoro dei quali si basava la complessiva riforma della scuola secondaria superiore progettata dal ministro Berlinguer negli anni '90 e mai completata, che avrebbe creato un biennio comune all'inizio della scuola media superiore.

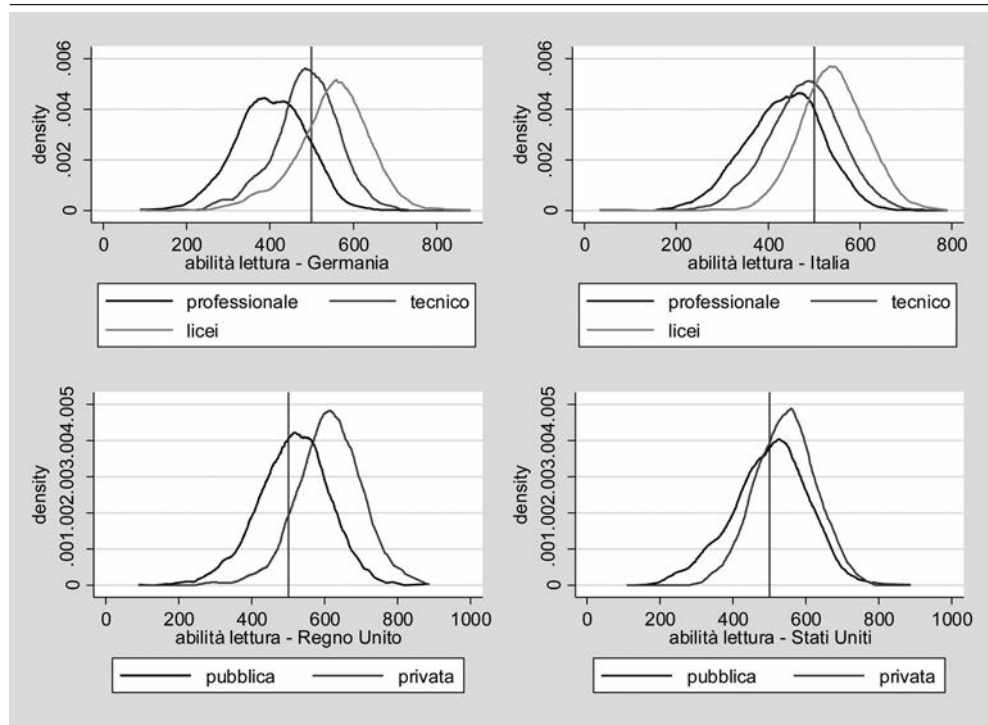
I criteri di selezione

Torniamo al confronto tra Italia e Germania. L'argomento che abbiamo appena esposto, applicato a questo confronto, porterebbe a ipotizzare che nel sistema scolastico tedesco, più differenziato di quello italiano, il peso delle origini sociali sulla scelta del percorso sia maggiore, e che quindi in Germania la scuola finisca per avere un più forte effetto di riproduzione intergenerazionale delle disuguaglianze sociali. Tuttavia, l'evidenza empirica va nel senso opposto. Abbiamo già visto sopra che in generale il livello di disuguaglianza della società tedesca è molto inferiore a quanto non sia in Italia. Come mai? Per quanto riguarda la scuola, una parte importante della risposta sta nei criteri di allocazione di bambini e bambine ai diversi percorsi. In Italia la scelta, che come sappiamo ha luogo a 14 anni, è prerogativa delle famiglie: al termine della scuola media inferiore gli insegnanti formulano delle indicazioni in merito al percorso da prendere alle superiori, ma queste indicazioni non sono in alcuna misura vincolanti, e quindi il criterio di allocazione è la decisione della famiglia, che evidentemente privilegia chi ha genitori più istruiti. I genitori delle classi superiori sono in media più istruiti, e

quindi dispongono di migliori informazioni in merito ai costi e ai benefici delle scelte scolastiche, e in loro è più forte la motivazione al mantenimento dello *status* sociale: per questo motivo, a parità di intelligenza è più probabile che una bambina di origini sociali elevate venga indirizzata dai genitori verso i percorsi liceali di quanto non accada per una bambina proveniente dalla classe operaia. Invece in Germania la scelta è prerogativa sostanziale degli insegnanti, le cui indicazioni sono nella maggior parte dei *Länder* completamente vincolanti per le famiglie. Gli insegnanti guardano in modo più oggettivo alle competenze dei loro alunni, e la loro scelta risulta quindi ispirata al principio di competenza in misura maggiore di quanto non accada quando la scelta è presa dalle famiglie. Il Grafico 7 mostra le conseguenze di questa situazione sulla distribuzione delle competenze cognitive, misurate dai *test* Pisa, sui quindicenni italiani e tedeschi, suddivisi a seconda del percorso scolastico che stanno seguendo. In Italia le tre curve, relative a licei, istituti tecnici e istituti professionali, sono tra loro molto più vicine di quanto non sia nel caso tedesco: questo significa che in Italia la quantità di liceali le cui competenze cognitive non sono distinguibili da quelle degli studenti degli istituti tecnici e professionali, rappresentata visualmente dall'area definita dall'intersezione delle curve, è maggiore che in Germania. L'allocazione dei giovani sui diversi percorsi scolastici sembra quindi essere più efficiente (e più equa) in Germania. Analogo tipo di stratificazione si ripropone nel mondo anglo-sassone tra scuole pubbliche e scuole private.

GRAF. 7

DISTRIBUZIONE DELLE COMPETENZE PER TIPO DI SCUOLA - INDAGINE PISA 2000

*L'istruzione terziaria e le Fachhochschulen*

Anche a livello di istruzione terziaria il sistema tedesco è molto più differenziato di quello italiano, e presenta un coordinamento molto più stretto con le aziende e il mercato del lavoro. Negli studi comparati sull'organizzazione istituzionale dell'istruzione superiore si distinguono di solito tre tipi di sistema, caratterizzati da un diverso grado di diversificazione delle scuole che li compongono. Ci sono in primo luogo i sistemi "unitari", composti da un solo tipo di istituzione, le tradizionali università. Ci sono poi i sistemi "binari", in cui alle università si affianca un sistema di scuole superiori di tipo tecnico-professionale, il cui prestigio è solitamente, ma non necessariamente, inferiore a quello delle università. Infine ci sono i sistemi "diversificati", che comprendono più di due tipi di istituzione. Il sistema italiano viene di norma classificato nel primo tipo: la grande maggioranza delle istituzioni che lo compongono sono università generaliste, che offrono corsi in tutte le discipline. Il sistema tedesco viene di solito definito binario, perché

la maggioranza delle scuole appartengono a uno di due tipi, le *università generaliste* e le *Fachhochschulen* (FH). L'esempio tipico di sistema diversificato è quello americano, che comprende università generaliste e specializzate, di ricerca e non; *colleges* di primo livello, di preparazione agli studi post-laurea (di secondo livello); scuole post-laurea professionalizzanti; *community colleges* (scuole locali) di primo livello, di durata biennale o quadriennale, e così via.

Nel sistema tedesco le università sono solitamente generaliste, come in Italia. Le *Fachhochschulen* (FH) sono state create all'inizio degli anni '70 riqualificando una serie di preesistenti istituti di formazione tecnica superiore (*Berufsakademien*, accademie professionali, con corsi in ingegneria, gestione aziendale, ecc.) che facevano parte del livello superiore del sistema duale di apprendistato. Il termine con cui i tedeschi traducono in inglese *Fachhochschulen* è *universities of applied sciences*, ma in Germania nessuno utilizzerebbe un'espressione di questo genere, perché le università sono chiaramente sovraordinate, in termini di prestigio, alle FH. Nella classificazione ISCED dei titoli di studio, usata nelle statistiche comparate, i diplomi delle FH sono comunque classificati come terziari (codice 5), mentre quelli delle *Berufsakademien* (che tuttora esistono nella gran parte dei *Länder*) sono classificati come "postsecondary not tertiary" (codice 4). Questa classificazione ha l'effetto di ridurre, nelle comparazioni, la percentuale di popolazione tedesca con titolo di studio terziario.

Le FH si distinguono dalle università perché presentano percorsi più brevi, didattica più orientata all'applicazione. Esse sono accessibili a 18 anni, anche con titoli secondari inferiori alla maturità, mentre nelle università si entra a 19 anni con la maturità (*Abitur*). Le FH, che sono in media più piccole delle università perché hanno un'offerta formativa più specializzata, hanno rapporti molto stretti con le aziende del territorio, sia piccole che grandi. Le materie più presenti nei programmi delle FH sono economia e business, ingegneria e architettura, scienze sociali e dell'amministrazione, e tipicamente i corsi comprendono due semestri di tirocinio (*Praktikum*). Le grandi aziende, però, hanno più spesso rapporti con università tecniche o con le facoltà tecniche ed economiche delle università generaliste. Oltre a università, FH e *Berufsakademien* esistono altri tipi di scuole superiori, specializzate nella formazioni per varie occupazioni. La distinzione tra tipi di istituto è connessa, come tipico del modello credenzialista tedesco, al valore occupazionale del titolo di studio: ogni tipo di scuola rilascia infatti titoli diversi e tra loro non sostituibili. Esistono le *Paedagogische Hochschulen* per l'insegnamento; le *Theologische Hochschulen* per i preti; le *Kunsthochschulen* per le belle arti; le *Verwaltungsfachhochschulen* per la pubblica amministrazione. Queste ul-

time (alla lettera “scuole specializzate per l’amministrazione”) sono riservate ai dipendenti pubblici, che le frequentano non solo per la formazione iniziale: i percorsi di sviluppo e di carriera nella pubblica amministrazione tedesca sono condizionati alla frequenza di specifici corsi e al superamento dei relativi esami, con un sistema molto simile, *mutatis mutandis*, a quello basato sul sistema duale che abbiamo descritto sopra per le carriere dei lavoratori dell’industria.

7. - Un confronto dei percorsi ed esiti tipici degli studenti nella carriera scolastica

Per provare a misurare la differenza degli esiti di due sistemi formativi, quello italiano e quello tedesco, abbiamo provato a ricostruire i percorsi formativi tipici nei due paesi, così per come sono sperimentati dalle generazioni più recenti. Le transizioni scolastiche possono essere studiate secondo due modalità: col metodo “per contemporanei” o con l’analisi della distribuzione dei risultati nella popolazione. Il primo metodo costruisce l’analisi partendo dai dati delle transizioni osservate in due anni contigui, e per questo ha il vantaggio della tempestività e della possibilità di una decomposizione più dettagliata (a livello territoriale o di indirizzo di scuola). Si espone però al grosso limite di non poter tener conto di ingressi/uscite dall’esterno (in quanto classifica come riduzione/aumento degli abbandoni quelle che sono entrate/uscite dall’esterno) e neppure dei rientri tardivi.

Il secondo metodo inferisce le transizioni e/o i mancati conseguimenti dalla osservazione della distribuzione dei titoli di studio conseguiti dalla popolazione (o da un suo sottoinsieme riferito a gruppi di età specifici). Esso ha il vantaggio di essere riferito allo stesso gruppo di persone che attraversa i diversi stadi della formazione, ma questo è anche il suo limite principale, in quanto quindi il sistema scolastico viene osservato in anni diversi seppur riferiti alla stessa popolazione. Poiché le informazioni provengono da indagini campionarie in cui gli intervistati vengono richiesti di indicare il loro titolo di studio più alto conseguito, spesso questo metodo è povero del dettaglio nei diversi passaggi di carriera.

Noi abbiamo deciso di utilizzare il secondo metodo perché era quello che ci garantiva la maggior comparabilità tra paesi, seppur con tutte le cautele del caso associate al comparare titoli di studio apparentemente simili per durata ma sicuramente diversi per contenuti didattici. In riferimento all’Italia abbiamo utilizzato la distribuzione dei risultati rilevati dai dati dell’Indagine della Banca d’Italia – indagine 2008, selezionando gli individui nella fascia d’età 31-40 anni. Sono individui nati tra 1968 e 1977, entrati quindi nella scuola elementare tra il 1974 e il

1983, nella scuola media tra 1979 e 1988, nella scuola superiore tra 1982 e 1991 e all'università tra 1987 e 2006. Le informazioni sui tassi di abbandono vengono invece dalla applicazione degli stessi dati stimati nel campione Isfol-Plus 2006 in riferimento alla stessa popolazione. Per quanto riguarda la Germania abbiamo utilizzato l'indagine GSOEP (German Socio-Economic Panel) nella *wave* raccolta nel 2009, sempre selezionando gli individui nella fascia d'età 31-40 anni. Sono individui nati tra 1969 e 1978, entrati quindi nella *Grundschule* (4 anni) tra il 1975 e il 1984, proseguenti nella *Orientierungsstufe* (2 anni) tra 1979 e 1988, entrati poi successivamente nei diversi indirizzi della scuola superiore tra 1981 e 1990:

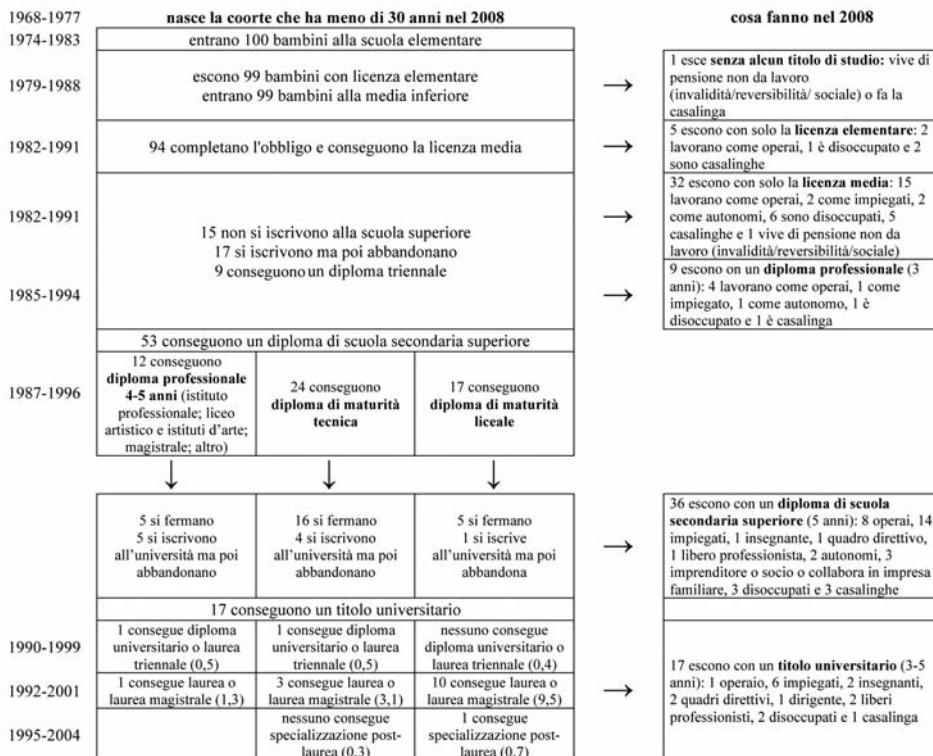
- a) *Gymnasium*+ *Gymnasiale Oberstufe*, corrispondente al nostro liceo, con durata di 4+3 anni, che si conclude con il conseguimento dell'*Abitur*, equivalente al diploma di maturità, e che dà accesso alla *Universität*;
- b) *Realschule*, corrispondente approssimativamente ai nostri istituti tecnici, con durata di 4 anni, cui fa seguito la possibilità di proseguire nelle *Fachoberschule* per ulteriori 3 anni, ottenendo poi l'accesso alla formazione postsecondaria di tipo tecnico rappresentata dalle *Fachhochschulen*, *Berufsakademien* e *Verwaltungsfachhochschulen*, di durata di 3-4 anni;
- c) *Hauptschule*, corrispondente approssimativamente al canale della formazione professionale ed apprendistato, di durata di 3 anni, con possibilità di prosecuzione nelle *Berufsfachschule* e nelle *Berufsschule + Betrieb*, di durata triennale¹.
Coloro che proseguono a livello terziario vi entrano tra il 1988 e il 2007.

Osservando i dati delle due tabelle seguenti ci si rende conto delle differenze dei sistemi scolastici tra i due paesi. Al di là dei diversi esiti occupazionali a favore dei giovani tedeschi, che può non essere necessariamente merito del sistema scolastico (27 italiani sono fuori dal mercato del lavoro, inattivi o disoccupati, contro 18 tedeschi), la cosa che colpisce maggiormente è la maggior scolarizzazione che si produce in Germania rispetto all'Italia. In Italia un terzo dei giovani conclude il proprio percorso scolastico con la licenza media (8 anni di scolarità), mentre in Germania il 44% ottiene almeno un titolo di scuola secondaria (10 anni di scolarità), a cui peraltro si aggiunge ulteriore formazione negli anni successivi (11 giovani tedeschi sono ancora in formazione).

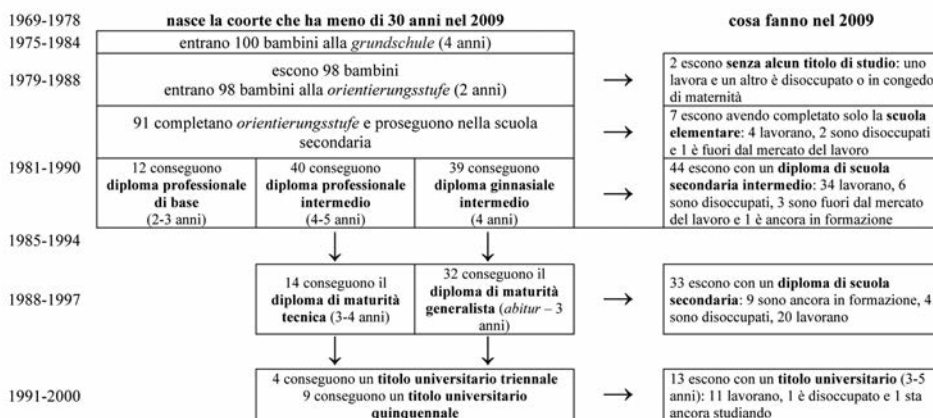
¹ La struttura del sistema scolastico tedesco è desunta dalla documentazione disponibile nel sito EURYDICE, *Strutture dei sistemi educativi europei 2011/2012*. Dato che l'ordinamento scolastico varia, come sappiamo, tra *Länder*, quello che riportiamo è un comportamento medio. In diversi *Länder* esistono forme di scuole più comprensive, intermedie tra i tre indirizzi principali (*Gesamtschulen*).

PERCORSI SCOLASTICI IN ITALIA E IN GERMANIA

Le transizioni nella scuola italiana – popolazione 31-40 anni



Le transizioni nella scuola tedesca – popolazione 31-40 anni



Fonte: elaborazione dati BANCA D'ITALIA e ISFOL-PLUS (Italia); GSOEP (Germania).

Non abbiamo dati dettagliati sugli abbandoni in corso di studi nel caso tedesco, ma quanto osserviamo dai titoli conseguiti ci permette di riconoscere che c'è una analogia fuoriuscita di persone senza completamento dell'obbligo nei due paesi (6 italiani contro 7 tedeschi), dopo di che i due paesi si differenziano: la Germania porta più giovani al conseguimento di un titolo intermedio (grazie alla maggior differenziazione dei titoli e della possibilità di ottenere certificazioni intermedie), ma meno persone al conseguimento di un titolo universitario (13 giovani tedeschi contro 17 italiani). Viceversa nel caso italiano l'esclusione dai percorsi secondari e terziari è più precoce, ma il tasso di prosecuzione finale è maggiore.

8. - Il processo di Bologna in Germania e in Italia

Abbiamo detto sopra che in Italia i governi faticano spesso a trovare una buona sintonia con la società, mentre in Germania la politica sembra avere una maggiore capacità di sintonizzarsi con la società e di accompagnarne le dinamiche. Un esempio recente di questa diversità può venire dal "processo di Bologna", la riforma pan-europea dell'istruzione terziaria. La Germania e l'Italia sono, con la Francia, i paesi che nel 1998 hanno avviato, con la dichiarazione della Sorbona, il processo di omogeneizzazione curriculare e organizzativa dei sistemi di istruzione terziaria europei, per facilitare da una parte l'integrazione tra istruzione e mercato del lavoro, e dall'altra la circolazione di individui e ricerche oltre i confini degli stati nazionali: si sarebbero così migliorate l'efficienza e l'efficacia dell'istruzione superiore.

Avviare un processo di riforma sovranazionale era un modo di aumentare la legittimità delle proposte riformiste. In Italia come in Germania, la legittimazione proveniente dal processo europeo mise a disposizione dei riformatori tutto il prestigio associato alla dimensione europea, aumentando il consenso verso le loro proposte da parte dell'opinione pubblica esterna dell'accademia. Ma in Italia questo non fu sufficiente per ottenere il consenso del mondo accademico, al punto che il governo fu indotto a concedere il decentramento locale delle modalità di reclutamento del personale universitario. Il nuovo sistema venne quindi implementato con condizioni molto favorevoli ai professori, e la coincidenza tra la sua introduzione e la creazione dei nuovi corsi a due livelli (3+2) diede ai professori la possibilità di operare una massiccia auto-promozione, non sempre basata sul principio di competenza, non dissimile per certi versi alle immissioni in ruolo *ope*

legis di 20 anni prima. Inoltre, le circostanze politiche in cui fu implementata in Italia la riforma (legge n. 509 del 1999) ne indebolirono uno dei grandi vantaggi potenziali, ovvero la possibilità di diversificare l'offerta formativa per meglio adattarla alla domanda sociale ed economica, creando le basi per un rapporto più stretto e cooperativo tra università e sistema economico. La legge imponeva infatti che il nuovo sistema venisse introdotto obbligatoriamente in tutte le università in tempi piuttosto brevi, a partire dall'anno accademico 2001/2002. Questa modalità di implementazione, centralista e *top down*, era in contraddizione rispetto al principio di autonomia introdotto dalla precedente riforma. E anche per questa ragione i nuovi trienni finirono per riprodurre molto da vicino i corsi già esistenti, perché c'era poco tempo per immaginare soluzioni innovative e perché tutte le discipline già esistenti nei curricula dovevano essere mantenute.

In Germania la riforma venne avviata dal governo federale con un emendamento alla legge quadro che regola il sistema dell'istruzione superiore (HRG) dell'agosto 1998. L'emendamento prevedeva diverse innovazioni, e soprattutto introduceva nell'ordinamento i nuovi titoli di *bachelor* e *master*, proposti dal processo di Bologna sul modello del doppio livello anglosassone. La riforma ottenne, diversamente che in Italia, il sostanziale consenso da tutte le forze politiche importanti. La legge prevedeva che i nuovi titoli venissero introdotti in forma sperimentale, accanto ai vecchi e non al loro posto, e lasciava che fossero le sedi a stabilire i tempi. Successivamente, di fronte a una certa riluttanza da parte del mondo accademico vennero gradualmente introdotti dei vincoli sulla tempistica, di modo che il processo di transizione al nuovo sistema venne realizzato nel corso di un decennio e a velocità variabile.

Inoltre, la natura federale del sistema (peraltro rinforzata proprio in questo periodo da due sentenze della corte costituzionale che, partendo da questioni specifiche, riaffermarono l'autonomia dei *Länder* nella gestione dell'istruzione superiore) consentì l'avvio di una serie di sperimentazioni istituzionali e organizzative: i ministeri dei *Länder* sono quindi diventati il motore del mutamento istituzionale dell'istruzione terziaria tedesca. Tra gli ambiti di *policy making* innovativo in cui i diversi governi regionali si sono differenziati nel corso degli ultimi anni, si possono citare:

- a) la forma giuridica delle istituzioni: alcuni *Länder*, per prima la Bassa Sassonia, l'hanno riformata in modo generalizzato, trasformando tutte le università in fondazioni di diritto privato; altri, come l'Assia, hanno trasformato la forma giuridica solo di alcuni istituti, agendo in modo selettivo; altri hanno mantenuto la forma giuridica tradizionale (analoga a quella italiana);

- b) inquadramento, retribuzione e struttura di carriera dei docenti, compresa l'organizzazione dei dottorati di ricerca;
- c) le tasse studentesche, che erano state abolite negli anni '70: alcuni *Länder* le hanno reintrodotte in modo sistematico, altri in modo selettivo (solo per fuori corso o seconde lauree), altri le hanno escluse;
- d) i programmi di incentivazione selettiva (alla ricerca, al trasferimento tecnologico, ai rapporti con il territorio), sviluppati dai diversi *Länder* anche in competizione tra loro.

È immediatamente evidente la diversa efficacia della riforma tedesca rispetto a quella italiana. La gradualità e flessibilità dell'implementazione e il decentramento del sistema hanno consentito di rinnovarlo in modo molto più efficace di quanto non sia accaduto in Italia. La riforma tedesca è durata 10 anni, e per molti aspetti è ancora in corso, mentre in Italia essa è sì stata forzosamente rapida, ma nel giro del decennio successivo sono state necessarie altre due riforme: prima la "riforma della riforma" introdotta nel 2004 dal ministro Moratti, e quindi la più recente riforma del 2011 del Ministro Gelmini.

9. - Flessibilità o beni pubblici? Una riflessione conclusiva

Verso la metà degli anni '90 sia in Italia che in Germania si discuteva molto di modelli di flessibilità e di cambiamenti della regolazione del mercato del lavoro: da questa discussione sarebbero poi nate le riforme, prima in Italia con il pacchetto Treu, quindi in Germania con i provvedimenti Hartz di cui abbiamo discusso sopra. In Germania le riforme si sono concentrate sulla flessibilità, perché la produzione dei beni pubblici necessari alla competitività economica non era in discussione: i beni pubblici in questione erano quelli di cui si è parlato in questo lavoro, in particolare la «qualificazione della forza lavoro, la cooperazione tra le parti sociali e più in generale la coesione sociale prodotta dal sistema duale di formazione professionale e dal collegamento credenzialistico tra scuola e mercato del lavoro» tipici del modello socio-economico tedesco.

Le politiche italiane degli stessi anni si sono mosse, come abbiamo visto, in una direzione diversa. Con la riforma del mercato del lavoro, la flessibilità giustamente richiesta dalle aziende è stata introdotta solo al margine, ed è quindi ricaduta prevalentemente sui nuovi entrati sul mercato del lavoro: questo ha inevitabilmente *indebolito il legame tra aziende e giovani* e disincentivato da entrambe le parti l'investimento in competenze e capitale umano, deprimendo

quindi la produzione di un bene pubblico decisivo per l'innovazione e la competitività delle aziende. Le riforme della scuola e dell'università sono state bloccate dall'eccesso di contrapposizione ideologica del sistema politico, perdendo in questo modo l'opportunità di migliorare l'efficienza e l'efficacia del sistema scolastico e di rinsaldarne il rapporto con il mercato del lavoro, beni pubblici fondamentali per imboccare la "via maestra" allo sviluppo, competendo sulla qualità nel quadro dell'economia della conoscenza.

Ora, nella situazione critica attuale, il nostro Paese dovrebbe avere la capacità di evitare gli errori commessi nel precedente ciclo riformista e non fare passare in secondo piano le riforme orientate alla produzione dei beni pubblici. In questo senso, la Germania può diventare un *benchmark* importante, a cui guardare non solo per quanto riguarda *le cose da fare*, ma anche, o forse soprattutto, per quanto riguarda *il modo in cui farle*, cioè la capacità della classe dirigente, non solo politica ma anche economica e culturale, di essere in sintonia con la società, di definire priorità condivise e di sviluppare politiche che sappiano mantenere il modello sociale europeo, e la coesione sociale che lo caratterizza, senza che questo vada a discapito dell'efficienza economica e dell'equità sociale.

10. - Per saperne di più

Sul confronto tra Italia e Germania nel contesto dell'attuale crisi economica europea e globale si è tenuto nell'estate scorsa (11-12 luglio 2011) un importante convegno economico dal titolo "Competitiveness in Europe: Confronting Italy and Germany", Roma. In particolare per questo lavoro abbiamo utilizzato le presentazioni fatte al convegno da Dullien, Fantacone e Mongeau Ospina, Möller, Niechoj, Saltari. Un altro contributo importante sul "miracolo tedesco" contemporaneo, e in particolare sulla riforma del mercato del lavoro, è quello degli economisti del lavoro Burda e Hunt (2011).

La classica comparazione tra il modello socio-economico europeo, ad alta intensità regolativa, e il modello liberista anglosassone si trova nei lavori dell'economista neo-keynesiano inglese Soskice pubblicati in Italia dalla rivista *Stato e Mercato* (1989, 2003). Una comparazione analoga, che dà più spazio al welfare state e al rapporto tra produzione e riproduzione sociale, è quella del sociologo danese-americano Esping-Andersen (2000). Un'analisi critica ma estremamente lucida delle trasformazioni in corso in Germania è quella di un altro sociologo dell'economia, Streeck (2009), secondo cui la Germania sta abbandonando l'economia sociale di mercato per avvicinarsi al modello anglosassone.

Sugli esiti della riforma del mercato del lavoro italiano esiste una letteratura ormai molto ampia e controversa. Limitandoci ai lavori empirici, una valutazione positiva della riforma è quella proposta da Bison, Rettore e Schizzerotto (2009), mentre valutazioni più critiche sono fornite da Ballarino e Barbieri (2012) e dal volume di Berton, Richiardi e Sacchi (2009).

La letteratura socio-economica sulla formazione professionale tedesca è sterminata, ma purtroppo ben poco è disponibile in italiano. Il primo lavoro che ha messo in luce l'influenza del sistema duale sul modello produttivo e sulle relazioni industriali tedesche è un contributo di un gruppo di economisti del lavoro francesi che risale agli anni '80 (ma la ricerca è di qualche anno prima): si tratta di Maurice, Sellier e Silvestre (1982). La teoria dei "vincoli benefici" è proposta da Streeck (1994). Il lavoro molto più recente di Thelen (2004), traccia la storia del sistema duale a partire dalle sue origini ottocentesche, confrontandone l'evoluzione con quella della formazione professionale in altri tre grandi paesi industriali. Gli sviluppi recenti e le criticità del sistema sono discusse in un paper della stessa Thelen e di Busemeyer (2008). La struttura del sistema duale è descritta nel database online *Eurydice* (http://eacea.ec.europa.eu/education/eurydice/index_en.php), una fonte preziosa di informazioni su tutti i sistemi scolastici europei. In italiano, la cosa migliore disponibile sul sistema duale è a tutt'oggi il capitolo sul caso tedesco del volume ricavato da un'importante ricerca comparata degli anni '90 sulla formazione delle risorse umane nelle "regioni motore" d'Europa (Bahn Müller, 1996).

Per quanto riguarda la scuola tecnico-professionale e la formazione professionale in Italia, la letteratura è in complesso lacunosa: mancano studi sistematici, sia in prospettiva socio-economica che dal punto di vista storico-istituzionale, mentre abbondano, in particolare sui media e sui numerosi siti specializzati, invece interventi prescrittivi basati su evidenza empirica debole (ricerche locali di tipo esplorativo-qualitativo) o su nessuna evidenza. Una fonte importante, soprattutto per la formazione professionale, sono i rapporti annuali dell'Isfol, scaricabili gratuitamente dal sito <http://www.isfol.it/>. Questi rapporti offrono la descrizione dell'esistente, mentre il medesimo istituto svolge anche numerose ricerche di approfondimento, i cui risultati sono pure reperibili sul sito web. Nel 2011 l'Isfol ha svolto la prima ricerca nazionale sulle carriere lavorative dei diplomati della formazione professionale triennale, i cui risultati saranno pubblicati a breve in volume. Gli esiti occupazionali e le carriere dei diplomati della scuola secondaria (sia tecnico-professionale che liceale) sono studiati dalle apposite indagini triennali dell'Istat, rapporti sulle quali sono reperibili sul sito dell'istituto (<http://www.istat.it/>). Sulle attività delle scuole non esistono ricerche nazionali. Il rapporto della ricerca

sulle scuole tecnico-professionali lombarde da cui sono tratte le informazioni riportate nella Tavola 3, è reperibile facilmente sul sito della Camera di Commercio di Milano, che ha finanziato la ricerca, inserendo in Google il nome dell'autore e il titolo. Il confronto tra le competenze dei 15enni tedeschi e italiani iscritti a diversi percorsi secondari è proposto da Checchi e Flabbi (2006).

Sulla riforma universitaria in Italia, due contributi recenti utili sono: Ballarino (2011*b*); Fondazione G. Agnelli (2012), il primo più attento alla prospettiva di lungo periodo, il secondo all'analisi degli esiti della riforma. Un confronto tra i due paesi è proposto da Ballarino (2011*a*). Il rapporto tra sistemi universitari e sistemi economici in 6 paesi europei è al centro del volume curato da Regini (2011).

La discussione su flessibilità e beni pubblici in Germania e in Italia negli anni '90 proviene da Regini (2000).

BIBLIOGRAFIA

- BAHNMÜLLER R., «Baden-Württemberg: il sistema duale di fronte alla crisi», in REGINI M. (a cura di), *La formazione delle risorse umane. Una sfida per le "regioni-motore" d'Europa*, Bologna, il Mulino, 1996, pp. 101-148.
- BALLARINO G., *La scuola tecnico-professionale lombarda e il mercato del lavoro: le iniziative delle scuole*, rapporto di ricerca per la Cciao Milano, 2008.
- .-, «Decentralization and Reform of Higher Education. Can Germany Be an Example for Italy?», *Italian Journal of Sociology of Education*, no. 2, 2011a, pages 46-71 (downloadable gratuitamente: <http://www.ijse.eu>).
- .-, «Le politiche per l'università», in ASCOLI U. (a cura di), *Il welfare in Italia*, Bologna, il Mulino, 2011b.
- BALLARINO G. - BARBIERI P., «Disuguaglianze e carriere lavorative», in CHECCHI D. (a cura di), *Disuguaglianze diverse*, Bologna, il Mulino, 2012.
- BERTON F. - RICHIARDI M. - SACCHI S., *Flex-insecurity. Perché in Italia la flessibilità diventa precarietà*, Bologna, il Mulino, 2009.
- BISON I. - RETTORE E. - SCHIZZEROTTO A., *Immobilità diffusa*, in CHECCHI D. (a cura di), Bologna, il Mulino, 2009.
- BURDA M. - HUNT J., «What Explains the German Labor Market Miracle in the Great Recession?», *IZA Working Paper*, no. 5800, 2011, disponibile su www.iza.org.
- CHECCHI D. - FLABBI L., «Mobilità intergenerazionale e decisioni scolastiche in Italia», in BALLARINO G. - CHECCHI D. (a cura di), *Sistema scolastico e disuguaglianza sociale*, Bologna, il Mulino 2006.
- ESPING-ANDERSEN G., *I fondamenti sociali delle economie postindustriali*, Bologna, il Mulino, 2000.
- FONDAZIONE G. AGNELLI, *I nuovi laureati. La riforma del 3+2 alla prova del mercato del lavoro*, Bari, Laterza, 2012.
- MAURICE M. - SELLIER F. - SILVESTRE J.J., *Politique d'éducation et organisation industrielle en France et en Allemagne*, Parigi, PUF, 1982.
- REGINI M., *Modelli di capitalismo*, Roma-Bari, Laterza, 2000.
- .-, (ed.), *European Universities and the Challenge of the Market*, Edward Elgar, Cheltenham, 2011.
- SOSKICE D., «Perché variano i tassi di disoccupazione: economia e istituzioni nei paesi industriali avanzati», *Stato e Mercato*, n. 27, 1989, pp. 333-378.
- .-, «Varietà dei capitalismi: alcuni tratti fondamentali», *Stato e Mercato*, n. 69, 2003, pp. 341-364.
- STRECK W., «Vincoli benefici: sui limiti economici dell'attore razionale», *Stato e Mercato*, n. 41, 1994, pp. 185-213.
- .-, *Re-Forming Capitalism. Institutional Change in the German Political Economy*, Oxford University Press, 2009.

THELEN K., *How Institutions Evolve. The Political Economy of Skills in Germany, Britain, the United States, and Japan*, Cambridge University Press, 2004,

THELEN K. - BUSEMEYER M.R., «From Collectivism towards Segmentalism», 2008, disponibile sul website del Max Planck Institut, www.mpifg.de.

Violare gli equilibri. Gli economisti italiani di sinistra nella crisi degli anni Settanta del Novecento

Piero Bini*
Università Roma Tre

*Between 1969 and the early nineteen-eighties a unique parabola took place in Italy that was both historical and theoretical. We are referring here to the birth, development and decline of a research project that had assumed Sraffa's book *Production of Commodities by Means of Commodities* as its basic reference. At the beginning of this debate, the characteristics of Sraffa's theory as an "open" scheme enabled it to produce a theoretical core having multiple derivations (both Keynesian and Marxian). Successively the paper considers the reasons which, since the mid-seventies, started to corrode this compound doctrinaire alliance.*

[JEL Classification: A11; B22; B51].

Keywords: Italian economic thought; Sraffa; Keynes; Marx; distributive conflict; income policy; stagflation.

* <p.bini@uniroma3.it>, Dipartimento di Scienze Politiche.

L'autore desidera ringraziare i molti studiosi che hanno avuto la gentilezza di proporgli suggerimenti e critiche in occasione dell'International Workshop "Economic Theories and Policies. A Historical Perspective (1945-2002)" svoltosi il 21-22 settembre 2011 presso l'Università di Roma Tre, dell'XI Convegno Aispe organizzato a Pisa, 1-3 dicembre 2011, e della Tavola rotonda sul pensiero di Ezio Tarantelli svoltasi presso l'Università Cattolica di Milano l'8 giugno 2012. Uno speciale riconoscimento di debito esprime nei confronti di Erik Buyst, Francesco Cattabini, Guglielmo Chiodi, Mariano D'Antonio, Antonio M. Fusco, Ivo Maes, Fabio Masini, Johan Lönnroth, Paolo Savona. La responsabilità di eventuali errori o imprecisioni è da attribuirsi esclusivamente all'autore.

1. - Introduzione

Gli anni Settanta del Novecento in Italia sono stati un decennio complesso, talvolta drammatico. Su di esso il velo della storia non si è ancora completamente posato. A seguito del crollo del sistema dei pagamenti internazionali basato sul dollaro (1971) e dell'improvviso aumento del prezzo internazionale del petrolio (1973), l'economia fu scossa nella sua matrice produttiva e deviata dal suo trend di crescita. Ne seguirono effetti negativi sul piano della disoccupazione, dell'inflazione, della finanza pubblica.

Dal punto di vista della politica, fu il decennio che vide la significativa avanzata elettorale del Partito comunista italiano (Pci) e la diffusione delle idee della sinistra. Questi venti della politica soffiarono fortemente anche sulla comunità degli economisti italiani. Molti di essi espressero varie critiche a questo o quell'aspetto del paradigma neoclassico, e mostrarono un orientamento favorevole a un maggior intervento dello stato nell'economia e a forme di programmazione e di controllo dall'alto.

Con questo scritto ci proponiamo di ricostruire questi nuovi indirizzi sia teorici che di politica economica, mettendone al tempo stesso in evidenza le articolazioni interne e la parabola evolutiva. Ci riferiremo all'insieme di queste posizioni critiche con l'espressione "paradigma della conflittualità".

Naturalmente non tutti gli economisti italiani si fecero coinvolgere in queste tendenze di pensiero. Una parte di essi, infatti, continuò a coltivare il tradizionale disegno dell'economia politica. Tuttavia, è tipico del periodo in esame che il complesso degli argomenti che ne facevano parte (come la sovranità del consumatore, la funzione positiva della proprietà privata, del risparmio, del profitto, dell'imprenditorialità, etc.) non suscitasse più il consenso di una volta, e anzi venisse omologato dalla maggioranza degli economisti di cui ci occuperemo come un pensiero obsoleto.

Abbiamo scelto l'espressione "paradigma della conflittualità" perché ci è sembrato che fosse in grado di rappresentare meglio di altre lo spirito dei tempi, e di evocare una serie significativa di elementi eterogenei, come le forti rivendicazioni salariali espresse dalle organizzazioni dei lavoratori, le vicende caratterizzate dall'antagonismo generazionale e dai movimenti studenteschi di protesta, la propensione degli intellettuali ad uscire dagli schemi di pensiero del passato, e così via.

Procederemo in questo modo. Nel secondo e nel terzo paragrafo presenteremo gli elementi di base del paradigma della conflittualità e la sua natura di prodotto scientifico composito, ovvero, come è stato detto, di "impasto dottrinario" (Gat-

tei, 1981, p. 80). Nel quarto paragrafo ci soffermeremo sui motivi che favorirono l'accoglienza dell'economia sraffiana in Italia a cavallo tra gli anni Sessanta e Settanta. Nel quinto paragrafo ricostruiremo alcuni aspetti significativi del dibattito teorico e di politica economica in Italia a metà degli anni Settanta, momento culmine del paradigma della conflittualità. Nel sesto e settimo paragrafo cercheremo di chiarire i motivi che portarono invece al suo declino, sia sul piano politico che su quello scientifico. Nell'ottavo paragrafo rivolgeremo l'attenzione a due economisti in particolare, Ezio Tarantelli e Luigi Spaventa, al cui allontanamento dal paradigma della conflittualità tra la metà degli anni Settanta e i primi anni Ottanta (dopo avervi precedentemente aderito) abbiamo attribuito il significato di una tendenza generale del pensiero economico italiano. Nelle Conclusioni svolgeremo la sintesi della nostra ricerca.

2. - La svolta conflittuale del '69 e la ricerca di un più ampio spazio teorico

Dal punto di vista dell'economia, gli anni Settanta iniziarono in Italia con il c.d. "autunno caldo" del 1969. Prese allora il via una lunga stagione di conflittualità e rivendicazioni salariali che raggiunse l'apice nel 1975. I dati seguenti, relativi agli incrementi percentuali medi annui della produttività e delle retribuzioni *pro-capite* nominali e reali nel periodo 1968-1975, mostrano una "politica dei redditi a rovescio". Il nesso tra incrementi salariali e produttività si era rotto.

TAV. 1

Anni	Retribuzioni nominali	Retribuzioni reali*	Produttività**
1968-69	6,8	4,8	5,9
1970	19,9	15,0	6,2
1971-72	10,9	5,6	3,8
1973-75	22,7	7,1	2,4

* Deflazionate con la variazione effettiva dei prezzi al consumo.

** Valore aggiunto per occupato.

Fonte: PROMETEIA e ISTAT, Tavola contenuta in ROSSI S. (2007, p. 9).

Anche nel campo delle idee economiche, il 1969 fece emergere una svolta, per quanto maturata negli anni precedenti. Ci riferiamo in particolare a due eventi: 1) lo svolgimento della X^{ma} Riunione scientifica della Società Italiana degli Economisti tenutasi a Roma dal 5 al 7 novembre 1969 (Società Italiana degli Economisti, 1974);

2) l'attuazione di un progetto di ricerca sia dell'Istituto nazionale per lo studio della congiuntura (Izzo, Pedone, Spaventa e Volpi, 1973), che del Consiglio Nazionale delle Ricerche (CNR, 1969)¹ sui problemi della politica economica di breve periodo in Italia.

La riunione della SIE del novembre 1969 ebbe come oggetto l'“Essenza e i limiti del marginalismo” e si svolse all'impronta della critica che Luigi Spaventa rivolse al neoclassicismo. Basandosi sul libro di Sraffa, *Produzione di merci a mezzo merci* (Sraffa, 1960), sottolineò l'infondatezza logica della funzione aggregata di produzione (neoclassica) perché viziata dalla relazione circolare tra misura complessiva del capitale e il suo saggio di remunerazione, il profitto. Di conseguenza, anche la teoria della distribuzione basata sulla produttività marginale dei fattori della produzione doveva essere abbandonata e, con essa – perché riconosciuta priva di generalità – la proposizione che afferma una relazione inversa tra saggio di salario reale e livello di occupazione².

Oltre a questa critica al *core* analitico del neoclassicismo, Spaventa sostenne la capacità della teoria di Sraffa di costituire la base di una nuova costruzione paradigmatica. Un punto centrale di questa teoria è che i prezzi relativi dei beni scaturiscono dalla configurazione tecnica della produzione e da una delle due variabili distributive fissata dall'esterno del modello (saggio di salario o saggio di profitto). Ciò equivale a sostenere che non esistono meccanismi endogeni di equilibrio delle variabili distributive, e che tale sistema teorico possiede – si cominciò a dire – un grado di “apertura” che la teoria neoclassica non ha. Per la “chiusura” del modello si sarebbe potuto-dovuto ricorrere ad elementi tratti da un menu di regole e scelte pubbliche non strettamente vincolate dalle relazioni strutturali dell'economia, bensì finalizzate ad obiettivi sociali ritenuti di volta in volta meritevoli di essere perseguiti sul piano politico (Spaventa, 1974, pp. 122 e ss.).

La difesa del neoclassicismo fu svolta da Del Punta (1974a). Da una parte, ne sostenne l'incisività metodologica nel rappresentare il processo economico in termini di scelte individuali svolte in condizioni di scarsità. Dall'altra, ribadì il sostanziale realismo della teoria marginalista della distribuzione:

¹ Gli economisti che vi contribuirono furono Giuseppe Campa, Augusto Graziani, Lucio Izzo, Rainer Masera, Antonio Pedone, Luigi Spaventa.

² Ciò è messo in evidenza dal manifestarsi del c.d. fenomeno del ritorno delle tecniche a cui pure si riferì in quella circostanza SPAVENTA L. (1974, pp. 112-118). Al lettore che voglia acquisire una conoscenza di base dell'opera di Sraffa, suggeriamo questi due riferimenti di carattere generale: RONCAGLIA A. (2005, pp. 435-466) e CHIODI G. e DITTA L. (2008).

«Ecco perché – dichiarò, riferendosi alle rivendicazioni salariali in corso – sono convinto che i sindacati non possano ottenere i livelli di salario che vogliono; così come non è consentito agli imprenditori di ottenere livelli di profitto *ad libitum*. Essendo pertanto anche persuaso del fatto che i saggi di remunerazione dei fattori della produzione non possano, sempre in un'economia di mercato, essere fissati esogenamente» (Del Punta, 1974*b*, p. 219).

Del Punta temeva che lo svolgimento del processo decisionale delle variabili distributive su basi socio-politiche (secondo l'esito pratico che a suo parere avrebbe prodotto lo schema sraffiano) sarebbe potuto sfociare in uno "scontro politico tra categorie e classi". Ne sarebbe derivato anche lo snaturamento in senso anti-democratico del sistema economico-istituzionale perché sempre più connotato da decisioni dall'alto, anziché delimitato dalle coordinate logiche (neoclassiche) dell'economia di mercato.

Il confronto tra Spaventa e Del Punta alimentò il clima polemico della riunione SIE del '69³. Tra i presenti vi erano, in posizione eminente, economisti della generazione più anziana. Essi rappresentavano una tradizione "classica" di studi economici per niente propensa ad abbandonare l'impianto teorico dell'equilibrio generale e gli strumenti del calcolo marginalista, e anche sostanzialmente "fredda" nei confronti della macroeconomia keynesiana⁴. Di contro, nella schiera dei giovani economisti, la critica arrivò a toccare anche la filosofia morale del marginalismo, per talune sue implicazioni pratiche mortificanti della personalità umana (si veda ad esempio Becattini, 1974).

In un giudizio di sintesi, questo dibattito accentuò le molte problematiche suscitate dai "limiti" del marginalismo, ma sicuramente non le risolse. Al tempo stesso però produsse l'effetto collaterale di accreditare l'economia sraffiana quale possibile paradigma alternativo. A questo risultato contribuì l'abilità di Spaventa nel far percepire il "grado di libertà" che questo schema presenta in merito alle variabili distributive e alla crescita, quale fondamento di una nuova teoria della politica economica; anzi di più: di una visione del mondo maggiormente ricettiva delle istanze sociali e culturali venute a maturazione a partire dagli anni Sessanta. Nuovi traguardi scientifici e importanti riconoscimenti professionali erano in vista per gli economisti disposti a partire da Sraffa.

³ CASAROSA C. (2004, p. 547) ricorda che dopo che Spaventa ebbe terminato la sua relazione, tra i presenti scoppiò "il finimondo".

⁴ Ci riferiamo ad economisti come Giovanni Demaria, Giuseppe Di Nardi, Giuseppe Palomba. Rinviemo il lettore interessato a cogliere alcune dinamiche dei processi formativi degli economisti italiani nel Novecento a FUSCO A.M. (2006).

3. - Progettare un *fine tuning* doppiamente virtuoso

L'altro evento scientifico a cui abbiamo attribuito il significato di svolta negli studi di buona parte degli economisti italiani – specie di coloro aventi un orientamento di sinistra – fu il progetto di ricerca promosso dall'Isco (Istituto nazionale per lo studio della congiuntura) nel 1968 sui problemi della “politica economica di breve periodo”. Nel corso del 1969, a questa iniziativa si aggiunse anche l'attività del “Gruppo per lo studio dei problemi della distribuzione, del progresso tecnico e dello sviluppo economico” costituitosi presso il Cnr. In entrambe queste ricerche è presente la forte insoddisfazione verso una politica economica che, per stabilizzare il ciclo e rispettare il vincolo della bilancia dei pagamenti, impiegava strumenti di controllo indiscriminato della domanda e determinava esiti deflazionistici. Questa critica conteneva tanto un giudizio negativo sulle manovre dei governi del tempo, quanto una presa di distanza teorica dalla c.d. sintesi neoclassica⁵. Data la natura di questo scritto, ci soffermeremo su questo secondo tipo di critica la quale si articolò in due motivi principali.

Il primo motivo si manifestò nella esigenza di estendere il principio keynesiano della domanda effettiva anche all'analisi dell'accumulazione e della crescita (si vedano al riguardo le argomentazioni di Fusco, 1997, pp. 103-112, e di Ciocca, 2009). I partecipanti alla ricerca Isco consideravano infatti l'obiettivo di un “elevato e regolare” saggio annuale di incremento degli investimenti (in primo luogo quelli di natura industriale) come prioritario rispetto a qualsiasi altro (si veda Izzo, Pedone, Spaventa e Volpi, 1973, p. 49). A tale riguardo, si tenga presente che la loro rappresentazione della funzione degli investimenti era del tutto in linea con il cap. XII della *General Theory* di Keynes, in quanto caratterizzata da aspettative imprenditoriali molto volatili e tali da impedire – a differenza di quanto sostenuto dagli economisti della sintesi neoclassica – il conseguimento del pieno impiego tanto nel breve quanto nel lungo periodo⁶. Per contrastare l'instabilità degli investimenti essi prospettarono l'integrazione di due tipi di strumenti, il primo co-

⁵ Come è noto – e tenendo conto del periodo storico di cui qui ci occupiamo – il *locus* classico della sintesi neoclassica è costituito dal manuale di SAMUELSON P.A. (1967).

⁶ Essi accentuarono in proposito la componente “soggettiva” delle decisioni di investimento: «Per motivi quasi sempre extra-economici, fasi di ottimismo e di fiducia nel futuro, in cui maggiore è la propensione ad assumere rischi e maggiori appaiono i benefici attesi, si alternano a periodi di pessimismo, in cui solo remunerazioni molto elevate sembrano sufficienti di fronte alle alee di un immobilizzo. Queste variazioni irregolari nelle reazioni a un dato stimolo (...) provocano spostamenti verso l'alto o verso il basso, e quindi una inerente instabilità» (IZZO L., PEDONE A., SPAVENTA L. e VOLPI F., 1973, p. 48).

stituito da una copiosa lista di misure fiscali e amministrative al fine di stabilizzare la propensione agli investimenti degli imprenditori privati; il secondo rappresentato dal sistema delle imprese pubbliche a cui si chiedeva di svolgere, oltre alle funzioni proprie di natura aziendale, anche un ruolo di compensazione anti-ciclica della domanda effettiva, nonché di realizzare un vero e proprio processo di socializzazione degli investimenti⁷.

Si trattava di un complesso disegno di politica economica a cui non era estranea una radice keynesiana “fondamentalista” (Maes, 2008)⁸. Nella saldatura che vi si prospetta tra breve e lungo periodo, questa impostazione comprendeva anche strumenti finalizzati ad affrontare taluni aspetti strutturali dell’economia italiana come lo squilibrio territoriale tra Nord e Sud, o il recupero di efficienza del settore agricolo e di quello della distribuzione commerciale. Tutti obiettivi oggettivamente giustificati dallo stato arretrato in cui tali settori si trovavano, ma anche molto ambiziosi, solo che si pensi alle modeste capacità operative del management pubblico italiano in gran parte delle sue espressioni⁹.

Il secondo motivo che testimonia la diversa prospettiva teorica degli economisti del gruppo di ricerca Isco e Cnr rispetto alla sintesi neoclassica, è costituito dal loro approccio al mercato del lavoro e alla questione salariale¹⁰. Essi respinsero il criterio di proporzionalità tra incrementi di produttività e incrementi salariali, e questo non perché non ne riconoscessero in astratto la funzionalità agli effetti della crescita economica nella stabilità dei prezzi interni. Ma perché lo ritennero una forma impropria di ingerenza pubblica nelle relazioni sindacali, e una sorta di legittimazione pseudo-scientifica di una distribuzione del reddito di lungo periodo di tipo *conservatore*.

⁷ La necessità di coinvolgere il sistema delle imprese pubbliche fu in particolare argomentata da GRAZIANI A. (1969) che era scettico sulla possibilità che nella esperienza italiana si potessero conseguire obiettivi ottimali solo grazie alla politica fiscale e monetaria.

⁸ Sul tema della diffusione di Keynes in Italia ci si permetta di rinviare anche a BINI P. e MAGLIULO A. (1999).

⁹ In merito alla funzionalità dei vari strumenti, merita di essere ricordato che gli estensori del progetto ponevano una maggiore enfasi sull’efficacia della politica fiscale piuttosto che su quella della politica monetaria, a cui si imputavano effetti asimmetrici (buona capacità di raffreddare l’economia, ma non altrettanto nello stimolarla) e indiscriminati (a causa della incidenza negativa che una restrizione monetaria produce in particolare sugli investimenti).

¹⁰ Nel manuale di SAMUELSON P.A. (1967) sono presenti vari riferimenti positivi alla politica dei redditi quale strumento ausiliario (in subordine cioè alla politica fiscale e monetaria) per elevare il livello di compatibilità tra tasso di crescita e stabilità dei prezzi.

In definitiva, pur nelle sue varianti interne, questo disegno di politica economica era fortemente orientato a valorizzare, estendere, qualificare l'intervento pubblico in economia, e anche a testimoniare una solidarietà di fondo con le rivendicazioni sindacali. L'intonazione politica era di sinistra. Ma ciò che rileva di più agli effetti della nostra ricerca, è che sulla questione salariale si istituì un collegamento tra questa proposta di politica economica e il "grado di libertà" contemplato nello schema teorico sraffiano in merito alla distribuzione, da declinarsi appunto in termini favorevoli ai lavoratori.

Siamo nel 1969. A seguito delle richieste sindacali di quell'anno, la struttura salariale conseguì (nel 1970) un primo grande balzo in avanti, ma non più di questo. Gli economisti sopra considerati non avevano e forse non potevano avere la percezione dei picchi che questa dinamica avrebbe avuto anche negli anni seguenti (vedi la Tav. 1). Le preoccupazioni in tal senso espresse dal marginalista Del Punta, di cui si è dato conto nel precedente paragrafo, non erano ancora le loro. Del resto, per essi, considerare il salario quale variabile *pro-tempore* "indipendente", non significava contemplarne aumenti arbitrari e incontrollati. Pochi anni più tardi, alcuni di questi economisti si sarebbero dovuti ricredere.

Uno tra i maggiori protagonisti di questa stagione di svolta degli studi economici in Italia fu indubbiamente Luigi Spaventa. Egli propose di ripartire dal binomio Sraffa-Keynes: il primo come rigorosa misura teorica del funzionamento di un sistema economico in grado di produrre un *surplus* (come nel sistema dell'economia classica)¹¹; il secondo come riferimento di una teoria della politica economica idonea a far acquisire agli investimenti, alla crescita e alle grandezze distributive, valori scaturenti da una analisi razionale e anche politica dell'economia italiana.

4. - Il fascino aggregante dell'economia sraffiana

Durante gli anni Sessanta, il dibattito tra le due sponde dell'Atlantico sulla teoria del capitale – meglio conosciuta come controversia tra le due Cambridge, o anche come dibattito sul ritorno delle tecniche – aveva dissodato il terreno su cui la pianta dell'economia sraffiana sarebbe cresciuta (vedi Birner, 2002). Per quanto riguarda l'Italia, una testimonianza della crescente attenzione verso di essa

¹¹ Oltre che nel campo della ricerca, SPAVENTA L. (1971) fu uno dei pochi a portare l'analisi sraffiana anche in quello della didattica.

è fornita dall'alta affluenza di giovani studiosi italiani presso la Facoltà di "Economics and Politics" dell'Università di Cambridge (UK) nel ventennio Sessanta-Settanta (vedi Ferrulli, 2012)¹². I loro principali punti di riferimento furono costituiti tanto dall'insegnamento neo-keynesiano di Joan Robinson, Nicolas Kaldor e Richard Kahn, quanto dalla giovane scuola italo-cantabrigense di Piero Sraffa, Pierangelo Garegnani e Luigi Pasinetti¹³. Una volta tornati in Italia, molti di quei giovani economisti parteciperanno al dibattito di cui qui ci occupiamo, risultando dei veicoli efficaci di trasmissione di idee economiche.

Come si è visto, l'economia sraffiana aveva fatto il suo ingresso ufficiale nel "salotto buono" degli economisti italiani nella riunione della Sie del 1969. Da allora le occasioni per approfondirne le caratteristiche metodologiche e analitiche si erano moltiplicate. Tra di esse merita un posto a sé il libro collettaneo curato da Paolo Sylos Labini, *Prezzi relativi e distribuzione del reddito*, pubblicato nel 1973 con un certo ritardo di anni rispetto alla elaborazione dei saggi ivi contenuti (Sylos Labini, 1973)¹⁴. L'interesse per questo libro si basa su vari motivi:

- 1) perché rappresenta una prima rassegna delle truppe dell'economia sraffiana tra la fine degli anni Sessanta e i primi anni Settanta (vedi Casarosa, 2004);
- 2) perché era scaturito dalle attività di ricerca dello stesso gruppo Cnr di cui abbiamo detto nel paragrafo precedente;
- 3) perché offre una prova della capacità di aggregazione dell'economia sraffiana nei confronti di economisti di vari orientamenti scientifici.

Crediamo che sia di una certa importanza soffermarci su questo terzo punto, ponendoci domande del tipo: verso quali studiosi si manifestò questa forza d'attrazione dell'economia sraffiana? Quali furono le principali motivazioni del loro interesse verso una struttura di riferimento (quella di Sraffa) così diversa da quella su cui originariamente si erano formati?

¹² Dai dati forniti nello studio citato si evince che, relativamente al periodo 1960-1974, gli economisti italiani iscritti come graduate students alla Facoltà di *Economics and Politics* di Cambridge furono 63, ovverosia il 12% di tutti gli iscritti. Le uniche presenze paragonabili a quelle italiane furono quelle degli studenti laureatisi presso lo stesso ateneo cantabrigense (pari al 25% del totale).

¹³ Come osserva GAROFALO G. (2005, p. 389), «Le influenze cambridgiane sono state trasmesse nel nostro Paese, tramite Pasinetti, all'ambiente della Cattolica di Milano e, tramite Garegnani, a Pavia e Modena».

¹⁴ Gli altri coautori furono: Pierangelo Garegnani, Luigi Pasinetti, Luigi Spaventa, Salvatore Biasco, Sebastiano Brusco, Guido Jesurum, Luca Meldolesi, Fernando Vianello, Enrico Zaghini.

Nel caso di Sylos Labini, la separazione sraffiana fra analisi dei prezzi relativi e analisi della distribuzione del reddito gli permetteva di sostenere la sua visione del processo economico caratterizzato da mercati oligopolistici (Sylos Labini, 1962) e dall'interagire in essi di variabili sociali, e di assecondare altresì il suo spirito riformatore (Sylos Labini, 1972, 2001).

In parte simili i motivi di Ezio Tarantelli. A partire dagli anni Settanta, egli si era dedicato a teorizzare il ruolo economico del sindacato dei lavoratori. Lo schema dell'economista italo-cantabrigense, là dove implica una relazione tra salario e saggio di profitto aperta a influenze esogene, rafforzava la sua idea sull'esistenza di uno spazio logico di manovra sindacale nella contrattazione salariale:

«La teoria della distribuzione del reddito più largamente accettata, almeno in Italia, è senza dubbio quella che fa capo a Sraffa (...) Questa teoria ha costituito un bagno liberatorio per tutti quegli economisti che credono, secondo me a ragione, che la distribuzione del reddito tra salari e profitti è una questione di potere relativo tra le due (e più) classi di reddito e non una variabile determinabile unicamente su basi tecnologiche» (Tarantelli, 1988c, p. 863).

In realtà, economisti come Sylos o Tarantelli non erano tanto interessati a sviluppare il modello di base di Sraffa (si veda in proposito Roncaglia, 2005), quanto a liberarsi dalla gabbia teorica della produttività marginale del lavoro. Per entrambi, inoltre, né la proprietà privata né il profitto erano istituzioni o categorie da mettersi seriamente in discussione, sebbene lo dovessero essere le modalità della loro formazione, e ciò nell'ottica delle loro proposte di cambiamenti istituzionali.

Un altro filone di convergenza verso la configurazione di Sraffa è testimoniato da quegli economisti keynesiani critici della sintesi neoclassica di cui abbiamo parlato nel paragrafo precedente. La caratteristica di sistema teorico "aperto" di questa configurazione dava coerenza alla prerogativa che essi riconoscevano al sindacato dei lavoratori di perseguire le proprie rivendicazioni salariali in piena *autonomia*. In breve, le compatibilità macroeconomiche erano da conseguire grazie al virtuoso governo delle grandezze che si trovavano sotto il controllo delle autorità pubbliche, e non di una dinamica salariale preconstituita.

Infine, dobbiamo considerare il pensiero economico marxista allora rappresentato in Italia da molti autori. Nei confronti dell'economia di Sraffa essi si divisero. Alcuni vi si opposero rilevandone la natura a-storica, e l'ulteriore distanza da Marx in ragione dei due seguenti argomenti. In primo luogo, dato che la costruzione sraffiana focalizza il conflitto tra capitalisti e lavoratori nel campo della distribuzione del reddito, si imputava ad essa di eludere la problematica dello

sfruttamento che si verifica nel processo produttivo di valorizzazione. In secondo luogo, essi osservavano che tale schema lascia indeterminata la fissazione delle quote distributive, facendola dipendere dalle alterne vicende dei conflitti di potere o da altre variabili extra-economiche. Esso non forniva dunque alcuna base scientifica a sostegno della tesi dell'*inevitabile* superamento del capitalismo¹⁵. L'anti-determinismo non allontanava l'economia di Sraffa soltanto dal marginalismo, ma anche dal marxismo ortodosso.

Invece, altri economisti di orientamento marxista sostennero un'interpretazione per certi versi opposta a quella ora considerata¹⁶. Essi apprezzarono il fatto che le carenze di analisi della teoria del valore-lavoro di Marx fossero state superate proprio grazie a Sraffa, legittimando nuovamente la ripresa dell'impostazione teorica classica e dello stesso Marx basata sulla nozione di sovrappiù. Inoltre, insieme al merito di aver dimostrato l'inconsistenza logica della teoria marginalista della distribuzione, essi attribuivano all'economia di Sraffa anche quello di aver fornito una base scientifica per sostenere di nuovo l'esistenza del rapporto conflittuale tra salari e profitti. Dagli stessi autori si procedeva poi ad ampliare il raggio della nuova costruzione paradigmatica, integrando le premesse sraffiane di natura microeconomica con quella parte del pensiero macroeconomico di Keynes che maggiormente evidenzia gli elementi contraddittori del sistema capitalistico¹⁷. Il senso di tutta questa operazione è sintetizzato nella seguente citazione:

¹⁵ Ricordiamo qui i nomi di alcuni autori delle critiche sopra sintetizzate: Domenico Mario Nuti, Mariano D'Antonio, Claudio Napoleoni, Lorenzo Calabi, Marina Bianchi, Guido Carandini, e altri ancora. Al tempo stesso occorre tener presente che il dibattito a cui ora ci si riferisce fu estremamente articolato e ricco di sfumature, dando luogo talvolta a cambiamenti di opinione da parte di taluni autori. La letteratura che ne scaturì fu ampia e a diversi livelli di impegno scientifico, concretizzandosi in monografie, raccolte di saggi, articoli su riviste scientifiche, fino a coinvolgere anche le redazioni di periodici e quotidiani di larga o limitata diffusione. Qui indichiamo solo una scelta di queste pubblicazioni: AUTORI VARI, 1974; BOTTA F., 1974; D'ANTONIO M., 1975; LAIBMAN D., 1978. Su alcuni sviluppi interni del dibattito marxista in Italia si veda GATTEI G., 1981.

¹⁶ Tra questi studiosi spiccano i nomi di Pierangelo Garegnani e di una serie ulteriore di autori, come Andrea Ginzburg, Ferdinando Vianello, Sebastiano Brusco e altri, allora facenti parte della Facoltà di Economia e Commercio dell'Università di Modena.

¹⁷ Ci si riferiva principalmente al principio della domanda effettiva e ad altri elementi (come il moltiplicatore) utili per comprendere il meccanismo della crisi. E ciò, si argomentava, in continuità con lo stesso Marx il quale, a differenza di Ricardo, «non ritenne mai che la domanda aggregata di prodotti tendesse ad adeguarsi alla capacità produttiva dell'economia» (GAREGNANI P., 1981, p. 114). Questa citazione è tratta da un articolo di Garegnani pubblicato originariamente il 4 agosto 1978 su *Rinascita*, settimanale del Partito comunista italiano, dal titolo «Sul contributo di Sraffa».

«È nostra convinzione (...) che il pensiero marxista debba appropriarsi del nucleo critico della teoria keynesiana, depurato dalle contaminazioni marginaliste; che presupposto di tale operazione sia rappresentato dai risultati teorici di Sraffa; che questi risultati, infine, vadano integralmente rivendicati alla tradizione marxista» (Ginzburg e Vianello, 1974, pp. 18-19).

Riassumendo: tra la fine degli anni Sessanta e l'inizio dei Settanta, l'opera di Sraffa catalizzò il consenso piuttosto che la dispersione delle idee. In suo nome, si formò una sorta di alleanza dottrinarie spuria che andava dagli economisti sraffiani in senso stretto, a quelli keynesiani che non aderivano alla sintesi neoclassica, a una qualificata rappresentanza di economisti marxisti, fino a includere studiosi dalla personale visione scientifica non strettamente omologabile ai suddetti filoni di pensiero, come nei casi di Sylos Labini e di Tarantelli. Questa alleanza si alimentò dell'idea che dall'incontro di correnti dottrinarie diverse si potesse ottenere un nuovo prodotto paradigmatico; e che questo servisse a sostenere scientificamente una distribuzione del reddito favorevole alla classe dei lavoratori, e un assetto sociale e istituzionale alternativo o comunque diverso dal capitalismo ad economia di mercato.

Questo stato di cose, la cui forza fu tale da marginalizzare in via provvisoria il neoclassicismo, è peraltro testimoniato da vari, ulteriori dibattiti svoltisi in Italia intorno alla metà degli anni Settanta. Tra questi, assume particolare rilievo quello che dette luogo alla c.d. "controversia Modigliani" di cui daremo sinteticamente conto nel prossimo paragrafo.

5. - La "controversia" Modigliani, la curva di Phillips in Italia, e il "paradigma della conflittualità"

La "controversia" Modigliani si verificò sullo sfondo degli effetti dello *shock* petrolifero del 1973 e dei fortissimi incrementi salariali della prima metà degli anni Settanta. Essa prese le mosse dall'applicazione in Italia, a partire dal gennaio 1975, di un meccanismo di adeguamento automatico e completo dei salari monetari all'incremento del costo della vita.

Secondo Modigliani, questo meccanismo avrebbe innescato una spirale prezzi-salari-prezzi insostenibile a lungo andare (Modigliani, 1976a e 1976b). In regime di cambi fissi, il rispetto del vincolo esterno avrebbe comportato l'adozione di misure restrittive e causato un ulteriore aumento del tasso di disoccupazione. Alternativamente, in regime di cambi flessibili, il sistema economico avrebbe assunto

un ritmo inflazionistico amplificato dalla spirale svalutazione-inflazione-svalutazione. In entrambi i sistemi di cambio, questo nuovo meccanismo della scala mobile avrebbe aggravato la crisi economica¹⁸. In conclusione, l'unica proposta per lui praticabile in quella fase di emergenza consisteva nella diminuzione dei salari reali¹⁹.

Questa indicazione fu tuttavia respinta dalla quasi totalità dei partecipanti a quella controversia²⁰. Tra questi, si segnalano in particolare Claudio Napoleoni, Federico Caffè, Antonio Pedone, Napoleone Colajanni, Augusto Graziani²¹.

Nel contesto di considerazioni a metà economiche e metà socio-politiche di quel dibattito, assunse un rilievo particolare l'elaborazione di Ezio Tarantelli. Pur condividendo gran parte dell'analisi di Modigliani, alla quale peraltro aveva contribuito come co-autore²², egli se ne allontanò in termini di indicazioni pratiche.

Secondo lui, nel corso degli anni Sessanta si era verificato un salto sociologico nelle relazioni economiche dovuto alle contestazioni espresse dalle nuove generazioni di lavoratori ad elevata scolarizzazione, frustrate e disilluse nelle loro aspettative da un sistema di relazioni sociali egemonizzato in modo conformistico dall'alleanza tra borghesia industriale e ceti medi. In Italia, a partire dal 1969, le loro espressioni di protesta, oltre che nelle Università e nelle piazze, si erano manifestate in modo particolarmente acuto anche nei luoghi di lavoro dando inizio alla stagione dell'"autunno caldo". Seguendo questa interpretazione, Tarantelli così concludeva: «L'autunno caldo non è un fenomeno di rincorsa salariale, ma uno scarto tra domanda e insufficiente risposta politica» (Tarantelli, 1978, p. 73). In altri termini, le forti rivendicazioni, oltre che ad ottenere alti salari, erano fi-

¹⁸ Modigliani non riteneva che si potesse realizzare a breve termine il correttivo costituito da un aumento significativo della produttività del lavoro, e ciò anche a causa della forte conflittualità che allora caratterizzava il sistema delle relazioni industriali in Italia. In proposito, indicò anche un motivo specifico: «il sistema attuale istituito con lo Statuto dei Lavoratori incoraggia in maniera paurosa l'assenteismo», (MODIGLIANI F., 1976*b*, p. 84).

¹⁹ Egli disse: «Il problema fondamentale e immediato da risolvere è oggi quello dell'eccessivo costo unitario del lavoro, che produce lo schiacciamento totale dei profitti, distrugge ogni incentivo agli investimenti e quindi blocca l'aumento della produttività. (...) Voglio essere anche più brutalmente franco ed esplicito: occorre abbassare il livello dei salari reali» (MODIGLIANI F., 1976*b*, pp. 82-83).

²⁰ Lo stesso MODIGLIANI F. (1999, p. 233) sintetizzò efficacemente il suo isolamento notando che: «Le accuse, soprattutto da sinistra, fioccarono come neve».

²¹ ASSO P.F. (2007) e CATTABRINI F. (2012) riportano le principali critiche di natura sia teorica che empirica rivolte da questi economisti alla ricetta Modigliani.

²² Ci riferiamo a MODIGLIANI F. e TARANTELLI E. (1973, 1975, 1977, 1979).

nalizzate a produrre una pressione sulle autorità di governo affinché attuassero un processo di riforme economiche, sociali e istituzionali. Questo punto di vista era condiviso da molti altri studiosi italiani del periodo e contribuì ad accreditare l'idea di una missione guida del sindacato dei lavoratori anche nell'ambito dei progetti politici finalizzati alla trasformazione e all'ammodernamento della società italiana. In questo clima di effervescenza ideologica, il concetto di salario quale "variabile indipendente" – derivato da uno schema teorico di radice classica – assunse anche il significato di una grandezza strumentale per determinare il cambiamento della classe di governo del paese e per subordinare ad essa l'economia di mercato.

Peraltro, nel campo degli studi economici, anche i tradizionali modelli teorici della regolazione economica furono messi in discussione. Lo schema IS/LM, ad esempio, non fu considerato errato in sé, quanto di scarsa rilevanza in una fase storica in cui la distribuzione del reddito – da tale schema non contemplata – era diventata la variabile fondamentale per comprendere le dinamiche sociali in corso e i più significativi meccanismi economici di aggiustamento.

Particolarmente interessante l'analisi che Tarantelli e Modigliani svolsero anche sulla curva di Phillips nell'esperienza italiana. A seguito del conflitto distributivo degli anni Settanta (il più acuto e di lunga durata dalla fine della seconda guerra mondiale) la curva di Phillips aveva perso la sua caratteristica di stabile relazione inversa tra tasso di disoccupazione e tasso di variazione dei salari. Questo *trade-off* si era modificato in senso sfavorevole in termini sia di disoccupazione che di inflazione, e ciò – secondo i due autori – come risultato di due *step* successivi. Il primo era costituito dallo *shock* salariale (forte differenziale di crescita del saggio di salario rispetto alla produttività *pro-capite*). Nei termini della curva di Phillips, ciò dava conto dello spostamento in alto di tutta la curva, segnalando per ciascun dato livello di disoccupazione un tasso di inflazione superiore a prima. L'origine del secondo *step* era da essi attribuito ad una (errata) politica monetaria restrittiva volta a contrastare l'inflazione da costi emersa nel primo *step*. Questo indirizzo deflazionistico, da una parte, produceva maggiore disoccupazione, ma, dall'altra, non risultava in grado di realizzare la disinflazione, dato che il suo effetto per così dire "disciplinante" risultava neutralizzato dalla forte combattività della "base operaia"²³. In conclusione, essi individuarono il fe-

²³ Per TARANTELLI E. (1978, p. 106) l'applicazione della ricetta monetarista in Italia avrebbe reso più acuto il conflitto sociale e ampliato "l'area dell'autonomia", cioè i movimenti di contestazione più radicali, allora molto diffusi nella società italiana.

nomeno della stag-flazione nel contemporaneo manifestarsi del conflitto industriale e della restrizione monetaria²⁴.

Dato il tema oggetto di questa ricerca, è da segnalare che negli scritti a firma singola di Tarantelli dedicati alla curva di Phillips, l'idea che si trattasse di una relazione sottoposta anche all'influenza di variabili extra-economiche per la cui stabilizzazione fossero soprattutto necessarie riforme politiche, risulta più netta che negli articoli a doppia firma. Tarantelli aveva infatti una forte propensione multidisciplinare. Altrettanto significative le finalità operative che egli intendeva sostenere con i suoi studi in argomento: si trattasse di comprendere come attenuare la risposta della disoccupazione alle misure monetarie della disinflazione, o di indicare un nuovo assetto istituzionale del mercato lavoro per renderlo maggiormente ricettivo delle istanze dei lavoratori (parleremo ancora di questo argomento nel par. 8).

In generale – e tornando al sodalizio scientifico tra Modigliani e Tarantelli – il loro approccio alla curva di Phillips fu di tipo keynesiano, criticando, in parallelo, le analisi di Friedman e di Phelps riguardo all'esistenza del tasso "naturale" di disoccupazione e alla sparizione della curva di Phillips nel lungo periodo²⁵.

Tuttavia, partendo da questa analisi comune, Modigliani e Tarantelli giunsero, come si è anticipato sopra, a conclusioni pratiche del tutto diverse. Modigliani cercò di svolgere una *moral suasion* nei confronti del sindacato dei lavoratori affinché accettasse il contenimento dei salari e il ritorno del sistema delle relazioni industriali a condizioni pre-autunno caldo. Tarantelli, invece, respinse questa proposta perché ritenuta antistorica: nonostante l'instabilità economica causata dalle forti rivendicazioni dei lavoratori di quegli anni, egli attribuiva ad esse il si-

²⁴ Nel diagramma di questa spiegazione in due tempi, emerge una curva di Phillips a "boomerang" che disegna il passaggio da una relazione inversa tra livelli di disoccupazione e incrementi salariali, ad una diretta. Come è ben noto, le economie dei paesi industrializzati negli anni Settanta furono principalmente caratterizzate dagli *shock* petroliferi. Così fu anche per l'Italia, dove peraltro si aggiunse, come si dice nel testo, una forte conflittualità salariale. Qui di seguito segnaliamo la serie degli incrementi percentuali annui del costo nominale del lavoro per unità di prodotto verificatisi in Italia nel decennio 1974-1983: 22,0 (1974); 34,2 (1975); 11,7 (1976); 16,6 (1977); 9,6 (1978); 13,1 (1979); 25,3 (1980); 20,0 (1981); 16,3 (1982); 12,1 (1983).

²⁵ Prendendo spunto dall'interpretazione fornita da uno dei principali studiosi della curva di Phillips a partire dagli anni Settanta – ci riferiamo a GORDON R.J. (1975, 1984, 2011) – possiamo dire che quello di Modigliani - Tarantelli fu un tempestivo contributo alla "resurrezione" di questo strumento di analisi in una fase in cui non solo i lavori di FRIEDMAN M. (1968) e PHELPS E.S. (1967 e 1968) a cui ci si riferisce nel testo, ma soprattutto quelli di Lucas sulle aspettative razionali – si veda LUCAS R.E. JR. (1972, 1973) – stavano conducendo, al contrario, a sostenere la proposizione della inefficacia della politica.

gnificato di una pressione *necessaria* al fine di cambiare lo *status quo* politico-sociale²⁶. Questo scollegamento tra analisi teorica (condivisa con Modigliani), e orientamento pratico (in opposizione a Modigliani) è un fatto che colpisce. Esso può essere compreso pienamente solo tenendo conto del rilievo assoluto che allora acquisì nel pensiero di molti intellettuali italiani – non solo in quello di Tarantelli – l’esigenza primaria della politica e dell’ideologia. Non era insolito allora sentir parlare (con implicita o esplicita adesione) dell’imminente «crollo definitivo del liberalismo» di fronte alla «crescita delle economie socialiste» (Frediani, 1974, p. 73), o dell’avvento a breve di una terza via tra economia di mercato ed economia pianificata in cui «impresa e mercato cessano di essere l’asse portante dello sviluppo economico» (Napoleoni, 1976, p. 183).

Intorno alla metà degli anni Settanta, non solo in Italia, ma anche a livello internazionale, non era insolito sentir parlare di una *crisis in economics* (Backhouse, 2010, pp. 152-163). Ma fu in Italia che l’unione di tendenze eterodosse come quelle sopra dette (riconducibili, secondo combinazioni variabili, a Sraffa, Keynes e Marx), parvero acquisire lo *status* di un nuovo assetto dottrinario, che qui abbiamo chiamato “paradigma della conflittualità”. Sebbene dai confini incerti, questo assetto alimentò dibattiti, convegni, riviste ed esercitò una forte attrattiva nei confronti dei giovani studiosi. I suoi principali esponenti furono ascoltati e ingaggiati per collaborare a periodici e quotidiani di grande tiratura, o per svolgere opera di consulenza presso istituzioni (non escluso la Banca d’Italia) e partiti politici.

Una prova di questa forza, a un tempo scientifica, culturale e ideologica, assunse anche forme negative come nel caso delle intemperanze verbali impiegate per screditare in modo talvolta sommario studiosi che non la pensavano allo stesso modo. Ci riferiamo qui ad economisti come lo stesso Modigliani o come Beniamino Andreatta (che era vicino alle posizioni dell’economista italo-americano), sminuiti come economisti “di destra”, e a cui si imputava di imprimere al pensiero keynesiano una valenza conservatrice se non addirittura autoritaria²⁷. Un tratta-

²⁶ Nella versione in lingua italiana (pubblicata su *Moneta e Credito* nel 1976) del loro articolo MODIGLIANI F. e TARANTELLI E. (1977), essendo consapevoli di questa loro divaricazione interpretativa, decisero “per espresso accordo” di non entrare sulla questione della “opportunità o meno” di “un ritorno ad un clima contrattuale pre-autunno caldo”: si veda al riguardo TARANTELLI E. (1988*b*, p. 443).

²⁷ Come nel giudizio di NAPOLEONI C. (1976, p. 114) secondo cui «il keynesismo, nato come la formulazione più avanzata della democrazia, si presenta viceversa come condizionato da un presupposto chiaramente autoritario»; oppure come in questo passo di CONVENEVOLE R. (1981, p. 431) in cui qualifica Modigliani «il nemico principale della classe operaia in questo paese (...) tra noi e lui c’è un abisso morale che ci separa».

mento non molto dissimile fu *a fortiori* riservato a quegli economisti che ancora si muovevano nel solco ortodosso (a-keynesiano o anti-keynesiano) del neoclassicismo²⁸, oppure che già coglievano nel monetarismo, come nel caso di Antonio Martino, la principale guida teorica di riferimento²⁹. Un fatto che rivela il clima ideologico del periodo è anche che il giudizio dei “conflittualisti” in merito a simili posizioni andava talvolta oltre gli aspetti teorico-tecnici, per stigmatizzarne a priori l’inadeguatezza sulla base della filosofia sociale di cui erano espressione: l’incapacità a comprendere il nuovo spirito dei tempi (ben rappresentato, ad esempio, dalla contestazione studentesca) e il carattere storicamente “progressivo” delle lotte del lavoro.

6. - Logica di emergenza e ciclo politico in un convegno economico del Partito comunista italiano

Inverno 1975-76. A seguito:

- 1) del primo *shock* petrolifero dell’autunno 1973;
- 2) della piena indicizzazione dei salari monetari al costo della vita stabilita nel gennaio 1975;
- 3) della svalutazione della lira tra la fine del ‘75 e l’inizio del ‘76,

l’economia italiana stava assumendo i caratteri di una situazione fuori controllo. Il forte aumento del costo reale del lavoro per unità di prodotto – insieme all’alto contenuto di importazioni della domanda interna, e a un eccessivo incremento della liquidità interna erroneamente perseguito nei mesi precedenti – aveva innescato (nella terminologia di Modigliani) una “spirale infernale”: inflazione, squilibrio della bilancia dei pagamenti, svalutazione del cambio, inflazione e così via³⁰.

Nel marzo 1976, in questo contesto di crisi conclamata, il Centro studi di politica economica del Partito comunista italiano (Cespe) tenne un Convegno sul

²⁸ Ci riferiamo ad un gruppo pur numeroso di economisti di varie generazioni. Solo per esemplificare: Giovanni Demaria, Epicarmo Corbino, Libero Lenti, Innocenzo Gasparini, Giuseppe Di Nardi, Tullio Bagiotti, Veniero Del Punta, Sergio Ricossa.

²⁹ Altri economisti, pur giudicando il monetarismo nel suo insieme una struttura di riferimento non soddisfacente, riconoscevano l’utilità interpretativa di alcuni suoi strumenti d’analisi (si veda, ad esempio, SAVONA P., 1981*b*).

³⁰ Su questa situazione si innestarono movimenti speculativi spinti anche da notizie allarmistiche sul “merito di credito” dell’Italia. La pressione sulla lira portò rapidamente all’esaurimento delle riserve di valuta e alla necessità di sospendere, il 20 gennaio 1976, il mercato dei cambi. La sospensione durò quaranta giorni (si veda al riguardo ROSSI S., 2007, pp. 27 e ss.).

tema «Crisi economica e condizionamenti internazionali dell'Italia» (Cespe, 1976)³¹.

Si era alla vigilia delle elezioni politiche del giugno 1976. Esse avrebbero confermato il balzo in avanti di oltre il 7% già conseguito dal Partito comunista italiano (Pci) nelle elezioni regionali del 1975. Il sistema politico centrato sul partito della Democrazia Cristiana appariva screditato. Nella opinione pubblica vi era una richiesta diffusa di moralizzazione e di maggiore efficienza nella conduzione del paese. Prendeva campo la convinzione che questi obiettivi potessero essere più agevolmente conseguiti grazie alla compartecipazione del Pci al governo del paese. Agli effetti di questo passo, pesavano tuttavia le pregiudiziali negative espresse dalla segreteria di stato USA e dalle cancellerie dei paesi europei facenti parte della Nato (Crainz, 2003, pp. 531 e ss.). Proprio al fine di allontanare questi timori internazionali, i dirigenti del Pci si prodigarono in dichiarazioni e interviste a favore sia del Patto Atlantico che dei sistemi ad economia di mercato.

Il convegno del Cespe presentò molteplici posizioni, a cominciare da quella nuovamente espressa nell'occasione da Modigliani. Pur costituendo un riferimento eccentrico rispetto ad una platea quasi totalmente orientata a sinistra, proprio alcuni suoi argomenti furono dialetticamente utilizzati dall'economista ufficiale del Pci, Eugenio Peggio, al fine di sostenere una mediazione interna al suo partito. Secondo questa linea occorre respingere «la teoria sostenuta in qualche settore della sinistra, che tendeva ad affermare il salario come “variabile indipendente” del processo produttivo» (Cespe, 1976, vol. 1, pp. 37-38); e confermare la collocazione internazionale dell'Italia anche dal punto di vista delle relazioni economiche. La conclusione di Peggio fu che:

«In generale, e nel medio termine, deve ritenersi che la dinamica del [costo del, *n.d.r.*] lavoro per unità di prodotto non possa differire sostanzialmente da quella che si verifica negli altri paesi con i quali l'Italia deve più competere. È questa la condizione necessaria per far sì che l'Italia possa continuare ad agire in un'economia aperta, e non debba fare concessioni di carattere protezionistico» (ivi, p. 38).

Al polo opposto di Modigliani si collocò Massimo Pivetti (Cespe, 1976, vol. 1, pp. 69-78). Anch'egli suonò, per così dire, su uno spartito keynesiano, ma la musica che ne ricavò fu totalmente diversa. Le sue argomentazioni possono essere così sintetizzate. Un livello insufficiente della domanda aggregata (come, secondo lui, l'eco-

³¹ Al Convegno parteciparono numerosi economisti accademici, molti dei quali iscritti o vicini al partito comunista italiano. Oltre a uomini politici del Pci, parteciparono anche alcuni esponenti di primo piano di altri partiti.

nomia italiana registrava) induce basse aspettative di profitto. Queste, a loro volta, costituiscono un incentivo alla esportazione dei capitali nella misura in cui determinano un abbassamento della propensione per gli investimenti interni. Occorreva perciò agire direttamente sul livello della domanda interna con un ulteriore incremento dei salari reali. Sugeriva poi di introdurre forme di razionamento selettivo dei beni di consumo all'interno e misure di contingentamento e altre di natura protezionistica nei confronti degli stessi beni, per ridurre drasticamente le importazioni. In breve, proponeva di attuare un radicale disegno di misure *import-substitution* al fine di abbassare il contenuto di importazioni della domanda interna, variare di conseguenza la composizione del prodotto interno lordo, e in definitiva innalzare il grado di compatibilità tra incremento dei salari e vincolo esterno. Sempre in questa prospettiva era necessario "ostacolare seriamente le esportazioni di capitale", e diminuire la dipendenza dell'Italia dall'estero e anche dal processo "dell'integrazione economica europea" (Pivetti, p. 77). Così facendo, Pivetti riteneva che il movimento operaio italiano avrebbe potuto realizzare più agevolmente il suo percorso verso l'egemonia politica, senza dover sottostare al "controllo sociale" implicito nell'uso degli strumenti della politica economica tradizionale³². Nuovamente, emergevano gli obiettivi del "paradigma della conflittualità".

Sul piano del metodo, l'analisi di Pivetti presenta un interesse ulteriore in quanto propone un collegamento esplicito tra ideologie e strumenti d'analisi, fin quasi a schematizzarne il significato in termini di categorie politiche. Semplificando: il cambio flessibile è di "destra" perché, consentendo la svalutazione, vanifica le rivendicazioni salariali dei lavoratori; il cambio fisso, insieme a *import-substitution* e controllo dei movimenti dei capitali, è invece di "sinistra", perché preserva il potere d'acquisto dei salari reali e il livello di occupazione, e così via. E ci riferiamo altresì al fatto che Pivetti considerava la sua proposta come un'alternativa globale da opporre alla politica monetaria di sostegno al capitalismo di mercato che egli attribuiva alla Banca d'Italia sotto la direzione di Guido Carli³³. Questo alludere a strumenti di analisi come mezzi di offesa o di difesa da

³² Pivetti aveva elaborato questa sua proposta già in precedenza, nell'ambito di un dibattito di economia marxista (si veda PIVETTI M., 1974).

³³ Questa critica di Pivetti era stata peraltro preceduta dall'analisi effettuata nel 1973 da GRAZIANI A. e MELONI F. (1980) sull'operato della Banca d'Italia durante gli anni Sessanta e i primi anni Settanta. I due autori vi sostenevano che le decisioni della Banca d'Italia riguardo all'inflazione e alla svalutazione del cambio erano state allora finalizzate a «restaurare l'equilibrio desiderato tra salari e profitti» e, per questa via, a mantenere «l'economia nazionale nelle mani dell'iniziativa imprenditoriale privata»: GRAZIANI A. e MELONI F. (p. 79); si veda in proposito anche NARDOZZI G. (1980).

usare nella battaglia politica, faceva parte della peculiarità di quei tempi: nell'Italia degli anni Settanta accadeva di frequente che l'analisi economica filtrasse anche nei dibattiti che si svolgevano in piccole sezioni di provincia di questo o quel partito, e ciò con la stessa intensità con la quale la passione politica entrava nelle decisioni monetarie prese nelle ovattate stanze della Banca d'Italia.

In questo dibattito in cui il vincolo esterno costituì il terreno privilegiato di scontro tra modi diversi di fronteggiarlo, tornò a farsi sentire Luigi Spaventa. Egli criticò l'opzione protezionistica di Pivetti (e di altri economisti presenti a quel Convegno, come Domenico Mario Nuti), riconfermando al contrario per l'Italia l'importanza del nesso tra crescita economica e commercio internazionale: «In questo mondo le nostre possibilità future di sviluppo dipendono dalla nostra capacità di esportare di più» (Cespe, 1976, vol. 1, p. 55). Al fine di rispettare il vincolo esterno, egli propose una "stretta fiscale" con il duplice scopo di frenare la domanda di importazioni e contestualmente ridurre il costo del lavoro ed aumentare i margini di profitto³⁴. Peraltro, Spaventa (analogamente a Modigliani, in questo) si distinse nettamente anche da quegli economisti che, come Claudio Napoleoni, finalizzavano l'azione di politica economica all'aumento della produttività da realizzarsi con una incisiva riforma dei settori c.d. parassitari (come la Pubblica amministrazione e il sistema della distribuzione commerciale). Infatti, gli effetti positivi di tale riforma non avrebbero aiutato a risolvere l'emergenza in corso, né, ancor meno, il problema costituito dal basso livello delle riserve valutarie. Il ricco ricettario del *fine tuning* dell'economia italiana, che Spaventa stesso, per primo, aveva contribuito a formulare (si veda il paragrafo 3), non era più applicabile. Per lui, la proposta di Napoleoni era "una ricetta per una cucina che non c'è più" (Cespe, 1976, vol. 1, p. 55). Si era giunti a un punto di svolta della politica economica: «Si dirà che non mi sono scostato da un ben vecchio "modello".....» (ivi).

Ai dati dell'emergenza economica che ispirarono la posizione di Spaventa, si aggiunse la logica particolare del ciclo politico che l'Italia stava attraversando, e

³⁴ Va osservato che questa formula fu almeno parzialmente realizzata negli anni immediatamente successivi. Ciò avvenne a seguito del fenomeno del *fiscal drag*, ovvero del gettito crescente che le aliquote progressive di imposta del sistema fiscale realizzarono sui valori inflazionati dei salari nominali lordi. E a seguito altresì della contemporanea decisione del governo, a partire dal 1977, di finanziare la riduzione degli oneri salariali indiretti (sociali e contributivi) a carico delle imprese. Gli effetti positivi di questa politica fiscale furono successivamente analizzati da GIAVAZZI F. e SPAVENTA L. (1989). Per una interpretazione diversa è da vedersi D'ADDA C. e SALITURO B. (1989).

che fu alla base delle posizioni espresse dai dirigenti del Pci. La loro accettazione dell'economia di mercato e di un sistema di liberi scambi internazionali non fu meramente strumentale. Ma sicuramente venne rafforzata in quella circostanza dalla specifica esigenza di ottenere, da parte dei paesi facenti parte del sistema occidentale, una patente di legittimazione per entrare nell'area di governo³⁵.

Nel Convegno del Cespe è perciò possibile individuare una battuta d'arresto, anzi un vero spartiacque, nei confronti di quel "paradigma della conflittualità" che molti economisti di sinistra avevano fino a quel momento cercato di alimentare.

7. - Il "paradigma della conflittualità" va in trincea. Il convegno di Pavia del 1978

Lo spartiacque a cui ci siamo appena riferiti è il passaggio da una situazione in cui il "paradigma della conflittualità" vantava in Italia un primato nella capacità di orientare le ricerche di economia, ad un'altra in cui la sua credibilità fu messa in discussione.

Sul piano della politica e dei rapporti sociali stavano analogamente emergendo criteri di compromesso, come il Convegno del Cespe del 1976 e le scelte del Partito comunista testimoniavano. Anche i sindacati dei lavoratori stavano evolvendo e, a partire dal 1976, avevano cominciato ad ammettere la necessità di moderare le proprie richieste. Una vera e propria svolta in tal senso fu infine annunciata dal segretario generale della Cgil (Confederazione Generale Italiana del Lavoro) Luciano Lama, in un'intervista rilasciata il 24 gennaio 1978 ad un quotidiano nazionale. In questa intervista troviamo i termini di una nuova linea di relazioni industriali:

«se vogliamo esser coerenti con l'obiettivo di far diminuire la disoccupazione, è chiaro che il miglioramento delle condizioni degli operai occupati deve passare in seconda linea (...) ci siamo resi conto che un sistema economico non sopporta variabili indipendenti (...) Ebbene, dobbiamo essere intellettualmente onesti: è stata una sciocchezza, perché in un'economia aperta le variabili sono tutte dipendenti una dall'altra» (Lama, 1978).

³⁵ Si vedano, in tal senso, gli interventi di Giorgio Amendola e Giorgio Napolitano in CESPE (1976, vol. 1, rispettivamente pp. 319-331 e pp. 257-263). Significativo anche quanto affermato da Antonio Giolitti, economista ed esponente del Partito socialista italiano, il quale rilevò esplicitamente la «professione di fede liberalizzatrice nel campo degli scambi internazionali fatta dai dirigenti del Pci da questa tribuna», cfr. Giolitti in CESPE (1976, vol. 1, p. 165).

Poche settimane dopo, a metà febbraio, si tenne una conferenza delle tre associazioni sindacali (Cgil, Cisl, Uil) al Palazzo dei Congressi dell'EUR di Roma, la quale recepiva quanto Luciano Lama aveva dichiarato nell'intervista citata (Federazione Cgil, Cisl, Uil, 1978)³⁶. Va però anche aggiunto che questi appelli al contenimento del costo del lavoro incontrarono notevoli resistenze da parte degli stessi lavoratori e delle loro organizzazioni di base che, in occasione dei rinnovi contrattuali a cavallo tra il '78 e il '79, rivendicarono aumenti sostanziosi dei livelli retributivi. Anche la situazione dell'ordine pubblico era allora molto critica.

In questo rovente clima sociale, va collocato storicamente il duro confronto (metodologico, teorico, politico) che caratterizzò il Convegno svoltosi a Pavia nel settembre del 1978 sul tema dei rapporti tra economisti e politici e fra decisioni pubbliche e teoria economica (Lunghini, 1981). Alcuni relatori segnalavano i modi attraverso i quali primari attori economici del tempo, come il Governo, la Banca d'Italia e la Confindustria, si stavano opponendo al paradigma della conflittualità (si veda Balducci e Marconi, 1981; Nardozi, 1981; Fotia, 1981). Critiche significative furono rivolte anche al sindacato unitario dei lavoratori che, dopo aver svolto nei primi anni Settanta un ruolo di rottura nel chiedere un "nuovo modello di sviluppo", riconosceva ora la necessità di ritornare a condizioni di profittabilità per le imprese e a maggiori ritmi di accumulazione (Campanella, 1981 e Valli, 1981).

Il tratto caratteristico del Convegno di Pavia, piuttosto che il conseguimento di una sintesi, fu però la frammentazione delle idee e l'emergere di numerosi contrasti fra gli stessi economisti che vi parteciparono³⁷. In vari interventi, oltre alla consapevolezza che le tesi liberiste stavano nuovamente prendendo campo (come rilevarono D'Antonio, 1981 e Vianello, 1981), si percepisce il senso di frustrazione derivante dalla constatazione delle insufficienze intrinseche al medesimo paradigma della conflittualità, come:

- 1) le difficoltà ad elaborare una nuova teoria della politica economica che ne fosse coerente espressione;
- 2) l'inadeguatezza della burocrazia e della macchina dello Stato nell'attuare con efficienza gli interventi pubblici ritenuti necessari;
- 3) la constatazione dell'incompleto processo di emancipazione della classe operaia dallo stato di subalternità nei confronti dei meccanismi di regolazione del capitalismo ad economia di mercato.

³⁶ Questo evento sindacale è meglio conosciuto come "la svolta dell'EUR".

³⁷ A questo proposito, la relazione di RAMPA L. (1981) offre interessanti considerazioni.

La relazione svolta da Augusto Graziani costituì la pietra angolare del Convegno (Graziani, 1981). Dal punto di vista teorico, egli denunciò il fatto che la proposizione secondo cui saggio di salario reale e livello di occupazione sono grandezze legate da una relazione funzionale inversa – relazione la cui fondatezza logica era stata negata dalla critica sraffiana (si rinvia in proposito al par. 2) – stava nuovamente motivando gli economisti (taluni anche di sinistra) a ragionare in termini di relazioni di compatibilità. Ciò di fatto significava subordinare la dinamica salariale ad obiettivi intermedi come la stabilità del livello dei prezzi, l'equilibrio della bilancia dei pagamenti, il conseguimento di un adeguato livello di investimenti. Ai sostenitori di questi rapporti di compatibilità – che, giusto per questo, Graziani chiamò *compatibilisti* o neo-marginalisti – egli attribuì il grave limite di non aver compreso la potenzialità euristica che il conflitto sociale potrebbe-dovrebbe assumere nella teoria economica. Ai già noti avversari di questa concezione (come Franco Modigliani, Guido Carli, Paolo Baffi), Graziani ne aggiungeva ora dei nuovi, provenienti appunto dalle fila degli economisti di sinistra, come Luigi Spaventa, Michele Salvati, Paolo Sylos Labini. Com'è facile immaginare, la sua relazione, invece che un esito chiarificatore, suscitò forti polemiche e accentuò le divisioni.

Le linee di demarcazione che si crearono in questa riunione degli “stati generali” della professione degli economisti di sinistra furono in effetti molteplici. Prima fra tutte, quella tra coloro che, a cominciare da Graziani, consideravano il compatibilismo (o neo-marginalismo) come una dottrina intrinsecamente conservatrice in fatto di salari reali; e coloro che invece rilevavano l'indeterminatezza del paradigma della conflittualità, mancante di un'analisi dei “centri di gravità” per ciò che attiene i salari e il saggio di profitto (in tal senso Salvati, 1981, pp. 422 e ss.). Come declinazione di questa contrapposizione, se ne manifestò un'altra tra quanti si fecero interpreti di una visione “idealista” del progresso scientifico in economia (cioè risultante da un processo di filiazione da idea a idea), e quanti al contrario ne sostennero una “materialista” (secondo cui sono le forze e i rapporti di produzione ad orientare in ultima istanza la ricerca scientifica). Nonostante l'impiego di un lessico da specialisti, il dibattito acquisì la caratteristica di un confronto ideologico a tutto campo, con al centro la questione della legittimazione o meno dei profitti e del capitalismo ad economia di mercato.

A questa tendenza alla drammatizzazione delle posizioni contribuì probabilmente anche un argomento di sociologia della scienza secondo cui, nel sottofondo di questa dialettica tra orientamenti diversi, si stava manifestando anche una specifica competizione per il conseguimento della *leadership* della professione italiana

degli economisti. Un testimone del tempo si è riferito implicitamente a questa motivazione ricordando certi artificiosi confronti tra quegli economisti che avevano accettato di prestare la loro collaborazione ai *policy makers* del tempo, e quanti al contrario (e tra essi, appunto, gli organizzatori del Convegno di Pavia) avevano scelto orgogliosamente di “autoproclamarsi indipendenti” (Lombardini, 1983, pp. 51-52).

Riferendosi ai “compatibilisti”, Graziani disse: “Chi si lascia guidare dalla correttezza formale opta inesorabilmente per la teoria dominante” (Graziani, 1981, p. 431). Le sue parole, per quanto signorilmente esposte, intendevano denunciare – riprendendo il tema introduttivo al Convegno svolto da Giorgio Lunghini (Lunghini, 1981) – il rischio di “strumentalizzazione inconsapevole” e di “asserimento al sistema” in cui incorrono quegli studiosi che, in nome di un’apparente neutralità degli strumenti, esauriscono le loro capacità intellettuali nella “specializzazione” e nel “tecnicismo” (Graziani, 1981, p. 431). Questo riferimento era chiaramente rivolto *anche* a quegli economisti che, pur avendo inizialmente condiviso un progetto di rinnovamento paradigmatico dell’economia con il fine del “rafforzamento concreto della classe lavoratrice”, ora se ne stavano allontanando. In termini di stile, in quel dibattito di fine anni Settanta, non mancarono giudizi sommari e momenti di virulenza polemica. Sta di fatto che su questo scisma a un tempo epistemologico, teorico e ideologico, si consumò la fine della parabola del paradigma della conflittualità.

8. - Anni Ottanta. La fuoriuscita dal paradigma della conflittualità negli studi di Ezio Tarantelli e di Luigi Spaventa

In questo ultimo paragrafo prima delle conclusioni, ci proponiamo di ricostruire – quasi fosse un test di verifica – l’evoluzione di due economisti che, dopo aver variamente contribuito all’accreditamento del paradigma della conflittualità, se ne allontanarono. Ci riferiamo a Ezio Tarantelli e a Luigi Spaventa.

Riguardo al primo, il lettore ricorderà quanto abbiamo detto a proposito della sua forte contestazione dello *status quo* politico e istituzionale italiano nel periodo di fine anni Sessanta primi anni Settanta, e ciò al punto da anteporre le ragioni di questa sua critica (nutrita politicamente e sociologicamente) ai risultati analitici conseguiti insieme a Modigliani nei loro studi congiunti sulla scala mobile e sulla curva di Phillips. Tuttavia, a partire dal 1977, Tarantelli inizia a disegnare un ruolo più propriamente teorico-economico del sindacato, meno finalizzato ideo-

logicamente³⁸. In questo passaggio verso una nuova teoria del mercato del lavoro, il suo maggior contributo deriva dall'idea che la politica dei redditi può costituire un efficace strumento di contenimento dell'inflazione a patto che le organizzazioni dei lavoratori riescano a realizzare un sufficiente grado di centralizzazione delle loro decisioni (si veda in proposito Tarantelli, 1988). La riduzione del numero dei giocatori sul mercato del lavoro avrebbe ridotto l'incertezza delle relazioni industriali, aumentando, di contro, la credibilità degli annunci antinflazionistici del sindacato medesimo. Grazie alla minore volatilità delle decisioni e delle aspettative, sarebbe stato più agevole attuare uno scambio politico tra moderazione salariale, da una parte, e riforme sociali e lotta all'inflazione, dall'altra³⁹.

È all'interno di questo schema di pensiero – dove anche l'approccio delle aspettative razionali svolge un suo ruolo – che egli fece sua ed approfondì un'idea che era stata precedentemente formulata da Paolo Savona e ulteriormente discussa all'interno del Centro Studi di Confindustria⁴⁰. La “geniale idea di Savona” – così l'avrebbe definita Guido Carli nelle sue memorie – era stata quella di “rovesciare la logica della scala mobile, legata al recupero dell'inflazione passata, e di agganciare i salari *all'inflazione programmata* futura” (Carli, 1996, p. 368): ovvero, prima si stabilisce un opportuno tasso di disinflazione relativamente al periodo preso a riferimento, dopodiché vengono applicati, nel corso del medesimo periodo, scatti di scala mobile coerenti con questo tasso precostituito.

Ezio Tarantelli, si diceva, si appropriò di questa impostazione; ne discusse all'interno di quel sindacato dei lavoratori (la Cisl) di cui era consulente; e soprattutto approfondì lo studio delle condizioni istituzionali che l'avrebbero resa possibile. Al fine di realizzare lo scambio politico tra moderazione salariale e rientro dall'inflazione, raccomandò l'attuazione di una sede tecnica e di un'impostazione politica della politica economica in virtù della quale il sindacato dei lavoratori potesse essere cooptato quale interlocutore riconosciuto nella definizione di misure riformatrici riguardanti, ad esempio, la politica fiscale, l'eliminazione di aree di rendita parassitaria, e simili (si veda in proposito Filosa e Rey, 1988).

³⁸ Sulla biografia scientifica di Tarantelli, rinviamo a MICHELAGNOLI G. (2012).

³⁹ Per dimostrare il nesso tra meccanismi istituzionali del mercato del lavoro e grado di performance macroeconomica del sistema economico, Tarantelli utilizzò il modello di decisioni in condizioni di incertezza di HEINER R. (1983). Per un approfondimento in tal senso si rinvia a GIOCOLI N. (1998).

⁴⁰ Come ricorda lo stesso Savona, questa idea era scaturita originariamente da una discussione avvenuta tra lui, Karl Brunner e Michele Fratianni a seguito di un seminario della Shadow Open Market Committee svoltosi presso l'Università di Rochester nel 1977; si veda al riguardo SAVONA P. (2008, p. 66 e 1981a, p. 228).

Riassumendo: a partire dalla fine degli anni '70 inizio anni '80, la parabola del suo pensiero rivela il passaggio dal paradigma della conflittualità (e dal salario quale variabile indipendente), ad una concezione di mercato del lavoro come luogo in cui realizzare una versione cooperativistica della dinamica sociale.

L'elaborazione di Tarantelli costituisce peraltro un caso esemplare in cui la forza delle idee economiche è riuscita a tradursi nel piano concreto delle politiche pubbliche⁴¹. In particolare, tra il 1992 e il 1993, i Governi allora in carica cominciarono a realizzare le prime forme credibili di quella concertazione triangolare, come da lui auspicata. Egli però non ebbe la possibilità di constatare l'applicazione delle sue idee in proposito. Proprio a causa di queste, era entrato nel mirino delle Brigate Rosse che lo assassinarono il 27 marzo 1985, poco dopo la fine di una sua lezione all'Università la "Sapienza" di Roma dove insegnava economia politica.

L'altro economista simbolo che abbiamo scelto per suffragare la parabola declinante del paradigma della conflittualità è Luigi Spaventa. Egli cominciò a manifestare i primi di segni di insoddisfazione verso l'economia sraffiana già nella prima metà degli anni Settanta⁴² avendo maturato che si trattava di un tipo di teorizzazione:

«according to which all we are interested in is normal long-run prices and that the rest is a kind of epiphenomenon about which we do not care. Now I think this is a very dangerous view, because it really confines economics to a very small sector» (Spaventa in Caravale, 1983, p. 194).

Del libro di Sraffa criticò in particolare l'assenza di analisi del fenomeno monetario e il fatto di ridurre il complesso processo dell'economia ad una ipostatizzazione di poche (e discutibili) ipotesi. Spaventa, che aveva inizialmente ravvisato nell'economia sraffiana un programma progressivo di ricerca, ora ne denunciava la valenza prevalentemente assolutista e regressiva. Di fronte a questa posizione, Mark Blaug – interlocutore di Spaventa in quella circostanza, e campione riconosciuto dell'antisraffismo (Blaug, 1974) – commentava così:

«Let me simply say that I was extremely pleased by practically everything that Prof. Spaventa said. I'm in agreement with almost all of it and particularly with his concluding comments on the importance of developing theories other than theories about long-run equilibrium prices» (Blaug in Caravale, 1983, p. 195).

⁴¹ Utili considerazioni al riguardo sono svolte da Bruno Chiarini nella sua Introduzione a TARANTELLI E. (1995).

⁴² Ciò risulta da alcuni riferimenti espressi da lui stesso in un dibattito con Mark Blaug svoltosi il 5 maggio 1976 presso l'Università di Perugia: si veda CARVALE G. (1983, pp. 182-210).

Attraverso studi successivi Spaventa ebbe l'occasione di motivare ulteriormente il suo distacco dall'economia di Sraffa, ritenuta responsabile di aver respinto *in toto* l'intera economia neoclassica, buttando via così «insieme a tanta acqua sporca, alcune creature utili e vitali per l'analisi» (Spaventa, 1982, p. 1056).

Molti anni dopo, Spaventa ha perfino accentuato questo suo parere negativo segnalando il grande dispendio di energie scientifiche che il tentativo sraffiano aveva comportato: «la critica finì per esaurirsi in se stessa...» (Spaventa, 2004, p. 575); e il fatto collegato di una tendenza all'isolamento del dibattito italiano rispetto a quello internazionale⁴³. Il suo ulteriore giudizio che in quegli anni vi sia stata in Italia una *lost generation* nella ricerca teorica è da ritenersi eccessivo, solo che si consideri come le strade della conoscenza siano a tal punto imperscrutabili e tortuose da non consentire bilanci precisi se non dopo un lungo tratto di tempo. Ma da questa nostra ricerca abbiamo ricavato gli elementi per sostenere che, nel periodo esaminato, si manifestò effettivamente, da parte di molti economisti italiani di sinistra, un atteggiamento di resistenza ad esplorare i nuovi territori della teoria, come il monetarismo, il modo in cui gli agenti formano le aspettative, il concetto di incoerenza temporale nella politica macroeconomica, etc.: un atteggiamento che, come la storia del pensiero economico insegna, indubbiamente non favorisce in generale il progresso scientifico⁴⁴.

9. - Conclusioni

Tra il 1969 e i primi anni Ottanta si svolse in Italia una singolare parabola, storica e teorica al tempo stesso. Ci riferiamo alla nascita, allo sviluppo e al declino di un "progetto" di ricerca che aveva assunto come riferimento di base il libro di Sraffa *Produzione di merci a mezzo merci*. Dopo che anche il dibattito internazio-

⁴³ Ricostruendo le attività svolte a cavallo tra gli anni Sessanta e Settanta dal "Gruppo CNR per lo studio dei problemi economici della distribuzione, del progresso tecnico e dello sviluppo", SPAVENTA L. (2004, p. 576) ne rivela le principali tendenze di ricerca e anche le distorsioni: «Pare oggi sconcertante non trovare traccia nelle discussioni teoriche del Gruppo delle questioni che si dibattevano oltre oceano e delle tendenze che vi si affermavano [si riferisce in particolare alla "nuova macroeconomia neoclassica", N.d.R.]: si continuava a combattere su un fronte oramai secondario nel dibattito corrente, mentre le armate avversarie dilagavano sul fronte keynesiano, lasciato indifeso».

⁴⁴ Un interessante studio su queste dinamiche della ricerca economica in Italia tra gli anni Settanta e Novanta del Novecento è stato effettuato da ASSO P.F. (2004) che ha preso a riferimento le pubblicazioni della *Rivista di Politica Economica*.

nale ne ebbe verificato la solidità logica, quest'opera vide accrescere il proprio spazio di manovra teorica, fino ad abbozzare i caratteri di una nuova costruzione paradigmatica. L'avvento di una nuova filosofia sociale, giusto a partire dagli anni Sessanta, contribuì a sostenere questa impresa scientifica fornendole un appropriato retroterra culturale e politico. Le sue caratteristiche di progetto teorico "aperto" le consentirono poi di svolgere un'azione aggregante nei confronti di varie correnti di pensiero economico.

Partendo da questi sviluppi, abbiamo individuato le ragioni del maggior radicamento della scuola sraffiana in Italia rispetto ad altri paesi. Oltre al pregio costituito dalla sua coerenza interna, essa risultò favorita da circostanze per così dire esterne. La prima è che questa scuola presentava forti elementi biografici di italianità. I suoi maggiori esponenti del tempo, oltre a Piero Sraffa, furono infatti Pierangelo Garegnani e Luigi L. Pasinetti. La seconda circostanza attiene alla presenza in Italia di un'ampia opinione pubblica influenzata dalle idee della sinistra e dalle scelte del Partito comunista italiano, ciò che favorì l'incontro del dibattito sraffiano con quello riguardante il destino del pensiero di Marx (Garofalo, 2005, p. 389).

Nel sistema di sinergie scientifiche che venne allora a determinarsi intorno allo schema sraffiano, un posto di rilievo spetta al dibattito riguardante il salario quale "variabile indipendente". Attraverso di esso si realizzò la saldatura tra uno dei motivi teorici di tale schema e l'aspirazione per un nuovo tipo di società. In questa rappresentazione, il conflitto distributivo trascese il significato di mero confronto tra domanda e offerta di prestazioni lavorative, per diventare rivendicazione anche di obiettivi di progresso sociale e civile, o perfino riflesso di istanze politiche radicali. Ne scaturì una concezione meta-teorica del salario come prezzo regolato prevalentemente da forze sociali e da rapporti di potere. Abbiamo chiamato questo complesso programma di ricerca "paradigma della conflittualità".

Nel corso di questo studio ci siamo poi soffermati sui motivi che a partire dalla metà degli anni Settanta cominciarono a corrodere questa composita alleanza dottrinaria. L'aggravarsi della crisi economica dell'Italia a metà anni Settanta – una "economia sotto assedio", fu definita⁴⁵ – certamente indebolì l'accreditamento di questo "paradigma". Inoltre, il maggiore sponsor politico di quel tentativo di rinnovamento scientifico, il Partito comunista, trovandosi (nel 1976) alla vigilia del suo ingresso nella maggioranza parlamentare, valutò come prioritario rompere quella spirale inflazione-svalutazione-inflazione che la politica salariale fin lì per-

⁴⁵ L'espressione è di Paolo Baffi, Governatore della Banca d'Italia dal 1975 al 1979.

seguita aveva contribuito ad alimentare. Gradualmente, poi, nelle relazioni fortemente conflittuali che allora intercorrevano tra i sindacati dei lavoratori e le associazioni padronali, cominciarono ad affiorare i primi elementi di distensione. La “svolta dell’Eur” non sarebbe potuta avvenire se non ci fosse stato un processo di avvicinamento tra i dirigenti delle due controparti del mercato del lavoro⁴⁶. A seguito di queste nuove dinamiche, nella stessa cultura economica della sinistra si produsse una frattura, di cui questo scritto ha cercato di cogliere alcuni aspetti.

Il “grado di libertà” che lo schema sraffiano contempla nell’analisi della distribuzione del reddito, accolto in una prima fase come punto di forza teorico e metodologico, fu avvertito successivamente come una caratteristica che privava lo studio dell’economia del solido ancoraggio costituito dall’analisi dell’equilibrio. Analogamente, la riflessione sulla necessità di integrare le variabili sociali nel lavoro dell’economista – una tendenza che il “paradigma della conflittualità” aveva incoraggiato – cominciò a risultare sospetta, come una modalità eccentrica, per non dire contraddittoria, di implementare il rigore logico dell’analisi economica. Dal punto di vista del dibattito tra economisti, il Convegno di Pavia (di cui al par. 7) costituì il tentativo di contrastare il ritorno di questi atteggiamenti epistemologici tipici dello schema dell’equilibrio generale. Ma esso non riuscì a invertire il senso del declino di quella esperienza scientifica e intellettuale.

Infine, riferendoci a due vicende particolari – quella di Ezio Tarantelli e di Luigi Spaventa – abbiamo esemplificato i percorsi personali di quanti, inizialmente tratti dal paradigma della conflittualità, poi se ne allontanarono. L’esperienza di studio di questi due autori testimonia che la maggiore *impasse* in cui venne a trovarsi questo assetto composito di idee fu costituita dalla sua difficoltà ad interpretare i processi reali dell’economia e, forse di più, a prospettare politiche e assetti istituzionali prontamente e efficacemente adattabili alle sfide di quei tempi.

Quest’ultimo giudizio ci riporta al tema generale di questa ricerca, di cui qui di seguito svolgiamo la sintesi conclusiva.

Iniziamo ricordando che il libro di Sraffa, *Produzione di merci a mezzo merci* offre una configurazione *pura* di relazioni economiche tale da poterne consentire un’analisi oggettiva indipendentemente da qualsiasi dimensione spazio-temporale.

⁴⁶ Per esemplificare, tra il Presidente della Confindustria del tempo, Guido Carli, e il segretario generale della Cgil, Luciano Lama, si instaurò un rapporto di stima che favorì questo avvicinamento. E ancora: l’idea di depotenziare il meccanismo della scala mobile sulla base di tassi decrescenti di inflazione programmata – idea che era circolata inizialmente all’interno del Centro Studi di Confindustria – influenzò anche il punto di vista di alcuni ambienti del sindacato dei lavoratori.

Ai fini della determinazione delle grandezze di un sistema economico *concreto*, questo schema di teoria pura richiede perciò di essere integrato con un secondo stadio di analisi caratterizzato da ipotesi riguardanti i reali comportamenti sociali e individuali⁴⁷.

Ebbene, nella specifica circostanza del dibattito tra economisti italiani avvenuto tra la fine degli anni Sessanta e l'inizio degli anni Settanta, la ricerca su questo secondo e più complesso stadio di analisi si caratterizzò con forti accenti ideologici di sinistra. Si voleva perseguire un assetto politico e distributivo favorevole ai lavoratori e ai salari, e un contesto di regole e di interventi pubblici finalizzati al contenimento, per non dire snaturamento, dell'economia di mercato al fine di realizzare forme di pianificazione economica. Questa declinazione accreditò un'interpretazione della teoria sraffiana (integrata da elementi della teoria keynesiana della domanda effettiva e da idee di provenienza marxiana) non solo come alternativa paradigmatica al neoclassicismo, ma anche come viatico scientifico alla modificazione dei rapporti sociali in senso fortemente "progressista". Del resto, la caratteristica di teoria pura dell'economia sraffiana non escludeva questo esito, anche se, data la sua natura di schema aperto, avrebbe potuto contemplarne molteplici altri. Lo stesso Spaventa, nel suo intervento alla riunione della Società italiana degli economisti del 1969, aveva rilevato la pluralità di ipotesi istituzionali (e distributive) alternativamente compatibili con il modello sraffiano di base (Spaventa, 1974, pp. 127-128). Sennonché, nello sposare la versione istituzionale "progressista" e conflittuale di cui abbiamo detto sopra, la gran parte dei protagonisti di quel dibattito finì per considerarla uno sbocco necessario dalla validità generale, non come una tra le opzioni possibili. In altre parole, si determinò una sovrapposizione indebita tra i due diversi piani d'analisi, il primo costituito da un sistema oggettivo di relazioni economiche naturali, il secondo sostenuto da giudizi di valore tipici dell'ideologia di sinistra. La partecipazione di economisti marxisti a questa operazione di rinnovamento paradigmatico non fece che rafforzare questo malinteso processo di identificazione.

In una prima fase caratterizzata da una congiuntura storica di forte contestazione del modello sociale precedente (fine anni Sessanta, primi anni Settanta), questa interpretazione favorì nella comunità italiana degli economisti forme convergenti di consenso. Invece, in una seconda fase (quella delle manifestazioni più

⁴⁷ Questa distinzione tra i due livelli di analisi – quello teorico fondativo, e quello istituzionale volto alla individuazione di adeguate forme di organizzazione sociale – è ciò che PASINETTI L. (2007) chiama *teorema della separazione*.

acute della crisi economica nella seconda metà degli anni Settanta), provocò piuttosto disorientamento intellettuale: le istituzioni pensate a partire dallo schema puro sraffiano (prima fra tutte, il salario-variabile-indipendente) stavano ostacolando l'uscita dalla crisi. L'effetto finale fu che l'obsolescenza dell'impianto politico-istituzionale trascinò nel discredito il suo schema teorico ispiratore.

Non è un caso che sia iniziato da allora un processo di emarginazione silenziosa del filone di studi che si richiama alla teoria di Sraffa, processo a cui riteniamo abbiano contribuito i motivi sopra richiamati⁴⁸.

Per certo vi è che ciò che venne a conclusione verso la fine degli anni Settanta primi anni Ottanta, fu l'idea di saldare l'elaborazione di un nucleo teorico dalle molteplici derivazioni (Sraffa, Keynes, Marx) con la visione di un nuovo tipo di società. Ma né il collante politico, né quello delle affinità scientifiche, furono sufficienti a fare di quel nucleo teorico composito e solo abbozzato, un progetto unitario di rinnovamento della scienza economica e delle istituzioni.

⁴⁸ In questi ultimi anni stanno tuttavia aumentando le ricerche finalizzate al superamento di questa emarginazione silenziosa. Si veda ad esempio il fascicolo speciale che il *Cambridge Journal of Economics* ha dedicato alle «New Perspectives on the Work of Piero Sraffa», vol. 36, Issue 6, November 2012.

BIBLIOGRAFIA

- ASSO P.F., «Dibattiti teorici, integrazione europea ed economia italiana 1970-2000», in BINI P. (a cura di), *I novant'anni della Rivista di Politica Economica (1911-2000). Teorie economiche, scelte politiche e cultura d'impresa nell'Italia del Novecento*, Fascicolo speciale della *Rivista di Politica Economica*, Roma, Sipi Editore, 2004, pp. 179-255.
- .-, «Franco Modigliani e l'Italia», in MODIGLIANI F., *L'impegno civile di un economista. Scritti editi e inediti sull'economia e la società italiana*, a cura di Pier Francesco Asso, Siena, Fondazione Monte dei Paschi, 2007, pp. 9-49.
- AUTORI VARI, *Marxismo ed economia. Un dibattito di "Rinascita"*, Venezia, Marsilio Editori, 1974.
- BACKHOUSE R.E., *The Puzzle of Modern Economics. Science or Ideology?*, Cambridge University Press, 2010.
- BALDUCCI R. - MARCONI M., «L'accumulazione del capitale nella visione del governo, della Banca d'Italia e della Confederazione generale dell'industria», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 71-108.
- BECATTINI G., «Intervento», in SOCIETÀ ITALIANA DEGLI ECONOMISTI, *Essenza e limiti del marginalismo nelle teorie economiche*, Atti della decima riunione scientifica svoltasi a Roma il 5-7 novembre 1969, Milano, Giuffrè Editore, 1974, pp. 186-193.
- BINI P. - MAGLIULO A., «Keynesianism in Italy. Before and after the General Theory», in PASINETTI L.L. - SCHEFOLD B. (eds.), *The Impact of Keynes on Economics in the 20th Century*, Cheltenham, UK, Northampton, US, Edward Elgar, 1999, pages 131-152.
- BIRNER J., *The Cambridge Controversies in Capital Theory. A Study in the Logic of Theory Development*, London and New York, Routledge, 2002.
- BLAUG M., *The Cambridge Revolution. Success or Failure?*, The Institute of Economic Affairs, 1974. L'edizione italiana di questo scritto fu pubblicata due anni dopo, BLAUG M., *La Rivoluzione di Cambridge*, Introduzione di Neri Salvadori, Napoli, Liguori Editore, 1976.
- BOTTA F. (a cura di), *Il dibattito su Sraffa*, Bari, De Donato Editore, 1974.
- CAMPANELLA F., «Scelte sindacali e teorie economiche», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 171-202.
- CARVALE G. (ed.), *The Crisis in Economic Theories*, contributi di KREGEL J. - BLAUG M. - PASINETTI L. - KALDOR N. - EATWELL J. - STEEDMAN I. e AL., Milano, Franco Angeli, 1983.
- CARLI G., *Cinquant'anni di vita italiana*, con la collaborazione di Paolo Peluffo, Roma-Bari, Laterza Editore, 1996.
- CASAROSA C., «Gli studi nel Regno Unito della seconda generazione», in GAROFALO G. - GRAZIANI A. (a cura di), *La formazione degli economisti in Italia (1950-1975)*, Bologna, il Mulino, 2004, pp. 535-554.

- CATTABRINI F., «Franco Modigliani and the Italian Left-Wing: The Debate over Labor Cost (1975-1978)», *History of Economic Thought and Policy*, no. 1, 2012, pages 75-95.
- CESPE (CENTRO STUDI DI POLITICA ECONOMICA DEL PCI), *Crisi economica e condizionamenti internazionali dell'Italia*, Atti del convegno svoltosi a Roma il 15-17 marzo 1976, Roma, Editori Riuniti, 2 vol, 1976.
- CHIODI G. - DITTA L. (eds.), *Sraffa or An Alternative Economics*, New York, Palgrave Macmillan, 2008.
- CIOCCA P., «Il contributo di via Nazionale», in GAROFALO G. - GRAZIANI A. (a cura di), *La formazione degli economisti in Italia (1950-1975)*, Bologna, il Mulino, 2004, pp. 579-598.
- .-, *Dopo Keynes, dopo Sraffa: il pensiero "critico" e l'economia italiana*, intervento al Convegno "Gli economisti postkeynesiani di Cambridge e l'Italia" svoltosi a Roma, 11-12 marzo 2009, dattiloscritto.
- CNR (CONSIGLIO NAZIONALE DELLE RICERCHE), *La politica economica a breve termine in Italia*, Napoli, Liguori Editore, 1969.
- CONVENEVOLE R., «Commenti», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 430-432.
- CRAINZ G., *Il paese mancato. Dal miracolo economico agli anni Ottanta*, Roma, Donzelli Editore, 2003.
- D'ADDA C. - SALITURO B., «L'economia italiana negli anni settanta e ottanta», *Rivista di Politica Economica*, aprile, 1989, pp. 317-344.
- DEL PUNTA V., «La validità degli schemi marginalisti tradizionali», in SOCIETÀ ITALIANA DEGLI ECONOMISTI, *Essenza e limiti del marginalismo nelle teorie economiche*, Atti della decima riunione scientifica svoltasi a Roma il 5-7 novembre 1969, Milano, Giuffrè Editore, 1974a, pp. 59-78.
- .-, «Replica», in SOCIETÀ ITALIANA DEGLI ECONOMISTI, *Essenza e limiti del marginalismo nelle teorie economiche*, Atti della decima riunione scientifica svoltasi a Roma il 5-7 novembre 1969, Milano, Giuffrè editore, 1974b, pp. 215-223.
- D'ANTONIO M. (a cura di), *La crisi post-keynesiana*, Torino, Boringhieri Editore, 1975.
- .-, «Commenti», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 402-407.
- FEDERAZIONE CGIL-CISL-UIL, «Proposte per una svolta di politica economica e di sviluppo civile e democratico», *Rassegna sindacale*, n. 2, 19 gennaio 1978, pp. v-xiv.
- FERRULLI S., *Economisti italiani a Cambridge (1960-1980). Processi formativi, aggiornamenti teorici e nuovi orientamenti di politica economica*, tesi di dottorato discussa presso la Scuola dottorale di Scienze politiche dell'Università di Roma 3 nell'anno accademico 2012, ciclo XXIV.
- FILOSA R. - REY G.M., «Presentazione», in TARANTELLI E., *L'utopia dei deboli è la paura dei forti. Saggi, relazioni e altri scritti accademici*, Milano, Franco Angeli, 1988, pp. xiii-xxiv.

- FOTIA G., «La posizione della Banca d'Italia sul problema della bilancia dei pagamenti», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 259-272.
- FREDIANI G., «Il senso della storia nella ricerca economica», in AUTORI VARI, *Marxismo ed economia. Un dibattito di "Rinascita"*, Venezia, Marsilio Editori, 1974, pp. 71-75.
- FRIEDMAN M., «The Role of Monetary Policy», *American Economic Review*, no. 58, 1968, pages 1-17.
- FUSCO A.M., «Gli studi di economia in Italia: momenti di riflessione teorica (1946-1996)», in ARCELLI M. (a cura di), *Storia, economia e società italiana 1947-1997*, introduzione di Cesare Geronzi, Roma-Bari, Laterza Editori, 1997, pp. 71-125.
- .-, «Sulla formazione degli economisti italiani nel corso del primo e del secondo dopoguerra», in FUSCO A.M., *I libri di "Economia Italiana". Cose, uomini, idee*, Napoli, Editoriale Scientifica srl, 2006, pp. 297-305.
- GAREGNANI P., *Marx e gli economisti classici. Valore e distribuzione nelle teorie del sovrappiù*, Torino, Einaudi Editore, 1981.
- GAROFALO G., «Gli economisti in Italia negli anni 1950-1975», *Economia Politica*, n. 3, dicembre 2005, pp. 381-397.
- GAROFALO G. - GRAZIANI A. (a cura di), *La formazione degli economisti in Italia (1950-1975)*, Bologna, il Mulino, 2004.
- GATTEI G., «Sovrappiù e plusvalore in Italia (1960-1980)», in GUIZZARDI G. - STERPI S. (a cura di), *La società italiana. Crisi di un sistema*, introduzione di A. Papisca, Milano, Franco Angeli, 1981, pp. 79-107.
- GIAVAZZI F. - SPAVENTA L., «Italia: gli effetti reali dell'inflazione e della disinflazione», *Rivista di Politica Economica*, luglio-agosto, 1989, pp. 5-43.
- GINZBURG A. - VIANELLO F., «Il fascino discreto della teoria economica», in AUTORI VARI, *Marxismo ed economia. Un dibattito di "Rinascita"*, Venezia, Marsilio Editori, 1974, pp. 15-26.
- GIOLCOLI N., «Neocorporatismo ed incertezza: la soluzione Tarantelli - Heiner», *Studi economici*, n. 2, 1998, pp. 49-80.
- GORDON R.J., «Alternative Responses of Policy to External Supply Shocks», *Brookings Papers on Economic Activity*, no. 6, 1975, pages 183-266.
- .-, «Supply Shocks and Monetary Policy Revisited», *The American Economic Review, Papers and Proceedings*, no. 74, 1984, pages 38-43.
- .-, «The History of the Phillips Curve: Consensus and Bifurcation», *Economica*, no. 78, 2011, pages 10-50.
- GRAZIANI A., «I problemi della politica a breve termine in Italia», in CNR (CONSIGLIO NAZIONALE DELLE RICERCHE), *La politica economica a breve termine in Italia*, Napoli, Liguori Editore, 1969, pp. 29-38.

- GRAZIANI A., «La teoria della distribuzione del reddito», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 285-340.
- GRAZIANI A. - MELONI F., «Inflazione e fluttuazione della lira», 1973, in NARDOZZI G. (a cura di), *I difficili anni '70. I problemi della politica economica italiana 1973/1979*, Sonzogno, Etas, 1980, pp. 59-80 e 341-351.
- HEINER R., «The Origin of Predictable Behaviour», *The American Economic Review*, vol. 73, 1983, pages 560-595.
- IZZO L. - PEDONE A. - SPAVENTA L. - VOLPI F., *Il controllo dell'economia nel breve periodo*, (prima edizione, 1970), Milano, Franco Angeli, 1973.
- LAIBMAN D. (a cura di), *Marx e Sraffa*, Parma-Lucca, Pratiche Editrice, 1978.
- LAMA L., «Intervista» rilasciata a Eugenio Scalfari, *La Repubblica*, 24 gennaio 1978.
- LOMBARDINI S., «Economia e politica: il problema della distribuzione dei redditi nella teoria conflittualista», in LOMBARDINI S., *Il metodo della scienza economica: passato e futuro*, Torino, Utet, 1983, pp. 51-68.
- LUCAS R.E. JR., «Expectations and the Neutrality of Money», *Journal of Economic Theory*, no. 4, 1972, pages 103-24.
- .-, «Some International Evidence on Output-Inflation Tradeoffs», *The American Economic Review*, no. 63, 1973, pages 326-34.
- LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981.
- .-, «Sui modi di produzione della "scienza" economica», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 5-12.
- MAES I., «The Spread of Keynesian Economics. A Comparison of the Belgian and Italian Experience (1945-1970)», *Journal of the History of Economic Thought*, vol. 30, no. 4, December 2008, pages 491-509.
- MICHELAGNOLI G., *Ezio Tarantelli. Economic Theory and Industrial Relations*, Berlin Heidelberg, Springer, 2012.
- MODIGLIANI F., «Attenzione ai pericoli della contingenza unificata», in TARANTELLI E. (a cura di), *Salario e crisi economica. Dalla "ricetta" Modigliani al dopo-elezioni*, Roma, Savelli editore, 1976a, pp. 71-74.
- .-, «Intervista negli Stati Uniti con Franco Modigliani», in TARANTELLI E. (a cura di), *Salario e crisi economica. Dalla "ricetta" Modigliani al dopo-elezioni*, Roma, Savelli editore, 1976b, pp. 81-86
- .-, *Avventure di un economista*, a cura di Paolo Peluffo, Bari, Laterza, 1999.
- .-, *L'impegno civile di un economista. Scritti editi e inediti sull'economia e la società italiana*, a cura di Pier Francesco Asso, Siena, Fondazione Monte dei Paschi, 2007.

- MODIGLIANI F. - TARANTELLI E., «A Generalization of the Phillips Curve for a Developing Country», *The Review of Economic Studies*, vol. XL(2), no. 122, 1973, pages 203-223.
- .-, *Mercato del lavoro, distribuzione del reddito e consumi privati*, Bologna, il Mulino, 1975.
- .-, «Market Forces, Trade Union Action and the Phillips Curve in Italy», *Quarterly Review*, Banca Nazionale del Lavoro, no. 31(1), 1977, pages 3-36.
- .-, «Determinanti strutturali e transitorie della mobilità del lavoro, “la congettura di Holt” e l’esperienza italiana», *Moneta e Credito*, n. 32(2), 1979, pp. 123-148.
- NAPOLEONI C., «Il capitalismo con le stampelle», in TARANTELLI E (a cura di), *Salario e crisi economica. Dalla “ricetta” Modigliani al dopo-elezioni*, Roma, Savelli editore, 1976, pp. 178-183.
- NARDOZZI G. (a cura di), *I difficili anni '70. I problemi della politica economica italiana 1973/1979*, Sonzogno, Etas, 1980.
- .-, «Accumulazione di capitale e politica monetaria: il punto di vista della Banca d’Italia», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 109-123.
- PASINETTI L.L., *Keynes and the Cambridge Keynesians. A “Revolution in Economics” to be Accomplished*, Cambridge University Press, 2007.
- PHELPS E.S., «Phillips Curve, Expectations of Inflation and Optimal Unemployment over Time», *Economica*, no. 34, 1967, pages 254-281.
- .-, «Money-Wage Dynamics and Labor-Market Equilibrium», *Journal of Political Economy*, no. 76, 1968, pages 678-711.
- PHILLIPS A.W., «The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom 1861-1957», *Economica*, no. 25, 1958, pages 283-299.
- PIVETTI M., «Esportazioni di capitali e salari reali», in AUTORI VARI, *Marxismo ed economia. Un dibattito di “Rinascita”*, Venezia, Marsilio Editori, 1974, pp. 27-36.
- RAMPA L., «Visioni e proposizioni teoriche nel dibattito a sinistra sulla politica economica (1971-1977)», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 125-170.
- RONCAGLIA A., *The Wealth of Ideas. A History of Economic Thought*, Cambridge University Press, 2005.
- ROSSI S., *La politica economica dell’Italia 1968-2007*, Roma-Bari, Laterza, 2007.
- SALVATI M., «Le difficoltà che si incontrano quando si vuol fare una critica all’economia politica. Con qualche esempio derivato dalle teorie dell’inflazione», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 341-353 e 419-426.
- SAMUELSON P.A., *Economics. An Introductory Analysis*, McGraw-Hill, seventh edition, 1967.

- SAVONA P., «Considerazioni sullo stato della modellistica economica in Italia: Note a margine dei due modelli CSC», *Economia Italiana*, n. 2, 1981a, pp. 213-235.
- .-, «Dalla politica della domanda alla politica dell'offerta. La via della Previsionale e Programmatica 1981», *Banche e Banchieri*, n. 8, 1981b, pp. 19-26.
- .-, «Guido Carli in Confindustria: maestro di pensiero e statista», in SAVONA P. (a cura di), *Guido Carli presidente di Confindustria 1976-1980*, Torino, Bollati Boringhieri, 2008.
- SOCIETÀ ITALIANA DEGLI ECONOMISTI, *Essenza e limiti del marginalismo nelle teorie economiche*, Atti della decima riunione scientifica svoltasi a Roma il 5-7 novembre 1969, Milano, Giuffrè Editore, 1974.
- SPAVENTA L., *Appunti di economia politica*, Roma, Bulzoni Editore, 1971.
- .-, «Significato e portata della critica alla teoria marginalista della distribuzione», in SOCIETÀ ITALIANA DEGLI ECONOMISTI, *Essenza e limiti del marginalismo nelle teorie economiche*, Atti della decima riunione scientifica svoltasi a Roma il 5-7 novembre 1969, Milano, Giuffrè Editore, 1974, pp. 101-130.
- .-, «Una svolta ad U della teoria economica», in *Scritti in onore di Innocenzo Gasparini*, vol. II, Milano, Giuffrè editore, 1982, pp. 1035-1058.
- .-, «Il Gruppo CNR per lo studio dei problemi economici della distribuzione, del progresso tecnico e dello sviluppo. Una infrastruttura di formazione e di ricerca», in GAROFALO G. - GRAZIANI A. (a cura di), *La formazione degli economisti in Italia (1950-1975)*, Bologna, il Mulino, 2004, pp. 555-578.
- SRAFFA P., *Production of Commodities by Means of Commodities. Prelude to a Critique of Economic Theory*, Cambridge University Press, 1960.
- SYLOS LABINI P., *Oligopoly and Technical Progress*, Cambridge Mass., Harvard University Press, 1962.
- .-, *Sindacati, inflazione e produttività*, Bari, Laterza, 1972.
- .-, (a cura di), *Prezzi relativi e distribuzione del reddito*, Torino, Boringhieri, 1973.
- .-, *Un paese a civiltà limitata. Intervista su etica, politica ed economia*, PETRINI R. (a cura di), Roma-Bari, Laterza, 2001.
- TARANTELLI E. (a cura di), *Salario e crisi economica. Dalla "ricetta" Modigliani al dopoelezioni*, Roma, Savelli Editore, 1976.
- .-, *Il ruolo economico del sindacato e il caso italiano*, Roma-Bari, Laterza, 1978.
- .-, *L'utopia dei deboli è la paura dei forti. Saggi, relazioni e altri scritti accademici*, Milano, Franco Angeli, 1988a.
- .-, «Forze di mercato, azione sindacale e la curva di Phillips in Italia» (1976), in TARANTELLI E., *L'utopia dei deboli è la paura dei forti. Saggi, relazioni e altri scritti accademici*, Milano, Franco Angeli, 1988b, pp. 413-446.
- .-, «La funzione del consumo e le "teorie alternative" della distribuzione del reddito» (1979), in TARANTELLI E., *L'utopia dei deboli è la paura dei forti. Saggi, relazioni e altri scritti accademici*, Milano, Franco Angeli, 1988c, pp. 863-879.

TARANTELLI E., *La forza delle idee. Scritti di economia e politica*, Roma-Bari, Laterza, 1995.

VALLI V., «Commenti», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 426-428.

-.-, *Politica economica. I modelli, gli strumenti, l'economia italiana*, Roma, La Nuova Italia Scientifica, 1989.

VIANELLO F., «Commenti», in LUNGHINI G. (a cura di), *Scelte politiche e teorie economiche in Italia 1945-1978*, Torino, Einaudi, 1981, pp. 399-402.

Variabili predittive dell'astensionismo elettorale: una modellizzazione della partecipazione al voto

Giuseppe Garofalo Silvia Pugliesi*

Università della Tuscia, Viterbo

Citizens may choose not to vote when the candidates are too similar to each other (indifference) or when they perceive little benefit from each candidate (alienation). The paper implements a unified model that captures the effect on voting behaviour of the candidates' (perceived) positions along the ideological scale, non-spatial factors (valence and partisanship), citizens' socio-demographic characteristics, their sensitivity to the political system and the mass media exposure. By applying this model to the 2008 national elections, the paper finds that the determinants of alienation and indifference differed considerably and that both the citizens' ideology and individual thresholds affected the voting probability.

[JEL Classification: D72].

Keywords: voting behavior; representativeness; voter utility/disutility; abstention by indifference; abstention by alienation.

* <garofalo@unitus.it>; <silviapugliesi@yahoo.it>, Dipartimento di Economia e Impresa.

1. - Introduzione

I fattori determinanti la decisione di partecipazione o astensione dal voto costituiscono il tema rilevante di un'ampia parte della letteratura sul comportamento degli elettori¹. L'affluenza alle urne nei Paesi sviluppati si è notevolmente ridotta negli ultimi anni² e l'Italia conferma questo trend decrescente: la partecipazione al voto ha subito un progressivo rallentamento passando dal 93% nel 1972 a poco più dell'80% nel 2008. Tale fenomeno allarma studiosi e *policy-makers* in quanto la partecipazione democratica, oltre a misurare la rappresentatività delle istituzioni e la qualità del processo democratico (Lijpart, 1997), influenza la distribuzione del reddito (Mueller e Stratmann, 2003), il grado di efficienza del settore pubblico (Borge *et al.*, 2008) ed il capitale civico (Putnam, 1993), a sua volta determinante la performance economica di un paese (Guiso *et al.*, 2010).

In letteratura è ben noto come la probabilità di votare sia influenzata da caratteristiche socio-demografiche, quali età, livelli di istruzione (Dee, 2003; Milligan *et al.*, 2003), reddito (Wolfinger e Rosenstone, 1980), informazione politica (Matsusaka, 1995; Ghirardato e Katz, 2002; Larcinese, 2009) ed esposizione ai mass media (George e Waldfogel; 2003, 2006).

Tuttavia il tema del legame tra le posizioni politiche dei partiti e le decisioni di partecipazione/astensione degli elettori ha avuto un ruolo marginale. A livello teorico, un'ampia parte della letteratura si è focalizzata su questo argomento all'interno della teoria della competizione politica spaziale, riconoscendo due principali cause di astensione: indifferenza e alienazione. L'elettore si astiene se è indifferente tra le politiche proposte dai partiti che appaiono così simili/vicine da non incentivare l'elettore a recarsi alle urne e contribuire con il proprio voto alla vittoria del partito preferito. Un elettore non-indifferente può ancora scegliere di astenersi se nessun partito è in grado di soddisfare sufficientemente i suoi interessi, vale a dire se nessuna proposta politica è tale da convincerlo a recarsi alle urne. Il cittadino che si astiene è disinteressato, alienato (astensione per alienazione): la proposta politica preferita, la più vicina, è comunque troppo lontana dalla sua politica ideale da garantire un livello minimo di utilità (utilità di aspirazione), inversamente correlata alla *tolleranza* che costituisce la distanza massima dal partito tale da persuadere il cittadino a votare.

¹ Per una rassegna si veda GEYS B. (2006).

² Per maggiori informazioni si può consultare il database sull'affluenza alle urne fornito dall'organizzazione IDEA (International Institute for Democracy and Electoral Assistance), disponibile all'indirizzo: <http://www.idea.int/index.cfm>.

Un'ampia parte della letteratura empirica sul comportamento elettorale non distingue le due fonti di astensione, lasciando irrisolta la questione sul modo in cui le variabili considerate aumentino la partecipazione al voto dei cittadini, soprattutto in riferimento alle due principali cause di astensione: le caratteristiche individuali aumentano la tolleranza degli elettori, vale a dire riducono la probabilità di astensione per alienazione, oppure riducono il valore soglia dell'indifferenza, diminuendo, così, la probabilità di astensione per indifferenza? Si conosce, inoltre, ben poco circa i legami e le interrelazioni esistenti tra affluenza, scelte degli elettori e posizionamento politico e ideologico dei partiti. L'evidenza empirica fornita da Zipp (1985); Thurner e Eymann (2000); Plane e Gershtenson (2004); Adams e Merrill (2003) e Adams *et al.* (2006), individua una componente di astensione politica, per cui l'astensionismo è funzione delle posizioni (percepite) dei partiti/candidati su alcuni rilevanti temi politici. In uno spazio politico unidimensionale, come in Katz (2007), l'astensione politica è squisitamente interpretabile come astensione ideologica, per cui l'impatto relativo di alienazione e indifferenza sull'astensionismo dipende dalla distanza ideologica (percepita) dai partiti/candidati. Coerentemente alla teoria spaziale del voto, questi lavori rilevano che la distanza ideologica e politica dai partiti/candidati influenza significativamente la probabilità di votare. Inoltre, il peso relativo della componente di astensione per indifferenza e alienazione è influenzato, in parte, da fattori non politici quali l'identificazione con il partito (*partisanhip*)³ e la valenza del candidato, vale a dire la competenza, l'integrità, il carisma, etc. (Stokes, 1963; Shepsle, 1991; Ansolabehere e Snyder, 2000; Groseclose, 2001).

Al fine di analizzare le interrelazioni tra la partecipazione al voto e le scelte dell'elettore, alcuni recenti lavori hanno utilizzato un modello consolidato della

³ L'influenza dell'identificazione con il partito sul comportamento dell'elettore è la caratteristica centrale del "modello Michigan" di voto (CAMPBELL A. *et al.*, 1960) in cui le scelte di voto sono il risultato di fattori di lungo termine – l'orientamento ideologico del cittadino e del contesto familiare/culturale in cui vive – e di fattori di breve termine, quali l'immagine del candidato, la campagna elettorale, le condizioni economiche. L'identificazione con il partito è, pertanto, un fattore psicologico, a lungo termine, che spesso deriva dai gruppi sociali cui appartiene il cittadino. Sull'impatto elettorale dell'identificazione con i partiti, la letteratura empirica fornisce risultati divergenti; alcuni autori ritengono che l'identificazione eserciti un'influenza dominante sul voto (CAMPBELL A. *et al.*, 1960; CONVERSE P. - PIERCE R., 1986, 1993; ENDERSBY J. - GALATAS S., 1997), altri sostengono che il voto sia guidato principalmente dal senso di appartenenza ai gruppi sociali (BAKER K. *et al.*, 1981; BARTLE J., 1998); infine, alcuni autori rilevano che l'impatto dell'identificazione con il partito varia significativamente con il contesto istituzionale (LEWIS-BECK M. e CHLARSON K., 2002).

partecipazione al voto e scelta dell'elettore, anche denominato modello indifferenza-alienazione, in base al quale la decisione di partecipazione o astensione e la decisione di quale candidato/partito votare sono incluse simultaneamente, vale a dire ogni cittadino sceglie simultaneamente se votare e per quale candidato/partito votare. Tale modello, sviluppato da Adams *et al.* (2001); Adams e Merrill (2003) e Adams *et al.* (2005)⁴, considera entrambe le cause di astensionismo e cattura l'effetto sul comportamento degli elettori di fattori politici e non-politici, spaziali e non-spaziali, quali le caratteristiche socio-demografiche dei cittadini, le predisposizioni nei confronti dei partiti e le valutazioni delle abilità/competenze dei candidati (valenza). Il principale scopo del modello di stima è rilevare i fattori determinanti l'astensione per alienazione e/o indifferenza e verificare l'esistenza di una componente significativa di astensione politica e ideologica.

Alcuni studi sul comportamento degli elettori italiani (Bellucci e Sagatti, 2012) evidenziano l'associazione tra astensionismo e variabili socio-demografiche (quali sesso, istruzione, professione, regione di appartenenza), nonché l'importanza dell'identificazione con il partito, la valutazione delle capacità del leader e la prossimità ideologica (Garzia e Viotti, 2012). Tuttavia, a nostra conoscenza, non risulta alcun contributo che applichi un modello di stima basato sull'utilità dell'elettore e sui valori soglia individuali di alienazione e indifferenza.

Il presente lavoro intende indagare le variabili predittive dell'astensionismo nelle elezioni parlamentari italiane del 2008, distinguendone il contributo alle due cause di astensione. In particolare, lo studio si propone di individuare:

- la relazione tra le caratteristiche individuali, socio-demografiche e attitudinali, e la probabilità di astensione per indifferenza e/o alienazione;
- l'impatto relativo di indifferenza e/o alienazione sull'astensionismo;
- l'influenza della dimensione ideologica e dei valori soglia individuali sulle due fonti di astensione.

L'articolo è così organizzato: la sezione 2 descrive il modello di stima, la sezione 3 introduce il lavoro e fornisce un'analisi descrittiva del *dataset* utilizzato; la sezione 4 riporta le stime e, infine, la sezione 5 riassume i principali risultati.

⁴ Questi lavori si basano sui contributi di SANDERS M. (1998, 2001); BURDEN B. - LACY D. (1999) e THURNER P. - EYMANN A. (2000).

2. - Una modellizzazione dell'astensione dal voto: i fattori indifferenza e alienazione

Traendo spunto dai lavori di Adams *et al.* (2003, 2006) sulle elezioni presidenziali americane, di Adams *et al.* (2005) sulle elezioni politiche nazionali in Francia, Gran Bretagna, Norvegia e di Katz (2007) sulle elezioni politiche in Brasile, adotteremo un modello di stima in base al quale la decisione di partecipazione o astensione e la decisione del candidato/partito cui assegnare il proprio voto sono incluse simultaneamente: ogni cittadino effettua una scelta tra $J+1$ diverse alternative: votare il partito j ($j=1, \dots, J$) o astenersi.

Il modello è a scelta discreta multinomiale (modello probabilistico) e segue l'approccio dell'utilità casuale (*RUM, Random Utility Models*), secondo il quale l'utilità di un'alternativa può esprimersi come la somma di una componente deterministica e casuale (il termine di errore).

In linea con la teoria spaziale del voto, si assume che gli elettori intervistati siano posizionati secondo le proprie preferenze lungo una scala ideologica sinistra-destra e che, pertanto, valutino i diversi candidati sulla base della *distanza ideologica*, ovvero la distanza dalla posizione (media percepita) del candidato. Oltre al fattore spaziale, l'utilità percepita del cittadino è influenzata anche da due fattori non-spaziali: la *valenza* del candidato, ovvero la qualità percepita dall'elettore, e l'*identificazione* con il partito (*partisanship*).

L'utilità dell'elettore i -esimo, la cui politica ideale è p_i , dal candidato j , con $j=1, \dots, J$, è definita come:

$$(1) \quad U_i^j = \alpha_j v_i^j + \gamma_j d_i^j - \beta_j (p_i - c_j)^2 + \varepsilon_i^j$$

dove c_j è la posizione (media percepita) del candidato j lungo la scala ideologica; v_i^j identifica la valenza del candidato j percepita dall'elettore i ; α_j rappresenta il peso della valenza del candidato nell'utilità del cittadino ed è un parametro da stimare; d_i^j misura la *partisanship*, ovvero l'identificazione dell'elettore con il partito di appartenenza del candidato e γ_j il relativo parametro da stimare.

In linea con la teoria spaziale del voto, il cittadino i riceve una disutilità pari alla distanza ideologica (quadratica) dal candidato⁵; il parametro β_j , da stimare,

⁵ Introdurre nelle funzioni di utilità la distanza assoluta o quadratica non cambia i risultati di stima: poiché, in base all'approccio dell'utilità casuale, la probabilità di scelta di un'alternativa dipende dalla differenza delle utilità, la scala delle utilità non è rilevante.

rappresenta l'importanza della distanza ideologica nell'utilità del cittadino e ε_i^j è il termine di disturbo casuale. La componente sistematica dell'utilità del cittadino i -esimo è indicata con S_i^j , vale a dire:

$$S_i^j = \alpha_j v_i^j + \gamma_j d_i^j - \beta_j (p_i - c_j)^2$$

L'elettore può astenersi per indifferenza o alienazione. Nel primo caso il cittadino non ritiene che i candidati siano sufficientemente diversi l'uno dall'altro da giustificare il *costo* (positivo) dell'atto di recarsi alle urne e contribuire, così, alla vittoria di un candidato. Nel secondo caso il cittadino si astiene in quanto nessun candidato è in grado di assicurargli un livello minimo di utilità, un'*utilità di aspirazione*. Pertanto, l'elettore sceglie di astenersi se:

- non c'è nessun candidato j tale da garantire al cittadino un'utilità sufficientemente più grande di quella fornita da ogni altro candidato (astensione per indifferenza). L'utilità ricevuta da ogni candidato è così simile a quella ottenuta da (almeno) uno degli altri candidati, da non compensare il costo di votazione (valore soglia dell'indifferenza). Il cittadino non percepisce sostanziali differenze tra un candidato e l'altro risultando, così, completamente disincentivato dal recarsi alle urne. D'altro canto, se l'utilità associata ad uno dei candidati consente di ottenere un surplus, sufficientemente elevato, sull'utilità ottenuta da ogni altro candidato, l'elettore è interessato alla vittoria elettorale del candidato preferito, per cui è motivato a recarsi alle urne ed assegnare il proprio voto. Formalmente il cittadino i è *indifferente* se non c'è un candidato j tale che:

$$(2) \quad U_i^j - U_i^k > I_i \quad \forall k \neq j, j = 1, \dots, J$$

dove I_i rappresenta il valore soglia dell'indifferenza per il cittadino i (il costo di votazione);

- non c'è nessun candidato j in grado di rappresentare adeguatamente gli interessi del cittadino (astensione per alienazione). Formalmente, il cittadino i è *alienato* se nessuno dei candidati gli garantisce un livello minimo di utilità, ovvero:

$$(3) \quad U_i^j < A_i \quad \forall j = 1, \dots, J$$

I valori soglia individuali dell'indifferenza I_i , e dell'alienazione A_i , dipendono da un vettore di variabili esplicative che include le caratteristiche individuali, socio-demografiche (età, istruzione, etc.), variabili attitudinali (informazione politica, interesse, etc.), di percezione del sistema politico (fiducia nella democrazia, etc.) ed esposizione ai mass-media, che ci si attende possano influenzare i valori soglia del cittadino e, quindi, il comportamento elettorale. Poiché la letteratura manca di contributi teorici o empirici che definiscano le variabili esplicative di sola pertinenza dell'alienazione, di sola pertinenza dell'indifferenza e relative ad entrambe, si assume che:

$$(4) \quad I_i = \exp(\beta^I X_i)$$

$$(5) \quad A_i = \beta^A X_i + \varepsilon_i^A$$

dove X_i è il vettore di variabili esplicative comune mentre i vettori β^I , β^A , esprimono, rispettivamente, l'influenza delle variabili considerate sui valori soglia di indifferenza⁶ e alienazione. Il valore soglia dell'alienazione è l'utilità di aspirazione e, in quanto tale, ha una componente sistematica:

$$S_i^A = \beta^A X_i$$

e una componente casuale, rappresentata dal termine di errore ε_i^A .

Ogni elettore sceglie simultaneamente se astenersi o votare uno dei candidati principali. Formalmente, l'insieme delle scelte di ogni elettore è:

$$\{V^1, \dots, V^j, \dots, V^J, NV\}$$

dove V^j , con $j=1, \dots, J$, assume valore 1 se l'elettore sceglie di votare il candidato j , e 0 altrimenti; analogamente, la variabile binaria NV assume valore 1 se l'elettore sceglie di non votare e 0 altrimenti. Assumendo che i termini di errore siano indipendenti e identicamente distribuiti si determinano le probabilità di scelta di

⁶ La forma esponenziale dell'equazione (4) garantisce il segno positivo del valore soglia di indifferenza. Inoltre, al fine di ottenere delle soluzioni in forma chiusa per le probabilità di scelta, il numero dei termini di errore casuale nel modello non può eccedere $J+1$. Pertanto, coerentemente a SANDERS M. (1998); ADAMS J. - MERRILL III S. (2003); ADAMS J. et AL. (2006), si omette il termine di errore nella specificazione del valore soglia di indifferenza.

ogni alternativa: $P_i(V^j)$ indica la probabilità che l'elettore i -esimo voti il candidato j , con $j=1, \dots, J$, e $P_i(NV)$ la probabilità che l'elettore si astenga. L'elettore i vota il candidato j se non è indifferente:

$$(S_i^j + \varepsilon_i^j) - (S_i^k + \varepsilon_i^k) > I_i \quad \forall k \neq j$$

né alienato, vale a dire:

$$S_i^j + \varepsilon_i^j > S_i^A + \varepsilon_i^A$$

Pertanto si ha che:

$$\begin{aligned} P_i(V^j) &= P(S_i^j + \varepsilon_i^j - S_i^k - \varepsilon_i^k > I_i \quad \forall k \neq j, S_i^j + \varepsilon_i^j > S_i^A + \varepsilon_i^A) \\ &= P(\varepsilon_i^k < \varepsilon_i^j + S_i^j - S_i^k - I_i \quad \forall k \neq j, \varepsilon_i^A < \varepsilon_i^j + S_i^j - S_i^A) \end{aligned}$$

I termini di errore sono indipendenti e identicamente distribuiti secondo la *distribuzione del valore estremo di primo tipo EV 1* (la distribuzione generalizzata dei valori estremi di Gumbel)⁷. Dopo alcuni passaggi matematici, la probabilità di votare il candidato j è la seguente⁸:

$$(6) \quad P_i(V^j) = \frac{\exp(S_i^j)}{\exp(S_i^j) + \exp(I_i) \sum_{k \neq j} \exp(S_i^k) + \exp(S_i^A)} \quad j = 1, \dots, J$$

e poiché:

$$\sum_j P_i(V^j) + P_i(NV) = 1$$

⁷ Le espressioni della funzione di ripartizione e di densità di una distribuzione generalizzata dei valori estremi di *Gumbel* sono le seguenti:

$$\begin{aligned} f(\varepsilon) &= \exp(-\varepsilon) \exp[-\exp(-\varepsilon)] \\ F(\varepsilon) &= \exp[-\exp(-\varepsilon)] \end{aligned}$$

⁸ Si veda ADAMS J. *et al.* (2005) e KATZ G. (2007) per una dimostrazione completa.

la probabilità che il cittadino i si astenga è⁹:

$$(7) \quad P_i(NV) = 1 - \sum_j P_i(V^j)$$

Il modello indifferenza-alienazione è stimato tramite il metodo della massima verosimiglianza. Il logaritmo della funzione di verosimiglianza (*Log-Likelihood*) è:

$$LL = \sum_{i=1}^N \left[\sum_j V^j \log P_i(V^j) + NV \log P_i(NV) \right]$$

dove $P_i(V^j)$ e $P_i(NV)$ sono costituite, rispettivamente, dalla (6) e (7). La massimizzazione della verosimiglianza restituisce le stime dei coefficienti e gli errori *standard*.¹⁰

La probabilità che l'elettore i -esimo sia alienato, vale a dire che nessun candidato sia in grado di garantirgli l'utilità di aspirazione, è definita come:

$$(8) \quad P_i(AL) = \frac{\exp(S_i^A)}{\sum_i \exp(S_i^j) + \exp(S_i^A)}$$

La probabilità che l'elettore sia indifferente, ovvero che nessun candidato sia percepito come sufficientemente diverso dagli altri, è:

$$(9) \quad P_i(IND) = 1 - \sum_j \frac{\exp(S_i^j)}{\exp(S_i^j) + \exp(I_i) \sum_{k \neq j} \exp(S_i^k)}$$

Utilizzando i coefficienti stimati si possono calcolare le probabilità predette, (6)-(9), per ogni elettore intervistato, nonché le probabilità che l'elettore sia *solo*

⁹ Il modello di stima, sebbene sia *conditional logit*, non impone la proprietà di Indipendenza dalle Alternative Irrilevanti (IIA), secondo la quale il rapporto tra le probabilità di due alternative non dipende dalla presenza/assenza di altre alternative. I denominatori delle probabilità (6) e (7) sono infatti diversi, per cui rimuovere o cambiare un'alternativa nell'insieme di scelta cambia il rapporto tra le probabilità delle alternative rimanenti (SANDERS M., 1998).

¹⁰ La massimizzazione della funzione di verosimiglianza è eseguita tramite il *software Stata 11.2* con il metodo *lf*. Il codice è disponibile su richiesta.

alienato, *solo* indifferente, alienato e indifferente. Le medie delle suddette probabilità consentono di stimare le percentuali di elettori che rientrano in ciascuna categoria.

3. - Un'applicazione del modello indifferenza-alienazione

Le elezioni politiche italiane del 2008, oltre ad essere le più recenti, costituiscono un caso di studio piuttosto interessante dal punto di vista dell'astensionismo e, in generale, del comportamento degli elettori. Le elezioni si sono svolte, infatti, a seguito dello scioglimento anticipato delle Camere avvenuto nel febbraio 2008 come conseguenza della crisi del governo di centro-sinistra, in carica dal 2006. La delusione degli elettori di questo schieramento e, in particolare, dell'estrema sinistra, a seguito della conclusione anticipata della legislatura, ha contribuito alla vittoria elettorale della coalizione di centro-destra (il Popolo della Libertà). La delusione degli elettori di sinistra, insieme alla scarsa fiducia riposta nella capacità di governo della nuova coalizione di centro-sinistra (composta dal Partito Democratico e dall'Italia dei Valori), si è manifestata sia sotto forma di astensione che di spostamento di quote significative di elettori verso il centro-destra. L'astensione nelle elezioni politiche del 2008 risulta essere dunque *selettiva*, in quanto ha colpito il centrosinistra più del centrodestra, e *punitiva* o "di protesta" poiché è nata da un clima di instabilità politica. La vittoria elettorale, con ampio margine¹¹, della coalizione di centro-destra, è stata in parte determinata dalla delusione dei cittadini, in particolare degli elettori di centro-sinistra, e dalla tendenza al voto (o non-voto) di protesta.

Allo scopo di analizzare le interrelazioni tra la partecipazione al voto, la dimensione ideologica e le caratteristiche individuali degli elettori utilizzeremo i dati dell'indagine campionaria post-elettorale svolta dall'Associazione Itanes (*Italian National Election Studies*)¹². Il *dataset* è composto da 3.000 osservazioni (in-

¹¹ La coalizione di centro-destra (il Popolo della Libertà) ha ottenuto 344 seggi nella Camera dei Deputati e 174 nel Senato della Repubblica, mentre la coalizione avversaria, di centro-sinistra, ne ha ottenuti 247 nella Camera e 134 al Senato. Il margine di vittoria della coalizione di centro-destra è stato significativamente più elevato rispetto a quello del centro-sinistra nelle elezioni politiche precedenti, del 2006.

¹² Il *dataset* è disponibile, su richiesta, al link: <http://www.itanes.org/dati/>. Il nostro lavoro utilizza solo i dati relativi alle elezioni della Camera dei Deputati. Poiché è necessario aver compiuto almeno 25 anni per poter eleggere un membro del Senato, contro i 18 per l'elezione della Camera, si è preferito concentrare l'attenzione sulle elezioni di quest'ultima, per poter analizzare l'effetto dell'età sul comportamento degli elettori senza alcun limite inferiore.

terviste) e contiene informazioni dettagliate sulle caratteristiche individuali, sulle preferenze ideologiche, sulle percezioni del sistema politico degli elettori intervistati. La *posizione ideologica* di ogni intervistato deriva dall'auto-posizionamento lungo una scala sinistra-destra da 1 a 10 (1 per estrema sinistra e 10 per estrema destra). La distribuzione degli elettori intervistati lungo la scala ideologica è riportata nel Grafico 1, dove è evidente la concentrazione degli intervistati al centro ed una maggiore frequenza di un orientamento politico di centro-sinistra.

La variabile dipendente è basata sulla risposta dell'intervistato alla domanda sul voto e indica se l'individuo si è astenuto o ha votato per uno dei quattro principali candidati: Fausto Bertinotti della coalizione di estrema sinistra (Sinistra Arcobaleno), Walter Veltroni della coalizione di centro-sinistra (costituita dal Partito Democratico e dall'Italia dei Valori), Pier Ferdinando Casini della coalizione di centro (Unione di Centro) e Silvio Berlusconi della coalizione di centro-destra (costituita dal Popolo della libertà e Lega Nord).

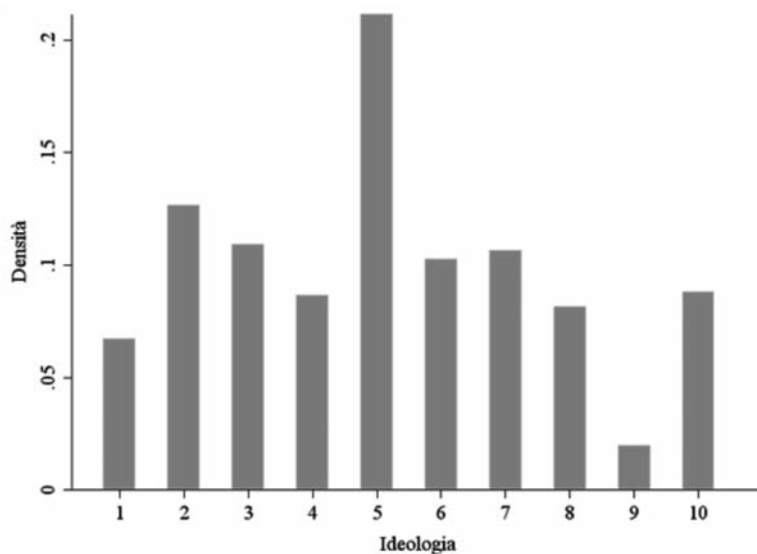
In linea con la letteratura empirica, il voto non valido (scheda in bianco o nulla) è considerato come astensione¹³. Inoltre, gli intervistati che dichiarano di aver votato o di essere intenzionati a votare uno dei candidati minori, non considerati nell'indagine, sono esclusi dal campione¹⁴.

¹³ A tale riguardo si veda LLAVAREDA A. (1991); POWER T. - ROBERTS T. (1995).

¹⁴ Si tratta degli intervistati che dichiarano di aver votato per il candidato della coalizione di estrema destra (Daniela Santanchè), del Partito Socialista (Enrico Boselli) e di altri partiti/coalizioni minori. Ciò comporta l'esclusione di 77 osservazioni su 3.000.

GRAF. 1

DISTRIBUZIONE DEGLI ELETTORI LUNGO LA SCALA IDEOLOGICA



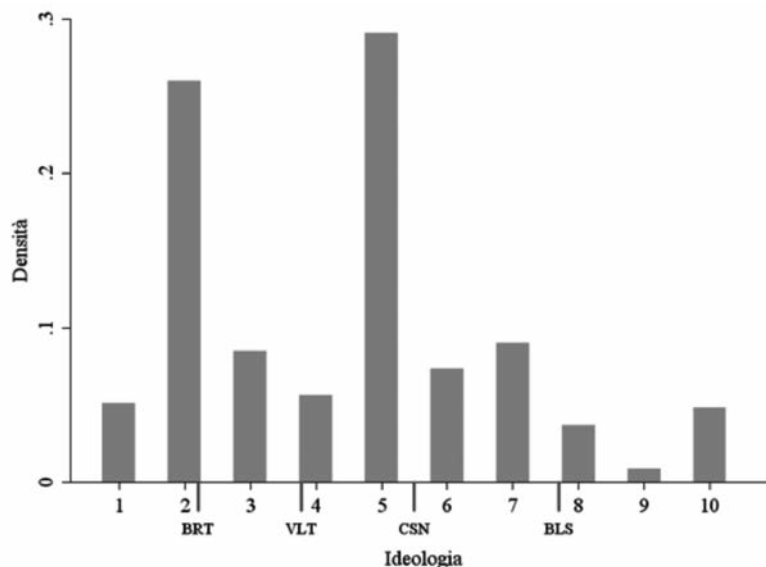
Nota: Il grafico riporta la distribuzione degli elettori lungo la scala ideologica (N=2.923).

Il Grafico 2 mostra la distribuzione degli intervistati astensionisti lungo la scala ideologica e riporta il posizionamento ideologico dei candidati. La collocazione di ogni candidato lungo la scala ideologica è la *media percepita*, ovvero la media del posizionamento attribuito al candidato dagli elettori intervistati.

L'astensione proviene soprattutto dal centro e, a conferma della componente di astensione selettiva, interessa prevalentemente l'orientamento ideologico di sinistra. In particolare, l'astensione presenta due picchi: uno in corrispondenza dell'estrema sinistra (il punto 2 della scala) e l'altro nel centro (il punto 5). Considerando le posizioni dei candidati, l'astensionismo interessa principalmente l'elettorato del candidato F. Bertinotti e P.F. Casini. Nel primo caso, ciò conferma la presenza di un'astensione punitiva, o di protesta, derivante dalla delusione per il fallimento del governo di sinistra varato nel 2006, che segue un analogo fallimento nel 1998 (il primo governo Prodi), in cui F. Bertinotti causò la fatale crisi di governo. Nel secondo caso, invece, l'astensione riflette in parte l'insoddisfazione popolare derivante dal clima di instabilità politica caratterizzante il biennio 2006-2008 e, in parte, la confusione dell'elettorato in quanto collocato tra i due maggiori avversari (W. Veltroni e S. Berlusconi) ma ideologicamente vicino al candidato centrista.

GRAF. 2

DISTRIBUZIONE DEGLI ASTENSIONISTI LUNGO LA SCALA IDEOLOGICA



Nota: Il grafico riporta la distribuzione degli elettori astensionisti lungo la scala ideologica (N=354) e le posizioni (percepite) dei diversi candidati (BRT=F. Bertinotti, VLT=W. Veltroni, CSN=P.F. Casini, BLS=S. Berlusconi).

La Tavola 1 riporta i dati sul voto nel campione: la coalizione di centro-destra ottiene la maggioranza di voti, seppur con un minore margine di vittoria rispetto ai dati osservati¹⁵. Ciò è dovuto ad una sovra-rappresentazione dell'elettorato di sinistra (in particolare del Partito Democratico e di Rifondazione Comunista) ed una sotto-rappresentazione dell'elettorato di centro (Unione di Centro) e centro-destra (Popolo delle Libertà). Tuttavia, si tratta di una caratteristica comunemente riscontrata nei campioni relativi alle diverse indagini elettorali: gli elettori ideologicamente orientati a destra mostrano una maggiore reticenza, rispetto agli elettori di sinistra, nel dichiarare il partito per cui hanno votato¹⁶.

¹⁵ La coalizione di centro-destra (il Popolo delle Libertà) ha ottenuto il 46,8% dei voti; la coalizione avversaria, di centro-sinistra, ne ha ottenuti il 37,6%.

¹⁶ Come suggerisce l'associazione Itanes, è preferibile non usare alcuna ponderazione sui dati relativi al voto. Infatti, la ponderazione sul voto produce marginali corretti sul sottocampione di coloro che hanno dichiarato il voto, ma genera un campione distorto sul campione totale, se – come è probabile – il rifiuto a rispondere alla domanda sul voto è correlato con gli orientamenti politici della persona. Considerando, inoltre, che il voto costituisce la variabile dipendente nel modello di stima, si ritiene corretto non usare alcuna ponderazione.

TAV. 1

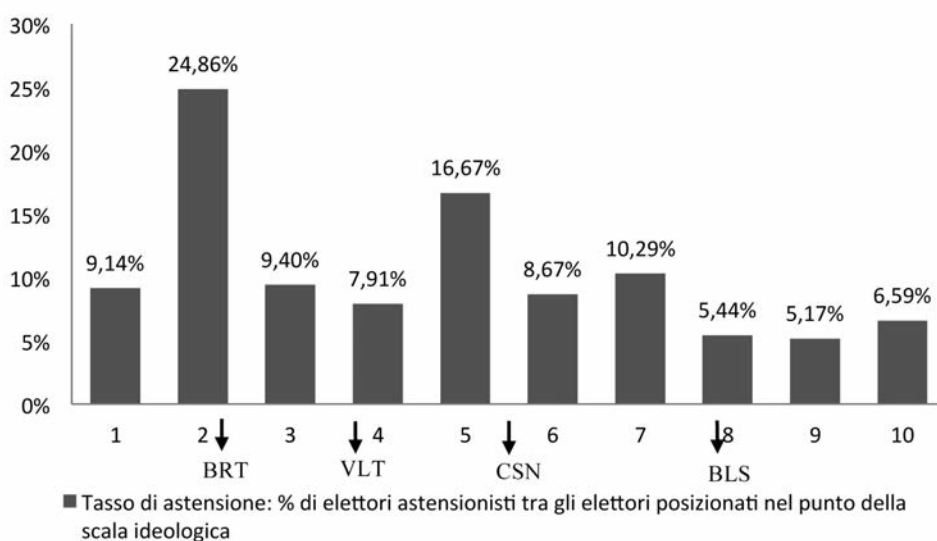
POSIZIONI IDEOLOGICHE DEI CANDIDATI, VOTI E ASTENSIONISMO

	<i>Posizione ideologica (media percepita)</i>	<i>Voti % elettori N=2.923</i>	<i>Voti % elettori attivi N=2.569</i>
S. Berlusconi	7,7	35,34	40,21
W. Veltroni	3,77	35,1	39,39
F. Bertinotti	2,2	14,1	16,04
P.F. Casini	5,49	3,4	3,85
Astensionismo		12,11	

Il Grafico 3 mostra il tasso di astensione per ogni punto della scala ideologica, calcolato come la frazione di elettori ivi collocati che hanno scelto di astenersi. È evidente come l'astensione colpisca soprattutto la regione periferica sinistra, in prossimità del candidato F. Bertinotti, così come la regione centrale. L'astensione è, piuttosto, decisamente inferiore tra gli elettori collocati nella regione periferica di destra, a conferma del fenomeno di astensione selettiva che ha colpito prevalentemente il bacino di attrazione del candidato F. Bertinotti.

GRAF. 3

TASSO DI ASTENSIONE



Nota: Il tasso di astensione è calcolato come la percentuale di intervistati, collocati nel punto della scala ideologica, che si dichiarano astenuti. Le frecce indicano le posizioni (percepita) dei diversi candidati (BRT=F. Bertinotti, VLT=W. Veltroni, CSN=P.F. Casini, BLS=S. Berlusconi).

Relativamente alle caratteristiche socio-demografiche del campione osservato, i dati mostrano come, in media, gli elettori astensionisti siano di sesso maschile, centro-meridionali, più giovani e meno istruiti rispetto agli elettori attivi (Tav. 2). Utilizzando le informazioni relative alle opinioni sul sistema politico, i partiti e la democrazia, nonché le risposte riguardanti il livello di informazione politica e di interesse politico degli intervistati, si è potuto costruire una serie di variabili attitudinali, che si crede possano influenzare le decisioni di partecipazione/astensione al voto:

- *Informazione politica*, calcolata come il numero di risposte corrette fornite dall'intervistato su tre domande riguardanti il sistema politico nazionale¹⁷;
- *Fiducia nei partiti, Fiducia nelle istituzioni* (Parlamento e Presidente della Repubblica), rappresentata dall'auto-posizionamento dell'intervistato lungo una scala da 1 (minima fiducia) a 4 (massima fiducia);
- *Complessità della politica*, catturata dal parere dell'intervistato in merito alla seguente affermazione: «Talvolta la politica sembra così complicata che non si riesce a capire che cosa sta succedendo»;
- *Disillusione verso la politica*, ovvero la valutazione dell'intervistato circa l'inefficacia della politica; il dato deriva dal grado di approvazione dell'affermazione: «Che governi la destra o la sinistra le cose non cambiano»;
- *Disillusione verso la democrazia*, vale a dire la percezione di essere ininfluenza sulle decisioni politiche del Governo; tale informazione è contenuta nel parere dell'intervistato in merito all'affermazione: «La gente come me non ha alcuna influenza su quello che fa il Governo»;
- *Interesse nella politica*; tale variabile misura il grado di interesse dell'intervistato nella politica ed è determinata dall'auto-posizionamento dell'intervistato lungo una scala da 1 (minimo interesse) a 4 (massimo interesse);
- le ultime due variabili mirano a catturare l'influenza sul comportamento degli elettori dell'esposizione ai mass media:
 - *Televisione* è una variabile binaria (*dummy*) che assume valore 1 se l'intervistato dichiara di aver guardato i dibattiti politici o i programmi televisivi di informazione politica, e 0 altrimenti;

¹⁷ In particolare, le tre domande si riferiscono al Presidente della Camera dei Deputati, il Ministro degli Esteri e all'elezione del Presidente della Repubblica.

TAV. 2

AFFLUENZA DEGLI ELETTORI INTERVISTATI, CARATTERISTICHE
SOCIO-DEMOGRAFICHE E VARIABILI ATTITUDINALI - STATISTICA DESCRITTIVA

	<i>Media (Deviazione Standard)</i>		
	<i>Elettori attivi</i> N=2.569	<i>Astensionisti</i> N=354	<i>Tot. elettori</i> N=2.923
Età	2,02 (0,69)	1,97 (0,76)	2,013 (0,70)
=1 per 18-34			
=2 per 35-64			
=3 per ≥ 65			
Sesso	0,47 (0,49)	0,52 (0,5)	0,48 (0,49)
=1 maschile			
=0 femminile			
Istruzione (da 0 a 4)	1,77 (1,3)	1,65 (1,34)	1,76 (1,31)
= 0 nessuna			
= 4 formazione universitaria			
Zona geografica	3,07 (1,58)	3,23 (1,59)	3,09 (1,58)
=1 Nord-Ovest			
=2 Nord-Est			
=3 Centro Nord			
=4 Centro Sud			
=5 Sud e Isole			
Informazione politica (da 0 a 3)	1,36 (1,18)	1,19 (1,21)	1,34 (1,19)
= 0 nessuna risposta corretta			
= 3 tutte le risposte corrette			
Fiducia nei partiti (da 1 a 4)	1,98 (0,69)	1,66 (0,64)	1,94 (0,69)
Fiducia nelle istituzioni (da 1 a 4)	3,08 (0,88)	2,75 (0,96)	3,04 (0,89)
Complessità della politica (da 1 a 4)	3,23 (0,97)	3,19 (1,02)	3,22 (0,97)
Disillusione verso la politica (da 1 a 4)	2,45 (1,07)	2,91 (1,05)	2,5 (1,07)
Disillusione verso la democrazia (da 1 a 4)	2,9 (1,05)	3,1 (1,01)	2,93 (1,05)
Interesse nella politica (da 1 a 4)	2,27 (0,88)	1,93 (0,88)	2,23 (0,89)
Televisione	0,83 (0,37)	0,65 (0,47)	0,81 (0,39)
= 1 ha guardato programmi politici e/o dibattiti televisivi			
= 0 altrimenti			
Quotidiani	0,63 (0,48)	0,46 (0,49)	0,61 (0,49)
= 1 ha letto articoli di politica			
= 0 altrimenti			

- *Quotidiani* è una variabile binaria che assume valore 1 se l'intervistato dichiara di aver letto articoli riguardanti i temi politici principali della campagna elettorale e 0 altrimenti¹⁸.

In generale gli astensionisti presentano, rispetto agli elettori attivi, un minor livello di informazione politica, fiducia nei partiti e nelle istituzioni, interesse nella politica e una minore esposizione ai contenuti politici dei mass media. Inoltre, l'astensionismo è direttamente legato alla disillusione verso la politica e la democrazia. I dati suggeriscono che il livello medio di complessità della politica percepita è elevato, in particolare tra gli elettori attivi. Ciò può derivare dalla maggiore attenzione e interesse per i temi politici da parte degli elettori attivi che, dovendo scegliere quale candidato votare, valutano attentamente le differenti politiche proposte in campagna elettorale apprezzandone, pertanto, la complessità.

Al fine di esaminare gli effetti sulla scelta di partecipazione/astensione al voto della dimensione ideologica, *partisanship*, caratteristiche degli elettori (socio-demografiche, di percezione del sistema politico ed esposizione ai mass-media), applichiamo il modello di stima indifferenza-alienazione nelle sue principali equazioni: l'utilità dell'elettore e i valori soglia individuali di indifferenza e alienazione. La decisione di partecipazione e la decisione su quale candidato/partito votare sono incluse simultaneamente, per cui ogni cittadino ha cinque scelte alternative: astenersi o votare uno dei quattro candidati principali. Formalmente, ogni elettore intervistato è rappresentato dal vettore:

$$(V^1, V^2, V^3, V^4, NV)$$

dove V^j , con $j=1, \dots, 4$, assume valore 1 se l'elettore dichiara di aver votato il candidato j , e 0 altrimenti; analogamente, la variabile binaria NV assume valore 1 se l'elettore dichiara di non aver votato e 0 altrimenti. Pertanto, il vettore contiene una sola variabile con valore 1, che costituisce la scelta dell'elettore. Le equazioni relative all'utilità dell'elettore e ai due valori soglia di indifferenza e alienazione, rispettivamente la (1), (4) e (5), sono di seguito specificate:

¹⁸ Un'ampia parte della letteratura analizza le distorsioni introdotte dai mass media (*media bias*) e i loro effetti sul comportamento degli elettori. Tra gli altri, si veda GROSECLOSE T. - MILYO J. (2005); DELLA VIGNA S. - KAPLAN E. (2007); GEORGE L. - WALDFOGEL J. (2006); GERBER A.S. *et al.* (2009); DURANTE R. - KNIGHT B.G. (2009). Tuttavia, gran parte dei lavori citati si concentrano sull'orientamento ideologico degli elettori e non sulla propensione alla partecipazione al voto che, invece, rappresenta uno dei principali temi d'interesse del presente lavoro.

$$(10) \quad U_i^j = \beta_{j_0} + \beta_{j_1}(\textit{partisanship}) + \beta_{j_2}(\textit{valenza del candidato}) \\ + \beta_{j_3}(\textit{distanza ideologica}) + \varepsilon_i^j \\ i=1, \dots, 4$$

$$(11) \quad I_i = \exp[\beta_4(\textit{Età}) + \beta_5(\textit{Sesso}) + \beta_6(\textit{Istruzione}) \\ + \beta_7(\textit{Zona}) + \beta_8(\textit{Informazione politica}) + \beta_9(\textit{Fiducia nei partiti}) \\ + \beta_{10}(\textit{Fiducia nelle istituzioni}) + \beta_{11}(\textit{Complessità della politica}) \\ + \beta_{12}(\textit{Disillusione verso la politica}) \\ + \beta_{13}(\textit{Disillusione verso la democrazia}) + \beta_{14}(\textit{Interesse politico}) \\ + \beta_{15}(\textit{Televisione}) + \beta_{16}(\textit{Quotidiani})]$$

$$(12) \quad A_i = \beta_{17}(\textit{Età}) + \beta_{18}(\textit{Sesso}) + \beta_{19}(\textit{Istruzione}) + \beta_{20}(\textit{Informazione politica}) \\ + \beta_{21}(\textit{Fiducia nei partiti}) + \beta_{22}(\textit{Fiducia nelle istituzioni}) \\ + \beta_{23}(\textit{Complessità della politica}) \\ + \beta_{24}(\textit{Disillusione verso la politica}) \\ + \beta_{25}(\textit{Disillusione verso la democrazia}) + \beta_{26}(\textit{Interesse politico}) \\ + \beta_{27}(\textit{Televisione}) + \beta_{28}(\textit{Quotidiani}) + \varepsilon_i^A$$

In particolare, la valenza è costituita dal punteggio, tra 1 e 10, che ogni intervistato attribuisce ai diversi candidati e l'identificazione con il partito (*partisanship*) è una variabile *dummy*, con valore 1 se l'intervistato dichiara di sentirsi vicino al partito e 0 altrimenti. Per poter garantire l'identificazione del modello è necessario normalizzare l'intercetta β_{j_0} per uno dei candidati (Casini). Inoltre, si è scelto di non imporre l'uguaglianza dei coefficienti relativi ai diversi candidati (intercetta, valenza e distanza ideologica). Ciò consente di esaminare il diverso impatto sull'utilità dei cittadini della valenza e della distanza ideologica tra i candidati¹⁹.

¹⁹ Le stime effettuate sotto il vincolo di uguali parametri tra i candidati, ovvero:

$$(\beta_{j_0}, \beta_{j_1}, \beta_{j_2}, \beta_{j_3}) = (\beta_{k_0}, \beta_{k_1}, \beta_{k_2}, \beta_{k_3}) \quad \forall k \neq j, j = 1, \dots, 4$$

non hanno prodotto sostanziali differenze nei risultati. Per poter ottenere un'analisi esaustiva, si è preferito evitare l'imposizione di tale vincolo.

4. - I risultati

La Tavola 3 riporta le stime dei parametri e gli errori *standard*. Il *test Wald* indica una buona significatività congiunta delle variabili incluse nel modello (all'1%) e la percentuale di previsioni corrette (PCP) riporta una buona capacità esplicativa delle variabili considerate.

Diversi parametri sono statisticamente significativi e mostrano il segno atteso. In particolare, in linea con i modelli teorici del voto, l'utilità degli elettori si riduce all'aumentare della distanza ideologica dal candidato. Inoltre, entrambi i fattori non-spaziali, la valenza (percepita) del candidato e l'identificazione con il partito, influenzano positivamente l'utilità; i coefficienti sono statisticamente significativi all'1% e ciò può spiegare, in parte, il motivo per cui gli elettori possono votare un candidato diverso da quello ideologico. Nel valutare i diversi candidati l'elettore tiene conto non solo di aspetti ideologici e di *partisanship* (la distanza ideologica ed il senso di appartenenza al partito) ma anche della valenza (percepita) del candidato, per cui il più vicino all'orientamento ideologico dell'elettore può risultare meno attraente rispetto ad un altro candidato.

Le stime dei parametri mostrano le caratteristiche individuali determinanti l'astensione per indifferenza o alienazione. Limitandoci alle sole relazioni statisticamente significative, la probabilità di astensione per indifferenza aumenta con il senso di disillusione dalla politica e con l'informazione dell'elettore. Sembra dunque che quanto più l'elettore sia informato sui temi legati al sistema politico, più ampia debba essere la differenza percepita tra i candidati per indurlo a votare.

L'informazione rende i cittadini scettici riguardo l'effettiva differenza tra i candidati. L'età e la fiducia nelle istituzioni riducono la probabilità di astensione per indifferenza.

L'età influenza invece positivamente l'utilità di aspirazione dei cittadini e, quindi, l'astensione per alienazione, così come la provenienza da regioni meridionali. L'interesse mostrato nei confronti dei temi politici ha un effetto negativo sull'utilità di aspirazione e, pertanto, influenza positivamente la predisposizione al voto. È sorprendente notare come il senso di disillusione dalla politica riduca la probabilità di alienazione, mentre aumenti la probabilità di indifferenza. La delusione dei cittadini, principalmente dovuta al clima di instabilità politica, si traduce da un lato nell'aspirazione ad un più basso livello minimo di utilità ottenibile e dall'altro in una maggiore sensibilità alle differenze tra i candidati.

Infine, relativamente all'influenza dei mass media, si rileva come l'esposizione ai programmi televisivi politici riduca significativamente l'utilità di aspirazione e, quindi, l'inclinazione ad essere alienati.

TAV. 3

IL MODELLO INDIFFERENZA-ALIENAZIONE NELLE ELEZIONI POLITICHE
ITALIANE DEL 2008 STIME DEI PARAMETRI
(ERRORI STANDARD)

	<i>Parametri Candidato</i>	<i>Valore soglia Indifferenza</i>	<i>Valore soglia Alienazione</i>
Età		-0,62*** (0,203)	0,691** (0,34)
Sesso		0,188 (0,231)	0,509 (0,336)
Zona geografica		-0,068 (0,068)	0,241** (0,096)
Istruzione		0,038 (0,089)	-0,151 (0,162)
Informazione politica		0,202* (0,12)	-0,068 (0,207)
Fiducia nei partiti		-0,177 (0,167)	-0,381 (0,239)
Fiducia nelle istituzioni		-0,159* (0,1)	-0,12 (0,149)
Complessità della politica		-0,142 (0,095)	-0,181 (0,158)
Disillusione verso la politica		0,275** (0,122)	-0,275* (0,158)
Disillusione verso la democrazia		0,099 (0,098)	0,070 (0,151)
Interesse politico		-0,038 (0,137)	-0,389* (0,224)
Televisione		-0,16 (0,268)	-0,766** (0,348)
Quotidiani		-0,245 (0,218)	-0,018 (0,366)
Intercetta		-0,17 (0,844)	5,232*** (1,372)
Valenza del candidato:			
Bertinotti	0,303*** (0,033)		
Veltroni	0,52*** (0,035)		
Casini	0,664*** (0,076)		
Berlusconi	0,449*** (0,032)		
Distanza ideologica:			
Bertinotti	-0,034*** (0,006)		
Veltroni	-0,061*** (0,010)		
Casini	-0,154*** (0,033)		
Berlusconi	-0,054*** (0,006)		
Identificazione partito:			
Bertinotti	0,953*** (0,268)		
Veltroni	2,319*** (0,192)		
Casini	3,62*** (0,428)		
Berlusconi	1,857*** (0,239)		
Intercetta:			
Bertinotti	3,891*** (0,564)		
Veltroni	2,786*** (0,566)		
Berlusconi	3,417*** (0,567)		

Log-likelihood (N=2.923) -2.416,222

 R^2 : 0,38Wald *test* (sui coefficienti dei valori soglia di indifferenza e alienazione):Wald Chi² (con 28 g.d.l.): 337,85***% previsioni corrette PCP¹: 69,24%; PRE (sul modello nullo)²: 52,43%

* significativo al 10%, **significativo al 5%, ***significativo all'1%.

¹ La percentuale di previsioni corrette è calcolata come il rapporto tra il numero di previsioni corrette ed il numero totale di previsioni.² Il modello nullo prevede che la variabile dipendente per ogni osservazione assuma il valore del risultato più comune nel campione.

Dai coefficienti stimati si determinano le probabilità che ogni elettore intervistato abbia votato per uno dei quattro candidati o si sia astenuto per alienazione, indifferenza o entrambi. Usando le medie delle suddette probabilità come risultati attesi, si ricava la frazione media di elettori che rientrano in ciascuna categoria. La Tavola 4 riporta le percentuali di voto e astensione osservate e stimate. È facile constatare come le previsioni corrispondano al comportamento aggregato nel campione di elettori.

TAV. 4

VOTI E TASSO DI ASTENSIONE OSSERVATI E STIMATI

	Osservato (N=2.923)	Stimato (N=2.923)
Voto Bertinotti	14,1%	14,08% (10; 18,15)
Voto Veltroni	35,1%	35,07% (33,84; 36,3)
Voto Casini	3,4%	3,33% (2,62; 4,03)
Voto Berlusconi	35,34%	35,31% (33,86; 36,75)
Astensione	12,11%	12,22% (10,71; 13,72)

Nota: gli intervalli di confidenza del 95% sono riportati in parentesi.

La Tavola 5 distingue gli elettori a seconda delle caratteristiche individuali e ne riporta la frazione media stimata di astensionisti, suddivisa nella componente di alienazione, indifferenza, alienazione e indifferenza.

TAV. 5

STIMA DELL'ASTENSIONISMO PER CAUSA E CARATTERISTICHE INDIVIDUALI

	<i>Alienazione</i>	<i>Indifferenza</i>	<i>Alienazione e indifferenza</i>	<i>Astensione</i>
Totale (N=2.923)	4,86 (2,4; 7,3)	6,43 (4,12; 8,75)	0,93 (0,56; 1,29)	12,22 (10,71; 13,72)
Età				
18-34 (N=706)	1,74 (0,05; 4,21)	11,58 (7,04; 16,13)	0,97 (0,09; 1,84)	14,29 (11,77; 16,82)
≥ 65 (N=744)	8,85 (4,9; 12,8)	2,44 (0,52; 4,36)	0,85 (0,31; 1,38)	12,14 (9,77; 14,51)
Sesso				
uomo (N=1.397)	4,85 (1,7; 8,01)	7,47 (4,06; 10,88)	1,16 (0,66; 1,67)	13,48 (11,59; 15,39)
donna (N=1.526)	4,86 (2,3; 7,42)	5,48 (3,54; 7,41)	0,71 (0,35; 1,07)	11,05 (8,93; 13,17)
Zona geografica				
Nord-Ovest (N=773)	3,25 (1,04; 5,46)	7,48 (4,92; 10,05)	0,74 (0,31; 1,16)	11,47 (9,59; 13,35)
Sud e Isole (N=862)	7,14 (3,25; 11,03)	5,7 (2,46; 8,94)	1,23 (0,73; 1,73)	14,07 (11,28; 16,86)
Istruzione				
nessuna (N=519)	10,48 (5,55; 15,4)	3,16 (0,92; 5,4)	1,25 (0,53; 1,95)	14,89 (11,31; 18,46)
formazione universitaria (N=304)	2,21 (0,39; 4,82)	8,83 (5,07; 12,59)	0,64 (0,08; 1,2)	11,68 (9,22; 14,15)
Informazione politica				
Minima (N=1.023)	7,53 (3,71; 11,34)	5,35 (2,61; 8,09)	1,18 (0,7; 1,66)	14,06 (10,76; 17,37)
Massima (N=725)	2,88 (0,48; 5,28)	7,2 (4,18; 10,23)	0,77 (0,3; 1,24)	10,85 (9,02; 12,68)
Fiducia nei partiti				
Minima (N=746)	8,37 (3,51; 13,24)	9,27 (4,47; 14,08)	2,05 (1,22; 2,88)	19,69 (16,63; 22,77)
Massima (N=44)	1,43 (0,8; 3,66)	2,23 (0,62; 3,84)	0,12 (0,01; 0,29)	3,78 (1,64; 5,93)
Fiducia nelle istituzioni				
Minima (N=215)	9,07 (2,65; 15,48)	11,24 (4,41; 18,07)	2,96 (1,45; 4,47)	23,27 (19,23; 27,3)
Massima (N=1.011)	3,26 (1,39; 5,13)	4,59 (2,83; 6,36)	0,44 (0,23; 0,64)	8,29 (6,7; 9,89)
Complessità politica				
Minima (N=277)	4,49 (0,49; 8,48)	8,76 (3,9; 13,61)	1,64 (0,59; 2,68)	14,89 (11,77; 18)
Massima (N=1.513)	5,58 (2,84; 8,33)	5,68 (3,32; 8,05)	0,88 (0,53; 1,23)	12,14 (9,8; 14,5)
Disillusione politica				
Minima (N=725)	3,77 (1,7; 5,84)	2,65 (0,93; 4,36)	0,36 (0,18; 0,55)	6,78 (5,37; 8,19)

./.

	<i>Alienazione</i>	<i>Indifferenza</i>	<i>Alienazione e indifferenza</i>	<i>Astensione</i>
Massima (N=634)	5,87 (1,78; 9,95)	11,63 (6,9; 16,37)	1,93 (1; 2,85)	19,43 (16,73; 22,13)
Disillusione democrazia Minima (N=383)	3,48 (0,69; 6,28)	4,96 (2,04; 7,89)	0,58 (0,2; 0,97)	9,02 (7,02; 11,04)
Massima (N=1.140)	5,55 (2,46; 8,63)	7,55 (4,35; 10,74)	1,28 (0,78; 1,79)	14,38 (12,16; 16,6)
Interesse politico Minima (N=718)	10,28 (5,53; 15,03)	6,68 (3,14; 10,23)	1,91 (1,17; 2,65)	18,87 (14,65; 23,1)
Massima (N=179)	1,72 (0,07; 3,53)	5,67 (2,87; 8,47)	0,41 (0,05; 0,86)	7,8 (5,72; 9,88)
Televisione Sì (N=2.365)	3,35 (1,52; 5,18)	5,98 (4,01; 7,95)	0,54 (0,29; 0,8)	9,87 (8,62; 11,14)
No (N=558)	11,24 (4,8; 17,69)	8,34 (2,96; 13,72)	2,54 (1,5; 3,57)	22,12 (17; 27,41)
Quotidiani Sì (N=1.784)	3,00 (1,06; 4,93)	5,80 (3,59; 8,02)	0,54 (0,25; 0,84)	9,34 (8,06; 10,63)
No (N=1.139)	7,77 (3,9; 11,64)	7,42 (4,11; 10,73)	1,52 (0,95; 2,09)	16,71 (13,61; 19,81)

Nota: gli intervalli di confidenza del 95% sono riportati in parentesi.

Le stime suggeriscono che sia l'alienazione che l'indifferenza determinano l'astensionismo nel campione di elettori, con il contributo dell'indifferenza nettamente superiore rispetto all'alienazione (il 52,62% contro il 39,77%). La componente di astensione per alienazione e indifferenza è invece marginale (7,61%). In generale l'indifferenza costituisce la fonte principale di astensione tra i giovani, gli elettori delle regioni nord-occidentali, i laureati, i più informati, interessati alla politica ed esposti ai mass media. L'alienazione è invece dominante tra gli anziani, gli elettori del sud e delle isole, gli elettori meno istruiti, informati, interessati ed esposti ai media. L'indifferenza è la fonte principale di astensione indipendentemente dal sesso e dalla fiducia riposta nei partiti o nelle istituzioni. Tuttavia, il contributo relativo dell'indifferenza aumenta con il grado di fiducia degli elettori, specialmente nei confronti dei partiti.

Coerentemente ai risultati delle statistiche descrittive, la valutazione della politica come materia complessa migliora la predisposizione al voto. L'indifferenza contribuisce all'astensione in misura nettamente superiore tra gli elettori che considerano la politica meno complicata. Nessuna sorpresa viene dalle variabili espressive della

delusione dei cittadini: l'astensione aumenta tra i disillusi e la causa principale è l'indifferenza. L'alienazione, tuttavia, domina tra gli elettori meno delusi dalla politica: il che conferma la relazione negativa tra disillusione e utilità di aspirazione, evidenziata dalle stime dei parametri. Infine, l'esposizione ai mass media migliora la predisposizione al voto. Il mezzo televisivo presenta un maggiore impatto sull'astensione, che aumenta del 124% tra i non-esposti contro il 44% dei quotidiani.

La Tavola 6 distingue gli elettori a seconda del candidato ideologico e ne riporta la frazione media (stimata) di astensionisti, suddivisa nella componente di alienazione, indifferenza, alienazione e indifferenza. Le percentuali stimate di elettori astensionisti nel bacino di attrazione di ogni candidato sono coerenti con i tassi di astensione osservati (Graf. 2). Infatti, l'astensionismo più elevato si registra in corrispondenza dei candidati F. Bertinotti e P.F. Casini, mentre la percentuale più bassa interessa il bacino di attrazione del candidato S. Berlusconi. Per quanto riguarda la causa di astensione, l'alienazione è dominante tra gli elettori di estrema sinistra (il 47,52% contro il 44,67% dell'indifferenza). L'indifferenza costituisce, invece, la maggiore fonte di astensione per gli elettori ideologici del candidato di centro-sinistra (60,63%), del candidato centrista (72,81%) e di centro-destra (il 54,11%). La più alta incidenza dell'alienazione si riscontra tra l'elettorato dei due candidati più esterni lungo la scala ideologica, Bertinotti (47,52%) e Berlusconi (40,18%), contro il 32,79% relativo a Veltroni e 21,87% per Casini.

TAV. 6

STIMA DEL TASSO DI ASTENSIONE PER CANDIDATO IDEOLOGICO E CAUSA

<i>Candidato ideologico</i>	<i>Alienazione</i>	<i>Indifferenza</i>	<i>Alienazione e indifferenza</i>	<i>Tasso di astensione</i>
Bertinotti (N=412)	8,16 (1,87; 14,44)	7,67 (4,68; 10,66)	1,34 (0,59; 2,1)	17,17 (13,41; 20,93)
Veltroni (N=1.025)	3,24 (1,43; 5,06)	5,99 (3,93; 8,04)	0,65 (0,35; 0,94)	9,88 (8,84; 10,92)
Casini (N=99)	2,55 (0,93; 4,18)	8,49 (5,73; 11,24)	0,62 (0,31; 0,93)	11,66 (9,88; 13,44)
Berlusconi (N=1.033)	3,52 (1,72; 5,32)	4,74 (3,05; 6,43)	0,5 (0,32; 0,68)	8,76 (7,78; 9,75)

Nota: gli intervalli di confidenza del 95 % sono riportati in parentesi.

Tali risultati confermano le teorie sul voto: l'astensione per alienazione aumenta con la distanza dal candidato più vicino mentre l'astensione per indifferenza aumenta al ridursi della differenza tra la distanza dal candidato più vicino e lontano (come in Katz, 2007). Ciò implica l'esistenza di una dimensione squi-

sitamente ideologica dell'astensionismo, vale a dire l'impatto relativo di alienazione e indifferenza dipende dalle posizioni (percepite) dei candidati. In altre parole il posizionamento dei candidati influenza la causa dell'astensione.

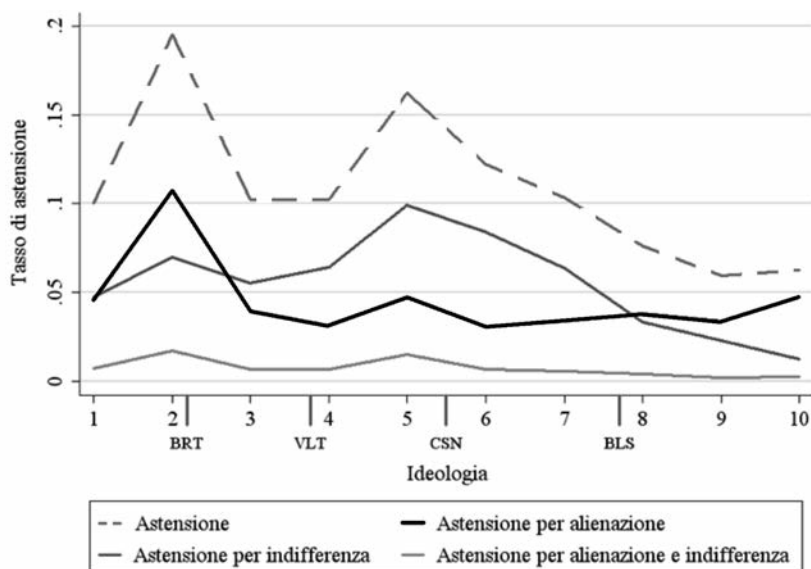
Infine, l'astensione causata sia da alienazione che indifferenza ha un ruolo marginale, soprattutto per i principali candidati, e raggiunge il suo massimo tra gli elettori ideologici del candidato Bertinotti, costituendo il 7,8% dell'astensione totale.

Al fine di esaminare la relazione tra l'orientamento ideologico dei cittadini, la causa di astensione e il posizionamento dei candidati, il Grafico 4 mostra il tasso di astensione stimato per ogni punto della scala ideologica e distinto a seconda della causa di astensione: indifferenza, alienazione o entrambe.

Gli intervistati la cui politica ideale appartiene all'intervallo [3,8] si astengono principalmente a causa dell'indifferenza, mentre gli elettori posizionati nelle regioni periferiche/estreme della scala ideologica si astengono soprattutto per alienazione.

GRAF. 4

TASSO DI ASTENSIONE STIMATO PER CAUSA E IDEOLOGIA



Nota: Il grafico riporta il tasso di astensione, calcolato come la percentuale stimata di astensionisti collocati nel punto della scala ideologica. Le frecce indicano le posizioni (percepite) dei diversi candidati (BRT=F. Bertinotti, VLT=W. Veltroni, CSN=P.F. Casini, BLS=S. Berlusconi).

Tra gli elettori con orientamento ideologico di centro-destra il maggiore impatto dell'astensione per indifferenza è netto. Ciò può essere dovuto, in parte, ad una minore inclinazione all'alienazione di tale elettorato e, in parte, alla percezione

di una maggiore affinità ideologica tra i candidati: poiché il partito, l'UDC, del candidato centrista e il partito, Forza Italia, si erano presentati alle elezioni politiche del 2006 nella stessa coalizione (il Popolo della Libertà), la probabilità che un elettore sia indifferente tra Casini e Berlusconi è sensibilmente più elevata rispetto alla probabilità di esserlo tra lo stesso Casini e il candidato di centro-sinistra Veltroni. L'indifferenza è la principale causa di astensione tra gli elettori moderati per almeno due possibili ragioni:

- al fine di effettuare la propria scelta, l'elettore moderato valuta almeno tre candidati: i due principali avversari e il candidato centrista; pertanto la probabilità di essere confuso, ovvero indifferente, aumenta;
- in linea con la teoria spaziale del voto, l'astensione per indifferenza aumenta in prossimità del *punto di mezzo*, vale a dire dell'elettore posizionato esattamente tra due candidati, che è indifferente per definizione. Infatti, l'astensione per indifferenza più elevata si registra al centro, in prossimità del punto di mezzo tra i due principali avversari, il candidato di centro-sinistra e di centro-destra.

Coerentemente ai modelli teorici sul voto, tra gli elettori periferici l'astensione per indifferenza si riduce all'aumentare della distanza ideologica dai due candidati principali. Ciò si verifica nell'intervallo [6,10], mentre nell'elettorato di centro-sinistra, oltre ad un calo dell'astensione per indifferenza meno marcato, si rileva un suo incremento in prossimità del candidato di estrema sinistra. Da un lato, la maggiore predisposizione all'indifferenza dell'elettorato di centro-sinistra è interpretabile come una forma di astensione punitiva o di protesta a seguito della delusione per il fallimento del governo Prodi; dall'altro, la mancata alleanza con il partito estremista fa sì che gli elettori orientati ideologicamente a sinistra si trovino di fronte a due candidati diversi, per cui la probabilità di essere indifferente aumenta significativamente. Comparando l'astensione per indifferenza nell'intervallo [1,4] con l'intervallo [6,10], è evidente come la presenza di un ulteriore candidato di sinistra aumenti la predisposizione all'indifferenza.

L'astensione tra gli elettori collocati all'esterno della regione centrale/moderata è principalmente guidata dall'alienazione, seppur con sostanziali differenze tra l'elettorato di estrema sinistra e destra. Nella regione periferica di destra l'astensione per alienazione aumenta con la distanza ideologica dal candidato più vicino (S. Berlusconi)²⁰. Invece, l'astensione tra gli elettori di estrema sinistra si riduce

²⁰ ADAMS J. *et al.* (2006), studiando l'astensione per indifferenza e alienazione nelle elezioni presidenziali americane 1980-1988, rilevano che «le tendenze dei cittadini ad astenersi per alienazione sono fortemente legate alle distanze percepite dai candidati. Tra i cittadini sostenitori

all'aumentare della distanza dal candidato F. Bertinotti. Tale risultato sembra incoerente con la teoria spaziale del voto per cui una maggiore distanza dal candidato aumenta il disinteresse del cittadino e, dunque, la probabilità di astenersi per alienazione. Tuttavia, questo è vero se tutti gli individui hanno lo stesso livello di utilità di aspirazione, ovvero se il parametro di tolleranza è uguale per tutti i cittadini²¹ ma, se gli elettori presentano utilità di aspirazione diverse e, dunque, la tolleranza è eterogenea, la probabilità di astensione dipende sia dalla distanza ideologica sia dalla tolleranza individuale; un elettore, sebbene ideologicamente vicino al candidato, può essere così poco tollerante da non votare, così come un elettore distante dal candidato può scegliere di votare se particolarmente tollerante. Il picco raggiunto dall'astensione per alienazione in corrispondenza della posizione ideologica del candidato di estrema sinistra deriva dalla scarsa tolleranza del suo elettorato, a conferma dell'astensione selettiva che ha colpito particolarmente tale elettorato.

5. - Conclusioni

Il lavoro implementa un modello di stima probabilistico in base al quale l'elettore sceglie se e (eventualmente) quale candidato votare in base alla propria utilità. Coerentemente al *background* teorico di riferimento, l'elettore si astiene se indifferente tra i candidati o alienato. Nel primo caso non c'è alcun candidato così diverso dagli altri da garantire all'elettore un surplus sufficientemente elevato. Nel secondo caso l'elettore è disinteressato in quanto nessun candidato lo rappresenta adeguatamente, ovvero è tale da assicurargli un livello minimo di utilità (utilità di aspirazione). Entrambi i valori soglia sono influenzati dalle caratteristiche individuali socio-demografiche, attitudinali, di percezione del sistema politico, nonché dall'esposizione ai mass-media. Il modello consente di andare oltre la semplice individuazione delle variabili predittive del (non) voto e di fornire interessanti spunti di riflessione riguardo l'eterogeneità dei valori soglia individuali e la relazione esistente tra l'astensione per alienazione o indifferenza e la distanza ideologica (percepita) dai candidati.

del partito Repubblicano la propensione ad astenersi per alienazione passa dal 30% circa per coloro che condividono il posizionamento ideologico del candidato George H.W. Bush a oltre il 60% per i cittadini più liberali e distanti da Bush».

²¹ Ciò è quanto assunto nei modelli di LLAVADOR H. (2000) e LEPPLE K. (2009).

Le stime relative alle elezioni politiche italiane del 2008 suggeriscono che le due fonti di astensione hanno determinanti diverse e che sono costituite in parte da fattori socio-strutturali, in parte da variabili attitudinali e di percezione del sistema politico, particolarmente sensibili al contesto socio-economico. In particolare, la probabilità di astensione per indifferenza è significativamente legata all'età, l'informazione politica, la fiducia riposta nelle istituzioni e la delusione dalla politica. La probabilità di astensione per alienazione dipende dall'età, la provenienza geografica, la delusione, l'interesse politico e l'esposizione ai mezzi di comunicazione televisiva.

In linea con la teoria spaziale del voto, la partecipazione al voto mostra un legame inverso con la distanza ideologica dal candidato. Tuttavia, il comportamento finale dell'elettore, vale a dire la scelta del candidato cui assegnare (eventualmente) il proprio voto, è influenzato sia dal fattore spaziale, la distanza ideologica, che da fattori non-spaziali, quali la valenza del candidato e l'identificazione con il partito di riferimento (*partisanship*).

I risultati più interessanti riguardano il legame tra il peso relativo dell'astensione per indifferenza o alienazione e la posizione (percepita) dei candidati. In linea con gli studi teorici sui modelli di competizione politica spaziale, la più alta componente di astensione per indifferenza si rileva al centro della scala ideologica, in prossimità del punto di mezzo tra i due principali avversari, di centro-sinistra e centro-destra. Inoltre, l'astensione per alienazione si riduce in prossimità del candidato ideologico; tuttavia, aumenta in presenza di elevati livelli di utilità di aspirazione. Tali considerazioni evidenziano un maggiore impatto della dimensione ideologica sull'inclinazione degli elettori all'indifferenza; le tendenze individuali all'astensione per alienazione, invece, sono fortemente influenzate dai propri valori soglia, a conferma delle complesse relazioni alla base del comportamento elettorale e dell'importanza dell'eterogeneità nella dimensione ideologica così come nei valori soglia di indifferenza e alienazione.

BIBLIOGRAFIA

- ADAMS J. - DOW J. - MERRILL III S., *The Political Consequences of Abstention Due to Alienation and Indifference: Applications to Presidential Elections*, Paper presentato all'incontro annuale di American Political Science Association, San Francisco, 2001.
- . -, «The Political Consequences of Alienation Based and Indifference-based Voter Abstention: Applications to Presidential Elections», *Political Behaviour*, no. 28(1), 2006, pages 65-86.
- ADAMS J. - MERRILL III S., «Voter Turnout and Candidate Strategies in American Elections», *Journal of Politics*, no. 65(1), 2003, pages 161-189.
- ADAMS J. - MERRILL III S. - GROFMAN B., *A Unified Theory of Party Competition. A Cross-National Analysis Integrating Spatial and Behavioral Factors*, New York, Cambridge University Press, 2005.
- ANSOLABEHERE S. - SNYDER J., «Valence Politics and Equilibrium in Spatial Elections Models», *Public Choice*, no. 103(3-4), 2000, pages 327-336.
- BAKER K. - DALTON R.J. - HILDEBRANDT K., *Germany Transformed: Political Culture and the New Politics*, Cambridge, MA, Harvard University Press, 1981.
- BARTLE J., «Left-Right Position Matters, but Does Social Class? Causal Models of the 1992 British General Election», *British Journal of Political Science*, no. 28(3), 1998, pages 501-529.
- BELLUCCI P. - SEGATTI P., *Votare in Italia. Dall'appartenenza alla scelta*, Bologna, Il Mulino, 2011.
- BORGE L. - FALCH T. - TOVMO P., «Public Sector Efficiency: The Roles of Political and Budgetary Institutions, Fiscal Capacity, and Democratic Participation», *Public Choice*, no. 136(3-4), 2008, pages 475-495.
- BURDEN B. - LACY D., *The Vote-Stealing and Turnout Effects of Third-Party Candidates in US Presidential Elections, 1968-1996*, Paper presentato all'incontro annuale di American Political Science Association, Atlanta, 1999.
- CAMPBELL A. - CONVERSE P. - MILLER W. - STOKES D.E., *The American Voter*, New York, John Wiley and Sons, 1960.
- CONVERSE P. - PIERCE R., *Political Representation in France*, Cambridge, MA, Harvard University Press, 1986.
- . -, «Comment on Fleury and Lewis - Beck», *Journal of Politics*, no. 55(04), 1993, pages 1110-1117.
- DOWNES A., *An Economic Theory of Democracy*, New York, Harper and Row, 1957.
- DEE T.S., «Are There Civic Returns to Education?», *Journal of Public Economics*, no. 88(9-10), 2004, pages 1697-1720.
- DELLAVIGNA S. - KAPLAN E., «The Fox News Effect: Media Bias and Voting», *The Quarterly Journal of Economics*, no. 122(3), 2007, pages 1187-1234.

- DURANTE R. - KNIGHT B.G., «Partisan Control, Media Bias, and Viewer Responses: Evidence from Berlusconi's Italy», *NBER Working Paper*, no. 14762, 2009.
- ENDERSBY J. - GALATAS S., *British Parties and Spatial Competition: Dimensions of Evaluation in the 1992 Election*, Paper presentato all'incontro annuale di Public Choice Society, San Francisco, CA, 1997.
- GARZIA D. - VIOTTI F., «Party Identification, Leader Effects and Vote Choice in Italy, 1990-2008», *World Political Science Review*, no. 8(1), 2012, pages 1-23.
- GERBER A.S. - KARLAN D. - BERGAN D., «Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions», *American Economic Journal: Applied Economics*, no. 1(2), 2009, pages 35-52.
- GEORGE L. - WALDFOGEL J., «The New York Times and the Market for Local Newspapers», *American Economic Review*, no. 96(1), 2006, pages 435-477.
- .-, «Who Affects Whom in Daily Newspaper Markets?», *Journal of Political Economy*, no. 111(4), 2003, pages 765-785.
- GEYS B., «Explaining Voter Turnout: A Review of Aggregate-Level Research», *Electoral Studies*, no. 25(4), 2006, pages 637-663.
- GHIRARDATO P. - KATZ J.N., «Indecision Theory: Quality of Information and Voting Behaviour», California Institute of Technology, Division of the Humanities and Social Sciences, *Working Paper*, no. 1106, 2003.
- GROSECLOSE T., «A Model of Candidate Location When One Candidate Has a Valence Advantage», *American Journal of Political Science*, no. 45(4), 2001, pages 862-886.
- GROSECLOSE T. - MILYO J., «A Measure of Media Bias», *The Quarterly Journal of Economics*, no. 120(4), 2005, pages 1191-1237.
- GUIO L. - SAPIENZA P. - ZINGALES L., «Civic Capital as the Missing Link», in BENHABIB J. - BISIN A. - JACKSON M.O. (eds.), *Handbook of Social Economics*, North Holland, Elsevier, vol. 1, 2010, pages 417-480.
- KATZ G., «Policy-Based Abstention in Brazil's 2002 Election», California Institute of Technology, *Social Science Working Paper*, no. 1288, 2007.
- LARCINESE V., «Information Acquisition, Ideology and Turnout: Theory and Evidence from Britain», *Journal of Theoretical Politics*, no. 21(2), 2009, pages 237-276.
- LEPPEL K., «A Note on the Median Voter Theory and Voter Alienation», *The Social Science Journal*, no. 46(2), 2009, pages 369-374.
- LEWIS-BECK M. - CHLARSON K., «Party, Ideology, Institutions, and the 1995 French Presidential Election», *British Journal of Political Science*, no. 32(3), 2002, pages 489-512.
- LIJPHART A., «Unequal Participation: Democracy's Unresolved Dilemma», *American Political Science Review*, no. 91(1), 1997, pages 1-14.
- LLAVADOR H.G., «Abstention and Political Competition», *Review of Economic Design*, no. 5(4), 2000, pages 411-432.

- LLAVAREDA A., *A democracia nas urnas: o processo partitario eleitoral brasileiro*, Rio de Janeiro, Rio Fundo Editora, 1991.
- MATSUSAKA J.G., «Explaining Voter Turnout Patterns: an Information Theory», *Public Choice*, no. 84(1-2), 1995, pages 91-117.
- MILLIGAN K. - MORETTI E. - OREOPULOS P., «Does Education Improve Citizenship? Evidence from the US and the UK», *Journal of Public Economics*, no. 88(9-10), 2004, pages 1667-1695.
- MUELLER D.C., *Public Choice III*, Cambridge University Press, 2003.
- PLANE D.L. - GERSHTENSON J., «Candidates' Ideological Locations, Abstention and Turnout in US Midterm Senate Elections», *Political Behaviour*, no. 26(1), 2004, pages 69-93.
- POWER T. - ROBERTS T., «Compulsory Voting, Invalid Ballots, and Abstention in Brazil», *Political Research Quarterly*, no. 48(4), 1995, pages 795-826.
- PUTNAM R. D., *Making Democracy Work*, Princeton University Press, 1993.
- SANDERS M., «Unified Models of Turnout and Vote Choice for Two-Candidate and Three-Candidate Elections», *Political Analysis*, no. 7(1), 1998, pages 89-115.
- . -, «Uncertainty and Turnout», *Political Analysis*, no. 9(1), 2001, pages 45-57.
- SHEPSLE K.A., *Models of Multiparty Electoral Competition*, Switzerland, Chur, Harwood Academic Publishers, 1991.
- STOKES D., «Spatial Models of Party Competition», *American Political Science Review*, no. 57(2), 1963, pages 368-377.
- THURNER P. - EYMANN A., «Policy-Specific Alienation and Indifference in the Calculus of Voting: A Simultaneous model of party Choice and Abstention», *Public Choice*, no. 102(1-2), 2000, pages 51-77.
- WOLFINGER R.E. - ROSENSTONE S.J., *Who Votes?*, New Haven, Yale University Press, 1980.
- ZIPP J.F., «Perceived Representativeness and Voting: an Assessment of "choices" versus "echoes"», *American Political Science Review*, no. 79(1), 1985, pages 50-61.

La convergenza dello sviluppo finanziario tra le regioni italiane dal 1890 ad oggi[◇]

Adriano Giannola*

Università di Napoli “Federico II”

Antonio Lopes**

Seconda Università di Napoli

Alberto Zazzaro***

Università Politecnica delle Marche

Based on the extensive statistical documentation recently released by Svimez (Association for the development of industry in Southern Italy), in this paper we examine the evolution of banking structures and lending activities across Italian regions during the last 120 years. In particular, we analyze whether Italian regions have experienced a process of financial convergence and whether this process has followed a similar pattern during the whole period or whether it has been influenced by the various regional industrial policy and bank regulatory regimes that have featured Italy from 1890 to 2011.

[JEL Classification: J21; O11; R11]

Keywords: financial convergence; geography of the Italian banking system; development policy.

[◇] Versioni preliminari di questo lavoro sono state presentate al Convegno SIS 2011 (Bologna, 2011) e al workshop della Società Geografica Italiana, “Sistema finanziario e spazio geografico in cambiamento” (Roma, 2011). Desideriamo ringraziare Marcello De Cecco e i partecipanti ai seminari per i commenti ricevuti. Ringraziamo in particolare Andrea Presbitero e Domenico Scalera per i puntuali e preziosi suggerimenti.

* <giannola@unina.it>, Dipartimento di Economia, Management, Istituzioni.

** <antonio.lopes@unina2.it>, Dipartimento di Scienze Politiche “Jean Monnet”.

*** <a.zazzaro@univpm.it>, Dipartimento di Scienze Economiche e Sociali e Money and Finance Research Group (MoFiR, Ancona).

1. - Introduzione

Sui divari di sviluppo economico tra le regioni italiane e sui processi di convergenza nel reddito *pro-capite* e in altre grandezze reali vi è una letteratura molto ricca e consolidata che ha affrontato diversi aspetti relativi alle determinanti e alle dinamiche temporali in una prospettiva di medio, lungo e lunghissimo periodo (Paci e Pigliaru, 1997; Daniele e Malanima, 2007; Fiaschi *et al.*, 2011; Dal Bianco, 2011; Malanima, 2012). Molto meno studiata è, invece, l'evoluzione territoriale dello sviluppo finanziario, e ciò sebbene le differenze nelle strutture e nei flussi finanziari regionali in Italia siano oggi, come in passato, molto evidenti e tutte le analisi empiriche condotte a livello regionale, provinciale e comunale abbiano confermato l'esistenza di un nesso causale tra il grado di sviluppo e di efficienza del sistema bancario e la crescita delle economie locali (Lucchetti *et al.*, 2001; Guiso *et al.*, 2004; Vaona, 2008; Vaona e Patuelli, 2008; Zago e Dongili, 2011). A più riprese, diversi autori hanno analizzato l'evoluzione dei sistemi bancari regionali, con particolare attenzione ai divari finanziari Nord-Sud e agli effetti della ristrutturazione del sistema bancario italiano durante gli anni Novanta (Galli, 1990; Bussetta e Sacco, 2001; Giannola, 2002; Colombo e Turati, 2004; Mattesini e Messori, 2004; Bongini e Ferri, 2005; Cannari e Panetta, 2006; De Bonis *et al.*, 2010; Nuzzo e Olivieri, 2011), ma tutti si sono limitati a proporre un'analisi di medio periodo, senza affrontare il tema della convergenza di lungo periodo dello sviluppo finanziario¹. Altri contributi sono stati dedicati alla ricostruzione dell'evoluzione del sistema bancario italiano nel lungo periodo, ma senza addentrarsi in modo specifico nella questione della dinamica territoriale dello sviluppo finanziario (Biscaini e Ciocca, 1979; Della Torre, 2000; Onado, 2003; De Bonis *et al.*, 2012).

Obiettivo di questo lavoro è esaminare l'evoluzione dell'articolazione territoriale del sistema bancario italiano e dell'attività delle banche dal 1890 ad oggi. Sulla base della ricca documentazione statistica recentemente diffusa dalla Svimez (Svimez, 2011a), verificheremo se nel corso degli ultimi centoventi anni le regioni italiane siano divenute finanziariamente più simili tra loro. In particolare, analizzeremo l'evoluzione delle strutture bancarie e delle politiche di raccolta e di impiego delle banche e cercheremo di stabilire se l'eventuale processo di convergenza finanziaria abbia seguito un percorso omogeneo nelle diverse fasi storiche

¹ Va detto che anche a livello *cross-country* le analisi sulla convergenza dello sviluppo finanziario e della struttura dei sistemi bancari sono piuttosto rare. Alcune eccezioni sono: BIANCO M., GERALI A., MASSARO R. (1997); BECK T.H.L., DEMIRGÜC-KUNT A., LEVINE R. (2009); AN-TZOULATOS A., PANOPOULOU E., TSOUUMAS C. (2010); DI GIACINTO V., ESPOSITO L. (2011).

che hanno caratterizzato le politiche economiche in ambito creditizio, industriale e regionale o se, al contrario, a queste politiche si siano accompagnate fasi alterne di convergenza, stabilità e divergenza dello sviluppo finanziario regionale in Italia.

Molto in breve, ciò che emerge dalla nostra analisi è che nel complesso vi è stato un processo di convergenza finanziaria tra le regioni italiane, ma che questo processo si è sostanzialmente limitato al primo ventennio del secondo dopoguerra, in linea con quanto avvenuto dal lato dell'economia reale (Daniele e Malanima, 2007). Successivamente, si è assistito a un certo grado di convergenza interna tra le regioni meridionali, ma il divario di sviluppo finanziario tra Mezzogiorno e Centro-Nord certamente non si è ridotto e in alcune fasi si è fortemente ampliato. Ciò è vero in modo particolare per le attività di impiego delle banche, mentre la raccolta, alimentata dalle politiche di redistribuzione dei redditi a favore del Mezzogiorno, ha evidenziato un andamento più costantemente convergente. A partire dall'inizio degli anni Duemila, una volta assestato il lungo e faticoso percorso di consolidamento e ammodernamento del sistema bancario italiano, gli impieghi nelle regioni del Sud rapportati al loro Prodotto interno lordo (Pil) hanno avuto un lieve incremento relativo rispetto a quanto avvenuto nelle regioni del Centro-Nord. Tuttavia, questo riavvicinamento sembra essersi bruscamente arrestato con l'insorgere della crisi finanziaria globale che ha determinato una forte restrizione del credito da parte delle banche italiane che ha investito in modo particolare le imprese localizzate in mercati serviti da istituti di credito funzionalmente distanti dalle regioni di insediamento, prevalentemente concentrati nelle regioni del Sud (Barboni e Rossi, 2012; Gobbi e Sette, 2012; Presbitero *et al.*, 2012).

Il resto del lavoro è organizzato nella maniera seguente. Nel paragrafo 2 discutiamo la periodizzazione delle politiche del credito e delle politiche regionali che adottiamo per analizzare l'evoluzione della convergenza finanziaria regionale. Nel paragrafo 3 presentiamo le nostre misure di sviluppo finanziario e i concetti di convergenza adottati. Nel paragrafo 4 analizziamo la convergenza finanziaria dal 1890 a oggi, mentre nel paragrafo 5 ci concentriamo sui processi di convergenza nei diversi periodi storici individuati nel secondo paragrafo. Alle conclusioni è infine dedicato il paragrafo 6.

2. - Le politiche per il credito e lo sviluppo regionale: una periodizzazione

La prima questione da affrontare riguarda la periodizzazione da adottare. Siccome l'oggetto della nostra indagine è l'articolazione territoriale dello sviluppo fi-

nanziario, alla tradizionale suddivisione temporale relativa agli interventi di riequilibrio regionale e di sviluppo del Mezzogiorno, ci è sembrato opportuno sovrapporre una periodizzazione riguardante le strategie di regolamentazione del sistema bancario da parte delle autorità di vigilanza. Da questo punto vista, possiamo distinguere tre periodi. Il primo periodo, che va dal 1890 fino alla legge bancaria del 1936, è fondamentalmente caratterizzato da un sistema bancario fragile e tendenzialmente instabile, privo di un'azione di indirizzo e di regolamentazione da parte di un'autorità monetaria. Il secondo periodo, che dal 1936 si protrae fino alla fine degli anni '80, è contraddistinto dal modello di regolamentazione strutturale e dai vincoli amministrativi all'attività di impiego delle banche, attraverso il quale la Banca d'Italia durante questo cinquantennio ha governato e guidato la stabilità del sistema bancario, il sostegno allo sviluppo industriale e la trasmissione della politica monetaria, e dalla forte prevalenza della proprietà pubblica delle banche. Successivamente, nel periodo che dagli anni '90 si estende fino ad oggi, si assiste ad una progressiva trasformazione del modello di vigilanza da parte della Banca d'Italia, la quale, anche a seguito di sollecitazioni provenienti dall'Unione Europea, spinge il sistema bancario italiano verso un assetto più competitivo, adottando un approccio di carattere prudenziale nel cui ambito si esplica la concorrenza tra gli intermediari, in gran parte passati ormai dalla sfera pubblica a quella privata, ai quali le modifiche legislative intervenute in quegli anni consentono un'ampia libertà nell'individuazione delle strategie aziendali, nella localizzazione degli sportelli e nella definizione della gamma dei servizi offerti alla clientela.

In relazione all'ambito delle politiche regionali, come documentano Daniele e Malanima (2007), al momento del decollo dell'economia italiana i divari di sviluppo tra le regioni italiane apparivano alquanto limitati e soprattutto non riguardavano le sole regioni meridionali, i cui livelli di reddito *pro-capite* si attestavano in taluni casi su grandezze in linea con la media nazionale (Puglia, Sardegna e Sicilia) o anche superiori ad essa (la Campania). Squilibri nella diffusione territoriale dello sviluppo economico cominciano a emergere con una certa evidenza nei primi decenni del Novecento, ma è solo con la costituzione dell'IRI nel 1933, e in particolare con la presidenza Giordani nel 1939 e la costituzione dell'Isveimer nel 1938, che si assiste a un primo embrionale disegno di riequilibrio territoriale dei flussi finanziari e di azioni concrete in favore dello sviluppo dell'attività produttiva nelle regioni del Mezzogiorno. A partire dal secondo dopoguerra vengono a delinearsi politiche regionali di sviluppo in senso pieno. Il periodo che va dall'istituzione della Cassa per il Mezzogiorno fino ai primi anni '70 è caratterizzato dal prevalere di un'impostazione di politica economica im-

prontata a interventi dal lato dell'offerta (infrastrutture prima e politiche industriali poi), volti a favorire l'ampliamento della base produttiva meridionale. Il ventennio che dal 1974 giunge fino alla soppressione dell'intervento straordinario nel 1992, si caratterizza invece per una progressiva perdita di importanza delle politiche attive di sviluppo che lasciano uno spazio crescente a interventi dal lato della domanda finalizzati al sostegno dei redditi nel Mezzogiorno. Gli anni compresi tra il 1992 e la fine del decennio sono caratterizzati da un significativo disimpegno delle politiche di riequilibrio territoriale, cui segue una fase, che di fatto si prolunga fino al 2007, in cui si riapre una stagione di interventi nell'ambito della cosiddetta "Nuova programmazione", diretti nuovamente a intervenire sui "vincoli" che gravano sul sistema produttivo meridionale. Infine, negli anni della crisi globale si assiste di nuovo a una brusca interruzione delle politiche per il Sud sacrificate in favore delle esigenze della stabilizzazione del debito pubblico e del finanziamento degli ammortizzatori sociali e degli scarni interventi anti-crisi.

Sovrapponendo le diverse fasi appena richiamate, possiamo pertanto suddividere i centoventi anni trascorsi dal decollo industriale ad oggi nei seguenti sei sotto-periodi: (i) 1890-1936; (ii) 1937-1950; (iii) 1951-1973; (iv) 1974-1989; (v) 1990-1996; (vi) 1997-2009.

2.1 *La nascita di un sistema bancario moderno e il finanziamento del decollo industriale (1890-1936)*

All'indomani dell'unificazione dello Stato italiano, il grado di sviluppo finanziario lungo l'intera penisola appare decisamente contenuto se confrontato con la ricchezza reale disponibile, tanto da far parlare all'epoca di deserto bancario (Della Torre *et al.*, 2008). Come ricorda Marco Onado (2003), al momento della dichiarazione del corso forzoso nel 1866 all'incirca il 90% delle transazioni monetarie avveniva ancora con l'impiego di moneta metallica, interi strati della popolazione italiana non avevano mai avuto alcun contatto con le banche e il credito in favore dei contadini in molte zone del Paese, specie nelle regioni del Sud, continuava a essere concesso prevalentemente in natura.

L'architettura del sistema bancario è formata da tre circuiti paralleli e praticamente non comunicanti tra loro:

- (1) gli istituti di emissione, le uniche banche a godere di un certo prestigio e a essere attivamente impegnate nel finanziamento dell'industria;
- (2) le aziende di credito ordinarie, prevalentemente impegnate in attività di carattere speculativo connesse al settore immobiliare;

(3) un limitato nucleo di Casse di Risparmio e le Banche Popolari, dedicate alla gestione del risparmio e al finanziamento delle imprese locali.

A questi si aggiunge, il risparmio postale e la Cassa Depositi e Prestiti. Praticamente assente è il mercato interbancario, così come manca una banca che di fatto svolga le funzioni di banca centrale. Ciò impedisce la circolazione della liquidità fra le singole banche e all'interno dell'economia, e contribuisce a spiegare le ricorrenti crisi di liquidità e i fallimenti che caratterizzano il sistema bancario italiano nel periodo post-unitario. La presenza di sei istituti di emissione rende estremamente difficoltoso il compito di favorire l'integrazione delle diverse realtà economiche locali in un unico sistema nazionale. Sebbene la Banca Nazionale godesse di una posizione di chiara preminenza tra le banche di emissione, sarà solo nel 1926 che verranno superate le diffidenze e le opposizioni al progetto di unificazione degli istituti di emissione.

Nel primo trentennio post-unitario si assiste a un rapido sviluppo del sistema delle banche Popolari che, al Nord come al Sud, vede più che decuplicare la sua presenza numerica (da 50 istituti nel 1870 si passa a 608 nel 1887) e quintuplicare la sua quota di mercato in termini di attivo controllato (Polsi, 1993; A'Hearn, 2000). Questo intenso processo di sviluppo sarà bruscamente interrotto dalla crisi della Banca Romana e del Credito Mobiliare nel 1893-1894 per riprendere poi vigore nel primo decennio del secolo successivo.

Con gli anni Novanta inizia a profilarsi la prima reale fase dello sviluppo dell'economia italiana nel senso di una trasformazione industriale della sua struttura produttiva. Il nesso tra sviluppo industriale e sviluppo finanziario è costituito dalle due grandi banche miste che nascono nel 1894 (Banca Commerciale Italiana) e nel 1895 (Credito Italiano) e che giocheranno, data anche l'assenza di un mercato dei capitali di rischio, un ruolo decisivo nella nascita dell'industria pesante concentrata nelle tre regioni del "triangolo industriale" (Confalonieri, 1976). L'avvio di questo processo è caratterizzato dall'accentuazione del divario tra le due aree del Paese. Sempre facendo riferimento ai dati riportati in Daniele e Malanima (2007), tra il 1897 e il 1907 il prodotto *pro capite* nel Centro-Nord cresce a un tasso medio del 2,1% contro l'1,4% nel Mezzogiorno; tale divaricazione nei livelli di sviluppo si accentua nel periodo 1908-1914 con tassi di crescita medi rispettivamente dell'1,4% e dello 0,4%.

Gli anni della prima guerra mondiale sono caratterizzati da forti oscillazioni degli impieghi e dei depositi; si riduce progressivamente il grado di patrimonializzazione del sistema bancario, con un contemporaneo aumento delle attività liquide, soprattutto titoli di stato a breve termine in relazione alle necessità di

finanziamento della crescente spesa militare. Lo sviluppo delle attività produttive orientate in misura prioritaria verso lo sforzo bellico determina un'ulteriore accentuazione del divario tra le due aree del Paese: se il prodotto *pro capite* meridionale era pari al 78,6% di quello centro-settentrionale nel 1914, esso scende al 73,7% nel 1918.

Terminata la guerra, oltre alle tensioni sociali e politiche che culmineranno nella liquidazione dello Stato liberale e nell'avvento del fascismo, l'industria italiana appare incapace di riconvertire le proprie strutture produttive dall'economia di guerra all'economia di pace (Ciocca, 2007). La scarsa pressione concorrenziale sul mercato dei beni porta i grandi gruppi industriali a utilizzare i profitti di guerra per rafforzare le posizioni di controllo attraverso un intenso processo di acquisizioni, senza ampliare la base produttiva del Paese (Lanzarone, 1946). Il gruppo piramidale e il meccanismo delle «scatole cinesi» diventano gli strumenti fondamentali attraverso i quali garantire il controllo a un nucleo ristretto di imprenditori e la banca mista ne è parte integrante. Negli anni Venti la commistione tra banca ed impresa tende ad accentuarsi e la crisi delle banche miste sfocia nella costituzione delle prime banche pubbliche di credito speciale (Crediop e Icipu). Tentativi di scalata delle banche da parte dei grandi gruppi industriali danno luogo a crescenti immobilizzi che causano, a loro volta, situazioni di illiquidità delle banche miste che tra il 1926 e il 1935 vedono ulteriormente diminuire sia il loro grado di patrimonializzazione sia la loro capacità di raccolta (Onado, 2003). Allo stesso tempo si assiste al rapido sviluppo lungo tutto il territorio nazionale delle banche private e della cooperazione bancaria.

Il restringimento della base produttiva e la progressiva contrazione del commercio internazionale spingono il divario Nord-Sud in direzione di un ulteriore allargamento. Tra il 1929 e il 1936, a fronte di un tasso medio annuo di crescita del prodotto *pro capite* dell'1% nel Centro-Nord, nelle regioni meridionali si registra un tasso di crescita negativo pari allo 0,4%; nel 1936 il prodotto *pro capite* meridionale è pari al 60% di quello centro-settentrionale.

Tra il 1926 ed il 1936 si succedono una serie di interventi legislativi che incidono profondamente su un sistema finanziario molto provato dalle vicende post belliche e dalla successiva crisi finanziaria mondiale. I provvedimenti legislativi del 1926 sanciscono che la Banca d'Italia rimane l'unico istituto di emissione attribuendole anche la funzione di organo di vigilanza sul sistema bancario; nel Mezzogiorno il Banco di Napoli ed il Banco di Sicilia perdono il privilegio dell'emissione e diventano istituti di credito di diritto pubblico. Nel 1933 la nascita dell'IRI sotto la presidenza di Alberto Beneduce conduce al passaggio alla sfera

pubblica di interi settori dell'industria italiana considerati strategici, mentre nel 1936 la Legge Bancaria pone fine all'esperienza della banca mista.

2.2 *Vigilanza strutturale, banca pubblica e doppia intermediazione (1937-1989)*

Con la legge bancaria del 1936 viene stabilita la separazione tra credito a breve e credito a lungo termine è imposto un freno severo alle commistioni tra banca ed impresa. Nel contempo si assiste alla significativa accentuazione della presenza pubblica nel settore creditizio: secondo le stime di Cotula e Garofalo (1976), la quota di mercato misurata dai depositi bancari riconducibile ad aziende di credito pubbliche che nel 1926 era pari al 27,7%, nel 1936 sale al 75,4%. Più in generale, nel sistema bancario italiano si assiste a una fortissima svolta in senso dirigista che finisce per ridurre «enormemente gli spazi dell'imprenditorialità all'interno del sistema bancario» sottoponendo ad autorizzazione amministrativa ogni scelta riguardante la localizzazione territoriale e l'acquisizione di partecipazioni in altre banche (Onado, 2003, p. 413)².

La normativa sulle banche emanata tra il 1926 e il 1936, unita all'azione degli organi di vigilanza e all'operare delle forze dei mercati favoriscono il processo di concentrazione del sistema e un aumento della dimensione media delle aziende di credito. Tra il 1936 e il 1939 il numero di sportelli per banca passa nel Centro-Nord da 4,2 a 4,5 e nel Sud da 2,9 a 3,4.

L'azione congiunta degli istituti di credito di diritto pubblico, degli istituti di credito speciale e delle aziende di credito, riconducibili indirettamente al settore pubblico attraverso l'azione dell'Iri, si configura come il primo disegno di riequilibrio territoriale dei flussi finanziari. A partire dal 1938, con la creazione dell'Isveimer per iniziativa del Banco di Napoli, si può contare sul finanziamento, sia con capitale di debito sia con partecipazioni al capitale di rischio, di iniziative industriali nel Mezzogiorno (Croce, 1999). Nel 1939, con l'assunzione della presidenza dell'IRI da parte di Francesco Giordani, si inizia a predisporre un programma concreto di interventi a favore delle regioni meridionali. Naturalmente

² A tale riguardo va osservato che nella concezione originaria di Beneduce che fu, di fatto, l'estensore della legge bancaria del 1936, i profitti delle banche sono subordinati alla funzione sociale svolta dal credito; quest'ultimo, infatti, costituisce «l'unico strumento veramente efficace col quale sia possibile dirigere, secondo i bisogni della nazione, lo sviluppo della sua economia. Il credito è quindi funzione dello Stato». Si vedano a riguardo le considerazioni che lo stesso Beneduce svolge in un suo appunto del 18 novembre 1935 dal titolo «Osservazioni sulle partecipazioni industriali dell'Iri e sull'organizzazione di un controllo bancario nazionale» riportate da D'ANTONE L. (2011, pp. 209-210).

gli effetti di queste iniziative, tenuto conto di una congiuntura economica caratterizzata dalla progressiva involuzione del sistema produttivo in senso autarchico in vista del coinvolgimento del Paese nel secondo conflitto mondiale, non furono in grado di influire in misura significativa sull'andamento dell'economia reale (Dell'Orefice, 2007). Alla vigilia dell'entrata in guerra dell'Italia il divario nel prodotto *pro capite* tra le regioni del Mezzogiorno e quelle del Centro-Nord si era ulteriormente allargato attestandosi al livello del 56% (Daniele e Malanima, 2007).

Negli anni immediatamente successivi alla fine del conflitto si assiste a un ulteriore rafforzamento del ruolo della Banca d'Italia nel governo del sistema bancario (Onado, 2003). In questo ambito, la politica del credito delle nostre autorità monetarie si sviluppa intorno a due grandi questioni di natura strutturale: l'articolazione territoriale del sistema bancario e la concorrenza tra banche, da un lato, il finanziamento degli investimenti a medio e lungo termine per la ricostruzione e lo sviluppo, dall'altro. La soluzione che viene allora immaginata e coerentemente portata avanti, nel rispetto della separazione tra banca e industria e della struttura frammentaria e decentrata della nostra industria manifatturiera, ha come nucleo centrale il sostegno alle banche a vocazione localistica – casse di risparmio, banche popolari e cooperative – e lo sviluppo di un sistema di istituti di credito speciale per il finanziamento della grande impresa (pubblica e privata) e la gestione degli incentivi pubblici (Albareto e Trapanese, 1999).

Dal 1947, alla Banca d'Italia è attribuito il potere di autorizzare l'apertura di banche e sportelli nel territorio nazionale nel rispetto della stabilità dei mercati locali del credito e delle esigenze di sviluppo delle singole province. La Banca interpreta a lungo in maniera restrittiva tale importante funzione, in modo particolare per quanto riguarda il numero di banche ritenuto già ampiamente sufficiente a soddisfare le esigenze della nostra industria³. In apparenza meno stringente è l'azione di controllo sul numero degli sportelli, che nei venticinque anni compresi tra il 1946 e il 1971 aumenta del 41%. Tuttavia, la struttura regionale del sistema bancario resta sostanzialmente congelata al 1936 (Tav. 1), con solo, in ossequio alle tesi sul decentramento finanziario di Donato Menichella, una consistente redistribuzione delle quote di mercato in favore delle banche locali

³ Tra il 1946 e il 1959 furono autorizzate 161 nuove aperture di banche, di cui 141 Casse rurali, 17 Banche popolari, 1 Cassa di risparmio e 1 Azienda di credito ordinario (ALBARETO G., TRAPANESE M., 1999). Tra il 1966 e il 1985, è persino attivo un esplicito divieto di costituzione di nuove banche.

e un freno all'espansione territoriale delle grandi banche di interesse nazionale e degli istituti di credito di diritto pubblico⁴. L'effetto (voluto) è ridurre la quota di mercato delle grandi banche⁵ per garantire una maggiore vicinanza funzionale e informativa del sistema bancario alle piccole e medie imprese che animavano tanta parte del nostro tessuto produttivo (Conti, 1997); ma l'ulteriore effetto – forse meno voluto, ma, comunque, tollerato dalle nostre autorità monetarie, come inevitabile costo collaterale per l'ottenimento del primo più importante obiettivo – è congelare la concorrenza sui mercati locali del credito assicurando stabili nicchie di mercato alle banche lì operanti. Ciò contribuisce a creare quella che Giuliano Amato con espressione efficace avrebbe poi definito “la foresta pietrificata”, un sistema di aziende di credito che per la parte maggiore è in mano al settore pubblico e alle sue logiche “non mercantili” e per l'altra parte, in mano ai privati, è posta sotto la tutela conservatrice della Banca d'Italia e del suo stretto controllo sulla struttura dei mercati locali del credito, anch'essa quindi isolata dalle regole e dai vincoli della libera competizione.

⁴ Questa scelta strategica è esplicitamente delineata nella Relazione annuale della Banca d'Italia per il 1948; una ricostruzione di tale politica è in CONTI G., FERRI G. (1997) e ALBARETO G., TRAPANESE M. (1999). Una testimonianza significativa riguardo al rigido controllo sull'apertura di sportelli da parte della Banca d'Italia e alla opposizione all'espansione territoriale dei maggiori istituti bancari è la Relazione del Consiglio di Amministrazione della Banca Commerciale Italiana per il 1954 scritta e letta da Raffaele Mattioli, nella quale si afferma: « La nostra rete di filiali in Italia non ha subito cambiamenti. Abbiamo insistito, e persisteremo, nella richiesta di qualche nuova filiale ed agenzia, giacché, per quanto le nostre domande siano modeste e pienamente documentate, nessuna autorizzazione ci è stata ancora concessa » (citato in RODANO G., 1983, p. 316).

⁵ Tra il 1945 e il 1960 la quota di mercato delle tre Banche di interesse nazionale in termini di totale attivo passa dal 28,6% al 21,45% e quella dei quattro Istituti di credito di diritto pubblico dal 22,95% al 21,28%. Significativa è la contrazione dei due grandi Banche meridionali, il Banco di Napoli e il Banco di Sicilia, che passa, rispettivamente, dall'8,43% al 5,15% e dal 4,75% al 3,87% (ALBARETO G., TRAPANESE M., 1999).

TAV. 1

SPORTELLI PER REGIONE (%)

	1936	1939	1946	1951	1956	1961	1966	1971	1974
Valle d'Aosta	0,56	0,60	0,21	0,19	0,18	0,26	0,28	0,29	0,30
Piemonte	8,74	9,04	9,66	10,08	9,86	9,62	9,62	9,35	9,11
Liguria	2,74	3,08	3,30	3,48	3,48	3,63	3,78	3,84	3,75
Lombardia	16,91	17,68	18,48	18,68	18,65	18,16	18,10	17,97	17,95
Trentino Alto Adige	4,04	4,37	3,67	3,66	3,74	3,61	3,54	3,76	3,91
Veneto	7,68	8,03	8,21	8,31	8,12	7,86	7,84	7,68	7,77
Friuli Venezia Giulia	5,89	4,36	4,54	3,24	3,19	2,98	2,94	2,86	2,81
Emilia-Romagna	9,86	9,99	9,62	10,16	10,23	10,31	10,00	10,11	9,95
Toscana	10,25	10,58	10,59	10,52	10,28	9,81	9,52	9,11	8,89
Marche	4,21	4,36	3,60	3,51	3,45	3,33	3,25	3,18	3,11
Umbria	1,80	2,00	2,05	2,07	2,04	1,93	1,88	1,79	1,75
Lazio	6,35	6,34	6,60	6,71	6,82	6,70	6,78	6,83	6,81
Molise	0,54	0,37	0,28	0,23	0,25	0,24	0,26	0,36	0,34
Abruzzi	2,66	2,53	1,90	1,82	1,72	1,81	1,88	1,99	2,06
Campania	4,00	3,89	3,71	3,76	3,78	3,90	4,27	4,20	4,24
Puglia	3,00	2,84	2,92	3,24	3,49	3,89	3,97	4,06	4,11
Basilicata	1,76	1,04	0,65	0,61	0,69	0,89	0,92	0,94	1,00
Calabria	2,64	2,24	2,22	2,20	2,17	2,12	2,07	1,97	1,94
Sicilia	5,59	5,83	6,74	6,44	6,80	7,88	8,05	8,72	9,11
Sardegna	0,77	0,87	1,05	1,09	1,05	1,07	1,03	0,99	1,09
Italia	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Fonte: SVIMEZ (2011a).

Con i tre piani sportelli realizzati negli anni Ottanta (1978, 1982 e 1985), si assiste a un parziale mutamento di rotta da parte della Banca d'Italia che si pone in maniera esplicita l'obiettivo di ampliare la concorrenza nei mercati locali del credito e di accrescere la dimensione media delle banche. Questo, però, senza modificare la prassi dirigista seguita nel passato né la morfologia del nostro sistema bancario, nella «consapevolezza che il concetto di concorrenza non è necessariamente sinonimo di *laissez-faire*» (Conigliani, 1990, p. 29) se non addirittura che nel mercato del credito la «concorrenza incontrollata conduce, oltre che alla instabilità, al monopolio» (Ciocca, 1982).

Il secondo nodo importante che viene affrontato negli anni della ricostruzione riguarda la questione del credito a medio e lungo termine che, se non può essere esercitato direttamente dalle banche di credito ordinario in ossequio alla legge bancaria del 1936, è assegnato a istituzioni o a sezioni specializzate che di fatto, però, sono sempre riconducibili ad aziende di credito ordinario: nascono così Mediobanca, i Mediocrediti regionali e, successivamente, il Mediocredito Centrale, Centrobanca

ed Efibanca, mentre il Banco di Napoli, il Banco di Sicilia e la Banca di Sardegna si dotano di una Sezione di credito industriale (Croce, 1999; Piluso, 1999).

Quello che si profila in quegli anni è, quindi, un sistema creditizio in cui la principale forma di provvista da parte delle aziende di credito è costituita dai depositi delle famiglie che detengono in questa forma la quota preponderante della loro ricchezza; simmetricamente, le banche finanziano attraverso gli impieghi a breve termine il sistema produttivo. Agli Istituti di credito speciale (Ics) è invece affidato il compito di finanziare gli investimenti del settore produttivo a medio e a lungo termine. Il finanziamento degli Ics è assicurato dall'emissione di strumenti obbligazionari che nel corso del tempo saranno collocati in misura prevalente presso le aziende di credito ordinario, dando luogo a quello che sarà individuato come il regime della "doppia intermediazione" e che caratterizzerà il sistema finanziario italiano fino ai primi anni '80 (Balducci *et al.*, 1979).

Nel Mezzogiorno gli Ics si configurano come veri e propri istituti di sviluppo e concorrono in misura decisiva all'ampliamento dei flussi creditizi verso questa parte del Paese⁶. Così, se nel 1957 gli impieghi degli Ics verso le imprese meridionali rappresentavano il 25% di quelli erogati nel Centro-Nord, nel 1974 essi superano il 43%. Nel corso di questi anni gli Ics opereranno in stretta collaborazione con la Cassa per il Mezzogiorno, intrecciando le proprie vicende gestionali con quelle dell'intervento straordinario. Si porrà quindi il problema dell'individuazione di un criterio di gestione dei fondi pubblici tale da conciliare la corretta gestione bancaria delle operazioni di impiego con la funzione di istituti di sviluppo ad essi assegnata, problema che sarà poi, per buona parte, alla base delle difficoltà del sistema bancario meridionale negli anni Ottanta e della sua successiva liquidazione negli anni Novanta (Giannola, 2002; Zazzaro, 2006).

2.2.1 *La ricostruzione postbellica: le politiche attive di sviluppo (1951-1973)*

L'istituzione nel 1950 della Cassa per il Mezzogiorno e l'intervento straordinario dello Stato a favore delle regioni meridionali rappresentano un elemento di svolta nel panorama della politica economica italiana del dopoguerra che insieme alle agevolazioni creditizie e alla costituzione di nuovi Istituti di credito a medio termine sono le componenti di un'unica strategia di politica economica che lega intervento pubblico e sistema bancario.

⁶ La legge n. 298 del 1954 riordinava il settore del credito a medio termine nell'Italia Meridionale attribuendo all'Irfis (Istituto regionale per i finanziamenti industriali in Sicilia) competenza per la Sicilia, al Cis (Credito industriale Sardo) per la Sardegna e all'Isveimer per il resto del Mezzogiorno. Su queste vicende, si rinvia a BAGELLA M. (1986).

Il carattere di maggior rilievo che qualifica le politiche di intervento è che queste fanno sì che il Mezzogiorno, nell'ambito dell'economia nazionale, passi da un ruolo eminentemente passivo (sovrappopolazione relativa, economia sostanzialmente chiusa e blocco politico conservatore) a un ruolo attivo del tutto nuovo sulla scena nazionale (Giannola *et al.*, 1990). Sia che si guardi ai primi interventi in agricoltura e nelle opere infrastrutturali negli anni Cinquanta, sia che si considerino le vicende successive della politica di industrializzazione negli anni Sessanta e nei primi anni Settanta, la linea di politica economica che prevale con progressiva chiarezza a partire dal 1948 fa del successo dell'intervento nel Mezzogiorno un fattore strategico per lo sviluppo del Paese.

Gli effetti sull'economia meridionale tendono a rovesciare la consolidata tendenza all'ampliamento del divario in termini di prodotto *pro capite* che, per la prima volta dalla fine della guerra, segnalano infatti un significativo e prolungato recupero passando dal 47,3% di quello del Centro-Nord nel 1951, al 67,3% nel 1971. La stessa inversione di tendenza la si osserva per gli investimenti fissi che evidenziano un tasso di crescita medio annuo più sostenuto nelle regioni meridionali rispetto al resto del Paese: 11,6% contro il 9,8%, tra il 1951 e il 1962; 6,7% contro il 5,1%, tra il 1963 e 1975⁷. Questa intensa fase di recupero si interromperà intorno alla metà degli anni Settanta, cioè con la conclusione della seconda fase della politica di industrializzazione e con il ritorno a un'impostazione delle politiche di sviluppo fondate sullo schema autopropulsivo dello sviluppo dal basso trainato dalla domanda.

2.2.2 *Il sostegno dei redditi e l'ampliamento dei divari territoriali (1974-1989)*

Significativamente diverso è il quadro che si profila nel corso degli anni Settanta e del decennio successivo. A seguito delle perturbazioni internazionali seguite allo *shock* petrolifero e alla fine del sistema dei cambi fissi, si apre una lunga stagione di instabilità contrassegnata da elevati tassi di inflazione, squilibri nella bilancia dei pagamenti e rallentamento della crescita che imporranno una brusca revisione della scala delle priorità della politica economica nella quale la ristrutturazione della grande impresa nelle regioni nord-occidentali e lo sviluppo delle imprese minori organizzate in distretti nell'Italia nord-orientale e centrale assumono un rilievo preminente (Graziani, 1998).

⁷ Le elaborazioni su serie di contabilità nazionale disaggregate territorialmente sono basate su dati ricostruiti in modo coerente dal 1951 al 2009 e contenuti in SVIMEZ (2011a, pp. 408-446).

In tale ambito le prospettive di un ulteriore sviluppo della base produttiva nel Mezzogiorno attraverso il rilancio di politiche attive di industrializzazione tendono progressivamente a ridursi (Giannola e Lopes, 1992). Se tra il 1970 ed il 1975 i contributi agli investimenti in conto capitale e in conto interessi rappresentano ancora circa il 5% dei trasferimenti alle regioni meridionali, negli anni che vanno dal 1976 al 1988 essi scendono a poco più del 3%. Il coinvolgimento delle imprese a partecipazione statale nel Mezzogiorno si riduce in misura drastica: fino alla metà degli anni Settanta la percentuale di investimenti che questo comparto di imprese effettua nelle regioni meridionali sul totale supera il 45%, nel periodo successivo, fino alla conclusione degli anni Ottanta, tale quota raggiunge a stento il 30%.

Il sostegno pubblico alla base produttiva del Sud perde incisività e le politiche di riequilibrio territoriale tendono a configurarsi prevalentemente come interventi di natura assistenziale a sostegno dei redditi e di una domanda che alimenta le importazioni dalle altre regioni del Paese, dando luogo a una dipendenza patologica delle regioni meridionali dal resto del Paese (Del Monte e Giannola, 1997). Ciò se da un lato assicura il consenso per le classi dirigenti nazionali e locali che gestiscono i flussi di spesa pubblica e l'appoggio dei ceti imprenditoriali settentrionali che traggono un indubbio vantaggio dalla presenza di un ampio mercato di sbocco al Sud, dall'altro non può impedire che il divario tra le due aree del Paese tenda ad ampliarsi di nuovo: il prodotto *pro-capite* meridionale, che nel 1975 aveva raggiunto il 66% di quello del Centro-Nord, subisce una rilevante contrazione relativa toccando nel 1988 il 59,4% e per poi ridursi ulteriormente al 59,3% nel 1992.

2.3 *Vigilanza prudentiale, liberalizzazione e privatizzazione del sistema bancario (1990- 2009)*

A partire dai primi anni '90 si attenuano progressivamente quei caratteri di discrezionalità fortemente dirigista nel rapporto tra banca centrale e banche ordinarie che avevano caratterizzato l'esperienza successiva alla legge bancaria del 1936.

La specificità di questo rapporto, come abbiamo notato in precedenza, non risiedeva tanto nella naturale preminenza gerarchica della banca centrale rispetto alle banche ordinarie, bensì nel fatto che tale preminenza entrava pesantemente in gioco nel determinare obiettivi, strategie e modalità dei comportamenti aziendali, incidendo quindi inevitabilmente sui livelli di efficienza gestionale delle aziende di credito e sulla efficienza allocativa del mercato creditizio nel suo complesso.

L'inversione di tendenza nel modello di vigilanza, da strutturale a prudentiale, rappresenta un tassello importante del generale processo di evoluzione della stra-

tegia della banca centrale sul piano della gestione della politica monetaria, e la coerente esplicitazione delle conseguenze che sul quadro istituzionale tale linea comporta. Il punto di partenza, stimolato dalle complesse vicende che seguono la crisi del sistema monetario internazionale all'inizio degli anni '70, è la valutazione dei costi e dei benefici collegati al mantenimento della sovranità su strumenti tipici della politica monetaria (tasso di cambio, tasso di interesse, controllo dei mercati dei capitali) rispetto all'alternativa – che progressivamente si afferma – di lasciare spazio al mercato, da un lato, e percorrere la strada dell'integrazione istituzionale a livello europeo, dall'altro.

Correlato a questa rivalutazione del mercato è l'avvio delle trasformazioni societarie da banche pubbliche a società per azioni, previsto nella legge Amato - Carli con la scorporazione delle fondazioni dalle banche e il trasferimento della proprietà di queste alle prime, e quello immediatamente successivo di una progressiva privatizzazione del settore bancario (Messori *et al.*, 2003). Nella stessa direzione si inserisce la politica che punta all'intensificazione della pressione competitiva sul mercato del credito. In questo quadro la riforma del sistema di vigilanza da strutturale a prudenziale sancita con l'emanazione del Testo unico in materia bancaria nel 1993 assume un ruolo fondamentale.

L'obiettivo di salvaguardare la stabilità del sistema creditizio e stimolare al contempo una maggior autonomia strategica degli operatori è perseguito con l'introduzione di coefficienti patrimoniali minimi obbligatori, oltre che con il progressivo smantellamento dei vincoli amministrativi sulla gestione del credito e sulle scelte di insediamento territoriale delle aziende bancarie. I coefficienti patrimoniali tendono al controllo della rischiosità dell'attivo della banca e pongono un limite alla capacità di crescita delle dimensioni operative svincolata da un parallelo sviluppo dei mezzi propri aziendali (Giannola e Lopes, 1996).

Lo scenario che si profila non sarà neutrale per il sistema finanziario e creditizio meridionale (Giannola, 2002; Zazzaro, 2006). Il deterioramento della struttura produttiva conseguente alla contemporanea liquidazione dell'intervento straordinario, si traduce in un peggioramento della qualità dell'attivo e in un aumento dei crediti in sofferenza per le banche del Mezzogiorno. In quest'area del Paese l'incidenza dei crediti in sofferenza sul totale degli impieghi tra il 1992 e il 1998 oscilla intorno al 20% degli impieghi, con punte particolarmente significative, superiori al 30% e al 40%, in regioni come la Calabria e la Sicilia, mentre nel Centro-Nord i crediti in sofferenza incidono mediamente nella misura del 6,3% degli impieghi all'economia (Mattesini e Messori, 2004).

Il deterioramento della qualità del credito nelle regioni del Sud, con il conseguente innalzamento del rischio per gli intermediari, si traduce di fatto in oneri maggiori e in una politica dei tassi di interesse attivi che troppo spesso tende a penalizzare proprio le imprese che godono di una migliore *performance*. Sono loro infatti a essere le uniche in grado di consentire il recupero del costo crescente delle sofferenze delle imprese in difficoltà che grava sul sistema bancario; sono le imprese migliori che subiscono una vera e propria discriminazione (statistica) nel mercato del credito dovuta alla pratica di impiegare il tasso medio di sofferenza nella regione come indicatore della rischiosità futura della clientela locale (Scalera e Zazzaro, 2001). Infine, la ricorrente pratica del multi-affidamento contribuisce ulteriormente ad affievolire la capacità della banca di valorizzare il patrimonio informativo dell'impresa affidata accentuando, al contrario, il carattere di opacità del rapporto tra imprese ed aziende di credito (Ferri e Messori, 2000).

La flessione e la ristrutturazione delle imprese locali meridionali determina un indebolimento del sistema creditizio locale, ossia proprio di quel gruppo di banche con una maggiore vocazione a sostenere l'attività e lo sviluppo delle imprese di minori dimensioni che operano sul territorio. In questa direzione agisce anche il modello di vigilanza prudenziale che, attraverso la griglia dei coefficienti minimi patrimoniali, impone vincoli operativi sempre più stringenti al crescere della quota delle sofferenze sull'attivo e, a maggior ragione, delle perdite patrimoniali.

In definitiva, è indubbio che i nuovi criteri adottati dalla Banca Centrale a garanzia della stabilità del sistema creditizio, in applicazione anche della Direttiva 89/647 della CEE, hanno inciso in misura tutt'altro che trascurabile sulla capacità operativa e di espansione delle banche con sede nel Mezzogiorno. Tali indicatori, infatti, basandosi in una prima fase sul coefficiente di rischiosità e dal 1993 sul coefficiente di solvibilità (*Cooke ratio*) e su quello dimensionale (*Gearing ratio*), hanno in concreto impedito a un sistema di aziende di credito gravato dal peso di crediti in sofferenza, di poter perseguire adeguate strategie di crescita aziendale e politiche di diversificazione produttiva. Inoltre, considerato il loro radicamento in un contesto operativo poco dinamico e con uno scarso sviluppo finanziario, le banche meridionali non hanno nemmeno avuto la possibilità di guadagnare efficienza attraverso un processo di ampliamento e riqualificazione della gamma dei servizi offerti.

2.3.1 *La liquidazione dell'intervento straordinario (1990-1996)*

La fine dell'intervento straordinario sancisce anche formalmente la conclusione di una stagione della politica economica nei confronti del Mezzogiorno che, a partire dalla seconda metà degli anni '70 si era caratterizzata in senso assistenziale

con azioni a sostegno dei redditi e sempre meno finalizzate all'obiettivo dello sviluppo della base produttiva.

Nel 1992, è definitivamente abolito l'intervento straordinario e contemporaneamente hanno inizio le azioni volte al risanamento della finanza pubblica. Infatti, i vincoli derivanti dall'adesione al Trattato di Maastricht richiedono ai paesi firmatari il rispetto di una serie di parametri che impongono, tra l'altro, l'adozione di politiche di contenimento dei disavanzi pubblici.

Dal punto di vista dell'evoluzione della struttura dell'industria si assiste al progressivo disimpegno della grande impresa pubblica e al sostanziale ridimensionamento dei progetti di investimento effettuati dai grandi gruppi privati: questi aspetti sono evidenziati dalla significativa flessione dell'occupazione nella grande impresa che vede diminuire i suoi addetti tra il 1981 ed il 1991 di oltre il 25%. Si apre quindi una fase in cui la dinamica della struttura produttiva meridionale grava in misura quasi esclusiva sulle *performance* dell'impresa locale che proprio in questi anni si vede esposta in misura crescente alla competizione delle imprese operanti nel resto del Paese; questo progressivo e drastico orientamento al mercato eserciterà un effetto significativo in termini di contenimento della base produttiva meridionale.

Il periodo compreso tra il 1991 e il 1996 è inoltre caratterizzato dalla tendenza del sistema produttivo meridionale a operare su una scala più contenuta che si manifesta in misura più intensa rispetto a quella riscontrabile nel resto del Paese. In tale lasso temporale, infatti, la dimensione caratteristica degli impianti industriali meridionali diminuisce da 35 a 28 addetti, mentre per le imprese nel Centro-Nord la riduzione è meno evidente con la dimensione media che passa da 40 a 37 addetti (Giannola, 1998).

Per ciò che concerne infine gli investimenti in infrastrutture questi tendono a ridursi già nel corso degli anni Ottanta attestandosi intorno al 15% delle importazioni nette: sulla base delle rilevazioni della Svimez si può ritenere che gli investimenti in opere pubbliche tra il 1990 ed il 2000 si riducono in termini reali di oltre il 12% (Giannola e Lopes, 2011).

Tutto ciò non può che condurre a un deterioramento del quadro macroeconomico caratterizzato da un tasso di crescita annuo dell'economia meridionale pari appena allo 0,9% contro l'1,5% di quello conseguito nel resto del Paese, che fa precipitare il Pil *pro capite* del Mezzogiorno rispetto al Centro-Nord al 56,1%, da una crescita dei consumi che si attesta a un modesto 0,9% annuo contro l'1,7% delle regioni centro-settentrionali e da una drastica caduta del processo di accumulazione che si traduce in una contrazione media annua degli investimenti fissi

lordi dell'ordine del 2,4% contro una crescita dell'1% degli investimenti effettuati dalle imprese nel Centro-Nord.

2.3.2 La "Nuova Programmazione" (1997-2009)

Alla conclusione repentina dell'intervento straordinario segue una profonda crisi dell'economia meridionale e una pressoché totale assenza di politiche di riequilibrio territoriale che riprendono corpo solo dopo il 1996 con l'avvio della cosiddetta «Nuova Programmazione» (NP).

La NP si colloca – e non poteva essere diversamente – nel solco delle politiche dell'offerta recuperando i principi ispiratori del primo intervento straordinario. In sostanza, essa si concretizza in forme di intervento attivo e discrezionale abbastanza tradizionali che vanno dalla fornitura di infrastrutture e programmazione urbanistica alle spese in formazione e ricerca, dalla valorizzazione delle risorse naturali e culturali all'erogazione di incentivi agli investimenti.

Tuttavia, ed è questa la novità, non più solo con azioni unilaterali da parte di amministrazioni (straordinarie) centrali, bensì coinvolgendo le amministrazioni (ordinarie), le istituzioni e i soggetti locali attraverso un articolato processo decisionale e allocativo che, partendo dalla definizione dei fabbisogni dal basso, prevede la contrattualizzazione delle relazioni amministrative, l'istituzione di forme di partenariato pubblico-privato e l'impiego di accorti ed efficienti meccanismi di selezione dei progetti.

Le ragioni dell'insuccesso sono state discusse in numerose occasioni e con accenti anche molto diversi (Rossi, 2005; Barca, 2006; Franzini e Giunta, 2008; Giannola, 2009; Viesti, 2009; Scalerà e Zazzaro, 2010). Qui è sufficiente ricordare che l'esito complessivamente insoddisfacente di questa nuova strategia di intervento in favore del Mezzogiorno non era imprevedibile⁸, stante l'azzardato tentativo di indirizzare le aspettative e i comportamenti degli agenti verso un sistema di valori basato su fiducia, *civicness* e senso di responsabilità facendo affidamento su strumenti di azione che, per poter avere successo, presuppongono proprio l'esistenza di aspettative e comportamenti che in quel sistema di valori già si muovono da tempo. Inoltre, con troppo ottimismo non si è tenuto conto dell'eccessiva burocratizzazione dell'intervento e delle tante occasioni per attività di *rent seeking* che questo avrebbe prodotto, nonché dei rischi di collusione tra i soggetti coinvolti nella negoziazione rivolta alla spartizione delle risorse pubbliche e dei pericoli di una loro ulteriore deresponsabilizzazione di fronte agli elettori.

⁸ Cfr. DEL MONTE A. (1998) e ZAZZARO A. (2001).

Tra il 1998 e il 2003 si apre una breve stagione di convergenza, dovuta più al minore rallentamento dell'economia meridionale rispetto a quella del Centro-Nord che non a un significativo slancio di crescita delle regioni del Mezzogiorno. Inoltre, la relativa migliore tenuta dell'economia meridionale appare attribuibile non tanto al dinamismo del limitato e fragile sistema locale d'impresе, quanto all'apporto dei fondi strutturali europei che, per il solo fatto di essere spesi anche se in modo spesso "discutibile", attivano comunque flussi consistenti di domanda aggregata.

D'altro canto, in termini di risorse le esigenze di finanziamento di una politica di contesto risultano clamorosamente mortificate se si considera quel che mostrano i consuntivi di spesa ripartiti territorialmente. A tale proposito, va ricordato⁹ che in fase di programmazione, era stato fissato l'obiettivo di destinare al Mezzogiorno il 45% della spesa in conto capitale (il 30% a valere su risorse ordinarie) e l'85% delle risorse per le aree sottoutilizzate, e di riequilibrare la spesa per lo sviluppo a favore degli investimenti pubblici. Ben diversamente, i dati elaborati dal Dipartimento delle Politiche per lo Sviluppo evidenziano che tra il 1998 e il 2006, l'obiettivo minimo di spesa in conto capitale a favore del Mezzogiorno non è mai stato raggiunto, così come l'obiettivo dell'85% della quota di spesa per le aree sottoutilizzate da destinare al Mezzogiorno (ad eccezione del 2002). In particolare, la quota di spesa in conto capitale, passa dal 41,3% del 1996 al 37% del 2006 con un graduale inesorabile declino negli anni successivi.

Quanto alla distribuzione della spesa pubblica in conto capitale tra investimenti e trasferimenti, i dati segnalano la tendenza al riequilibrio tra i due capitoli di spesa. Nel Mezzogiorno nel 1996 gli investimenti erano pari a 8,8 miliardi di euro (a prezzi costanti del 2005), a fronte dei 12,6 destinati agli incentivi, mentre nel 2003 entrambi i capitoli di spesa si attestavano intorno agli 11 miliardi di euro. Lo squilibrio iniziale, dunque, si è tramutato in un sostanziale equilibrio a parità di spesa totale: l'obiettivo del riequilibrio territoriale è stato perseguito in un quadro di riduzione relativa di risorse da destinare al Mezzogiorno.

Non stupisce quindi che negli anni successivi al 2004 anche la debole tendenza alla convergenza evidenziata nel quinquennio precedente si interrompa bruscamente; il tasso di crescita medio annuo del prodotto nel Mezzogiorno tra il 2004 e il 2007 è pari alla metà (0,8%) di quello registrato nel resto del Paese (1,6%). La stessa dinamica dei consumi penalizza le regioni meridionali rispetto al Centro-Nord: essi crescono al Sud nel periodo considerato ad un tasso annuo dello 0,5% contro l'1,3% del resto del Paese.

⁹ Si veda ancora la ricostruzione fornita in GIANNOLA A. e PETRAGLIA C. (2007).

È quindi inevitabile un nuovo ampliamento del divario tra le due circoscrizioni: se nel 2004 il reddito *pro capite* è il 59,9% di quello delle regioni centro-settentrionali, nel 2007 esso scende al 57,4%, e ciò nonostante il fatto che il Mezzogiorno abbia avuto crescita demografica nulla nell'ultimo decennio e la popolazione del Centro-Nord sia cresciuta intorno al 6% soprattutto per l'effetto dell'immigrazione, ossia per l'afflusso di individui con redditi molto inferiori alla media dell'area. Le aspettative riposte in questa nuova fase della politica di intervento non si sono quindi realizzate e diventa ancora più pressante per il Sud il tema del «deficit di crescita rispetto al Centro-Nord»¹⁰.

3. - Sviluppo finanziario e convergenza

In termini ampi, la nozione di sviluppo finanziario fa riferimento alla capacità del sistema finanziario di mobilitare il risparmio e far affluire le risorse monetarie e finanziarie nelle mani dei soggetti che le sanno impiegare in maniera più produttiva. Si tratta, com'è evidente, di una nozione irrimediabilmente vaga e incerta, essendo molteplici i mercati, le istituzioni e gli strumenti che compongono il sistema finanziario e, soprattutto, essendo molteplici le funzioni che il sistema finanziario è chiamato a svolgere per favorire la corretta allocazione delle risorse monetarie e reali (Levine, 2005).

A misura di questa indeterminatezza, nella letteratura economica sono stati proposti numerosi indicatori di sviluppo finanziario che considerano solo alcuni caratteri del sistema finanziario. In questo lavoro, avendo a riferimento il sistema bancario e le regioni, prenderemo in esame cinque indicatori già ampiamente utilizzati in letteratura.

I primi due indicatori riguardano l'articolazione territoriale del sistema bancario regionale e la sua struttura organizzativa. Il primo è il tradizionale tasso di bancarizzazione, sportelli/popolazione, che misura la diffusione del sistema bancario nel territorio¹¹. Il secondo, invece, prende in considerazione il grado di vicinanza funzionale delle banche al territorio¹², la loro autonomia decisionale,

¹⁰ Si vedano a tale riguardo le considerazioni contenute in SVIMEZ (2006).

¹¹ A tale riguardo si utilizza talvolta anche il rapporto Sportelli/Pil sulla base del presupposto che il risultato dell'attività del sistema creditizio che si esplica attraverso l'istallazione di sportelli sul territorio è rappresentato dal livello di produzione a sua volta approssimato dal Pil.

¹² Sulla nozione di distanza funzionale e distanza operativa, cfr. ALESSANDRINI P., CROCI M., ZAZZARO A. (2005) e ALESSANDRINI P., PRESBITERO A., ZAZZARO A. (2009).

misurata dalla quota di sportelli nella regione appartenenti a banche aventi il proprio centro direttivo (la sede legale) nella regione stessa.

Gli altri tre indicatori che analizziamo sono:

- (a) il rapporto impieghi/Pil – l'indicatore di sviluppo finanziario più spesso impiegato nella moderna letteratura su finanza e crescita –;
- (b) il rapporto (impieghi+depositi)/sportelli – una misura semplice anche se un po' rozza della produttività del sistema bancario regionale –;
- (c) il rapporto impieghi/depositi – una misura della capacità di mobilitazione del risparmio locale all'interno del circuito produttivo regionale.

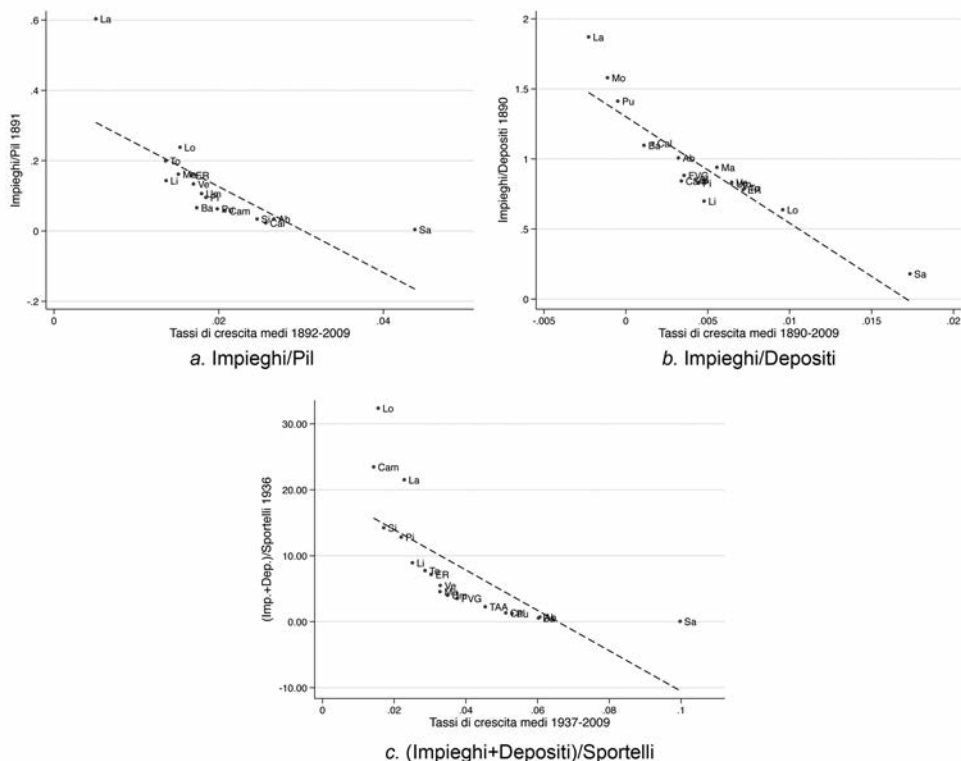
Infine, per valutare la convergenza dello sviluppo finanziario nelle regioni italiane faremo riferimento ai due classici concetti di *convergenza beta* e *convergenza sigma* impiegati nella letteratura sulla crescita economica. In particolare, diremo che c'è *convergenza beta* quando le regioni che inizialmente hanno sistemi finanziari meno sviluppati aumentano il loro sviluppo finanziario a tassi di crescita più sostenuti rispetto alle regioni finanziariamente più sviluppate. Diremo invece che c'è *convergenza sigma* quando la dispersione nei livelli di sviluppo finanziario tra le regioni si riduce nel tempo. Com'è noto, i due concetti non sono coincidenti, la *convergenza beta* essendo solo necessaria alla *convergenza sigma*: ad esempio, le regioni potrebbero convergere nei livelli di sviluppo finanziario, ma se la variabilità iniziale dello sviluppo finanziario tra le regioni fosse inferiore alla deviazione standard degli *shock* successivi, la varianza potrebbe evidenziare un *trend* temporale crescente (Barro e Sala-i-Martin, 1992).

4. - Uno sguardo d'insieme

Se si guarda all'articolazione territoriale dell'attività di impiego e di raccolta delle banche italiane nel corso dell'intero periodo che dal 1890 arriva ai nostri giorni, ciò che sembra emergere è che in Italia lo sviluppo finanziario ha seguito un effettivo processo di convergenza tra regioni, con quelle meridionali che sono divenute finanziariamente più simili alle regioni del Centro-Nord.

Come mostra il Grafico 1, le regioni che ad inizio del periodo si attestavano su livelli più bassi della media per ciascuno degli indicatori di sviluppo finanziario considerati, hanno sperimentato tra il 1890 e il 2009 dei tassi di crescita medi degli stessi indicatori superiori a quelli delle regioni che inizialmente risultavano finanziariamente più sviluppate. Ciò è particolarmente evidente per il rapporto impieghi/Pil (cfr. Graf. 1a), che nelle regioni del Sud, non superando mai il 7%

CONVERGENZA BETA 1890-2009



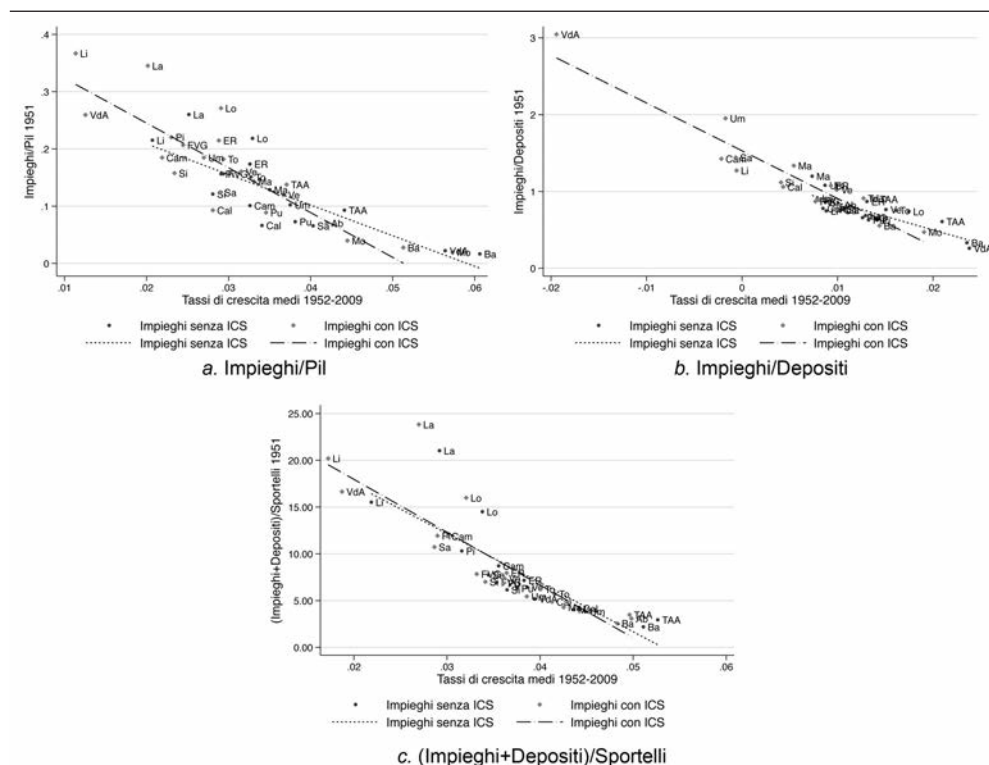
Nota. A. Impieghi/Pil. Gli anni esaminati sono: 1891, 1895, 1900, 1905, 1910, 1929, 1936, 1938, 1939-2009. Fino al 1951 sono escluse: Valle d'Aosta, Trentino Alto Adige, Friuli Venezia Giulia e Molise. Fino al 1951 la serie del Pil regionale è stata ricostruita a partire dai dati sul Pil *pro-capite* regionale a prezzi 1911 riportati in DANIELE V. e MALANIMA P. (2007) e convertiti in euro e dai dati sulla popolazione media annuale per regioni della SVIMEZ (2011a); nel caso di dati mancanti è stata utilizzata la popolazione osservata nell'anno immediatamente precedente, $t-x$, aumentata per il tasso di crescita medio annuale tra questo e l'anno immediatamente successivo a quello mancante, $g_{t-x,t+y}$, ossia: $Pop_t = Pop_{t-x} (1 + g_{t-x,t+y})^x$. La serie degli impieghi è quella pubblicata dalla SVIMEZ (2011a) calcolata in valori del 1911 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009). Per il periodo 1951-2009 le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). **B. Impieghi/Depositi.** Gli anni esaminati sono: 1890, 1895, 1900, 1905, 1910, 1929, 1936, 1938, 1939-2009. La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. Fino al 1920 sono escluse la Valle d'Aosta e il Trentino Alto Adige e per il 1936 è esclusa la Valle d'Aosta. **C. (Impieghi+Depositi)/Sportelli.** Gli anni considerati sono: 1936, 1939, 1946, 1951, 1956, 1961, 1966, 1971, 1975-2009. Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a); fino al 1966 è escluso il Molise, mentre per il 1936 è esclusa anche la Valle d'Aosta.

nel 1890, è cresciuto nel periodo a tassi annuali decisamente superiori al 2% (in particolare la Sardegna è cresciuta a tassi superiori al 4%), e che nel Lazio è cresciuto solo a tassi del 5‰ facendo perdere alla regione nel corso del tempo la sua centralità finanziaria. Andamenti non dissimili si ritrovano nella capacità di intermediazione degli sportelli regionali e nella capacità di mobilitazione del risparmio locale (cfr. Graff. 1b e 1c), anche se è interessante notare che nel 1890 era nelle regioni meridionali che il rapporto impieghi/depositi assumeva i valori maggiori a evidenziare le scarsissime capacità di risparmio dell'area.

La stessa tendenza si ritrova nel periodo che va dal 1951 ad oggi, con Liguria e Lazio che perdono terreno a favore delle altre regioni del Centro-Nord e di quelle del Mezzogiorno (cfr. Graff. 2a - 2b - 2c).

GRAF. 2

CONVERGENZA BETA 1951-2009



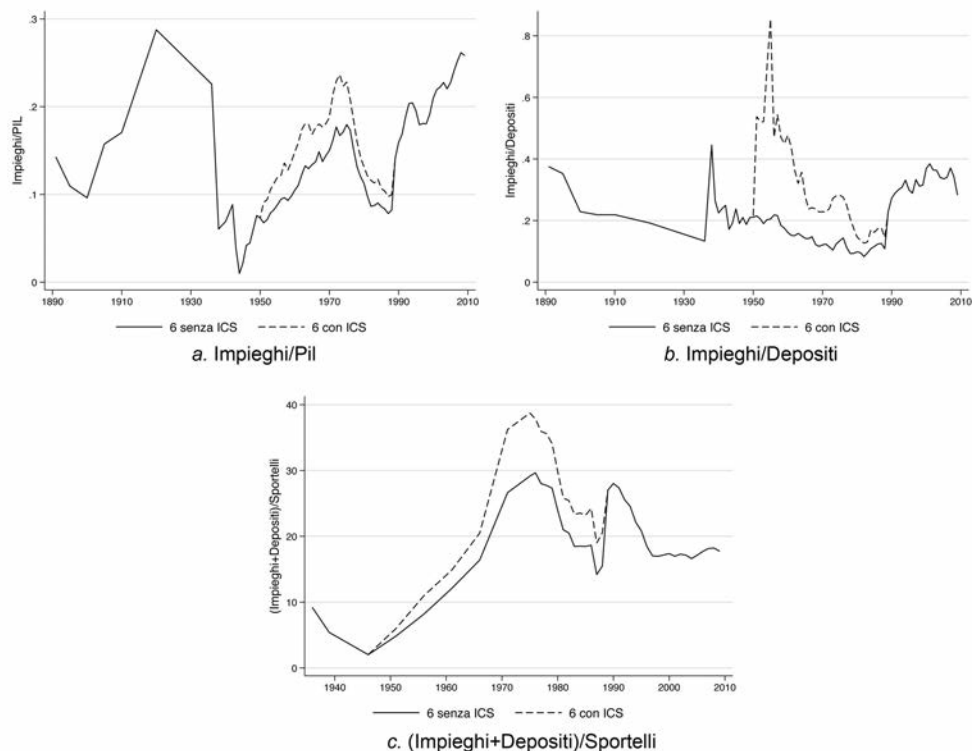
Nota: **A. Impieghi/Pil.** Le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). Dal 1951 al 1988 gli impieghi sono calcolati sia al lordo sia al netto degli impieghi degli Istituti di Credito Speciale. **B. Impieghi/Depositi.** La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. **C. (Impieghi+Depositi)/Sportelli.** Gli anni considerati sono: 1951, 1956, 1961, 1966, 1971, 1975-2009. Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a); fino al 1966 è escluso il Molise.

Tuttavia, questo avvicinamento relativo medio non sempre si è accompagnato a una minore dispersione dello sviluppo finanziario tra le regioni nel corso del tempo (cfr. Graff. 3a - 3b - 3c). In termini assoluti, la dispersione cross-regionale dello sviluppo finanziario è a inizio periodo inferiore o leggermente superiore (impieghi/depositi, Graf. 3b) a quella che si osserva nel 2009. Una contenuta diminuzione della deviazione standard è osservabile per quanto riguarda le risorse totali intermedie dagli sportelli locali (Graf. 3c) nel periodo 1950-1980, e per il rapporto impieghi/depositi (Graf. 3b) nel periodo 1971-1990. L'ammontare degli impieghi bancari regionali rapportati al Pil regionale, invece, assume una variabilità crescente lungo tutto il secondo dopoguerra, calcolando gli impieghi sia al lordo che al netto dell'attività degli Istituti di Credito Speciale, e raggiunge il suo massimo proprio nel 2009 (Graf. 3a).

Ma soprattutto, le differenze assolute nello sviluppo finanziario tra Mezzogiorno e Centro-Nord si sono fortemente allargate. Ciò sta evidentemente a indicare che il processo di convergenza medio nasconde andamenti diversi nel corso del tempo. In particolare, a partire dagli anni Novanta, con l'abbandono della regolamentazione strutturale da parte della Banca d'Italia, la liberalizzazione dell'apertura di nuovi sportelli e l'avvio del processo di privatizzazione e consolidamento della nostra industria bancaria, si è assistito a una divaricazione tra Nord e Sud del Paese tanto nelle strutture dei sistemi bancari locali, quanto nell'attività di impiego delle banche nelle due aree.

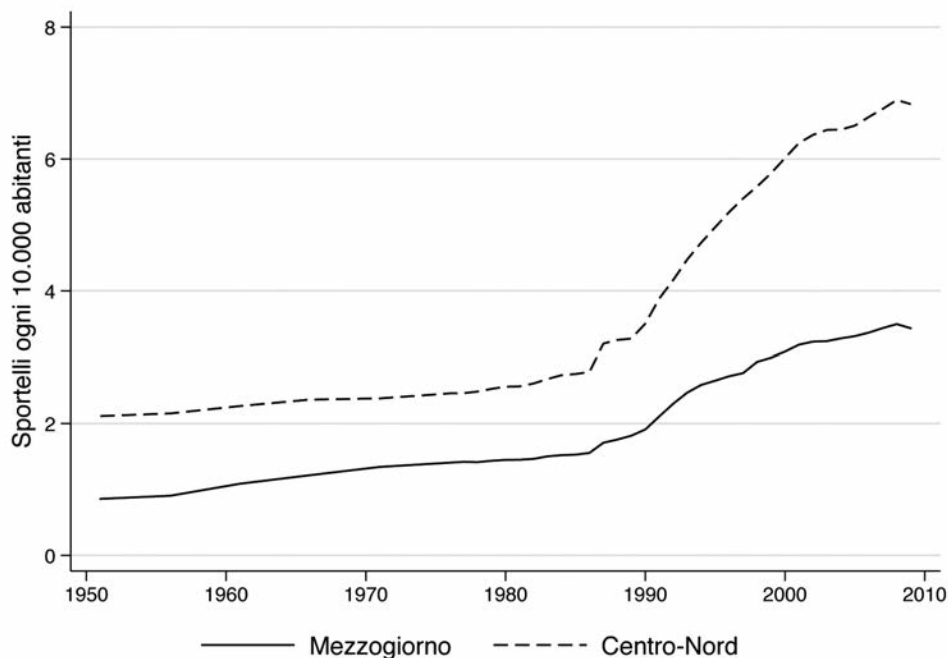
Come mostra il Grafico 4, nel 1971 il Mezzogiorno poteva contare su poco più di uno sportello bancario ogni diecimila abitanti, mentre nel Centro-Nord la presenza bancaria era quasi doppia, superando i due sportelli per diecimila abitanti. Questa differenza si è mantenuta pressoché costante in termini assoluti fino al 1986 (l'anno d'avvio dell'ultimo piano sportelli), per poi divenire sempre più ampia nel corso dei vent'anni successivi.

CONVERGENZA SIGMA 1890-2009



Nota: A. Impieghi/Pil. Gli anni esaminati sono: 1891, 1895, 1900, 1905, 1910, 1929, 1936, 1938, 1939-2009. Fino al 1951 sono escluse: Valle d'Aosta, Trentino Alto Adige, Friuli Venezia Giulia e Molise. Fino al 1951 la serie del Pil regionale è stata ricostruita a partire dai dati sul Pil *pro-capite* regionale a prezzi 1911 riportati in DANIELE V. e MALANIMA P. (2007) e convertiti in euro e dai dati sulla popolazione media annuale per regioni della SVIMEZ (2011a); nel caso di dati mancanti è stata utilizzata la popolazione osservata nell'anno immediatamente precedente, $t-x$, aumentata per il tasso di crescita medio annuale tra questo e l'anno immediatamente successivo a quello mancante, $g_{t-x:t+y}$, ossia: $Pop_t = Pop_{t-x} (1 + g_{t-x:t+y})^x$. La serie degli impieghi è quella pubblicata dalla SVIMEZ (2011a) calcolata in valori del 1911 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009). Per il periodo 1951-2009 le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). Dal 1951 al 1988 gli impieghi sono calcolati sia al lordo sia al netto degli impieghi degli Istituti di Credito Speciale. **B. Impieghi/Depositi.** Gli anni esaminati sono: 1890, 1895, 1900, 1905, 1910, 1929, 1936, 1938, 1939-2009. La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. Fino al 1920 sono escluse la Valle d'Aosta e il Trentino Alto Adige e per il 1936 è esclusa la Valle d'Aosta. **C. (Impieghi+Depositi)/Sportelli.** Gli anni considerati sono: 1936, 1939, 1946, 1951, 1956, 1961, 1966, 1971, 1975-2009. Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a); fino al 1966 è escluso il Molise, mentre per il 1936 è esclusa anche la Valle d'Aosta.

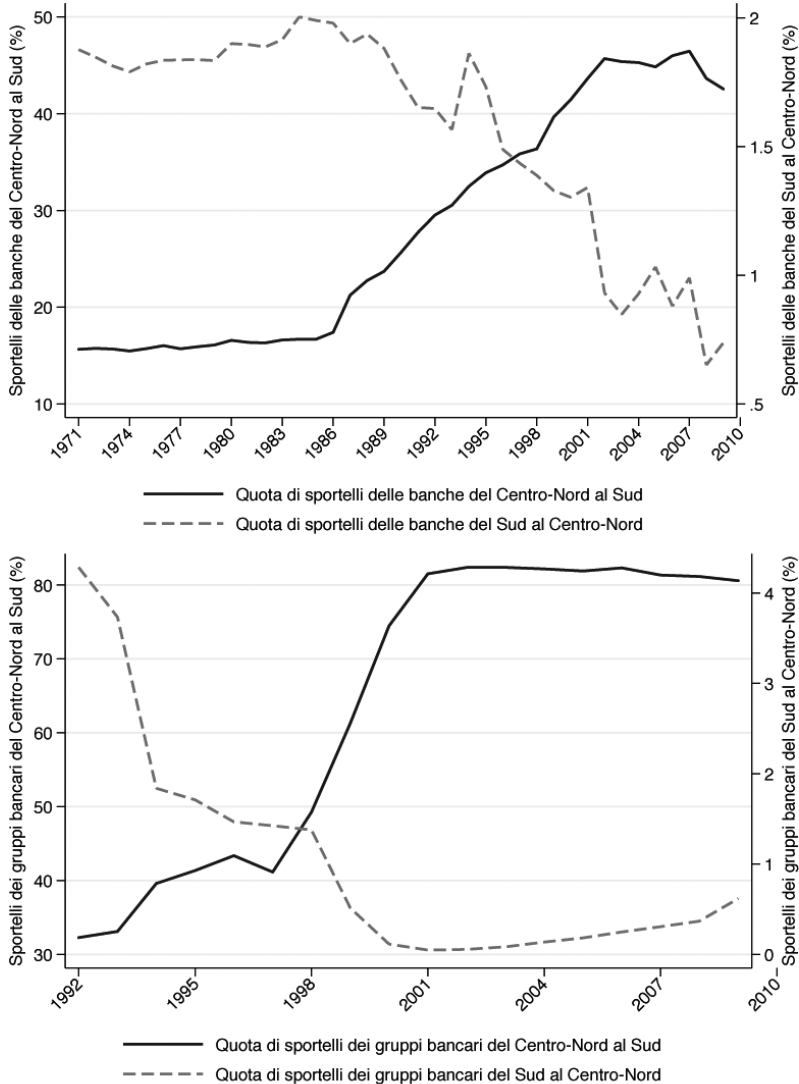
GRADO DI BANCARIZZAZIONE



Nota: Gli anni esaminati sono: 1951, 1956, 1961, 1966, 1971, 1975-2009. Le serie degli sportelli e della popolazione sono contenute in SVIMEZ (2011a). Nel caso di dati mancanti è stata utilizzata la popolazione osservata nell'anno immediatamente precedente, $t-x$, aumentata per il tasso di crescita medio annuale tra questo e l'anno immediatamente successivo a quello mancante, $g_{t-x:t+y}$, ossia: $Pop_t = Pop_{t-x}(1 + g_{t-x:t+y})^x$.

Ciò che evidenzia ancor di più il processo di crescente integrazione finanziaria passiva delle regioni meridionali è però la penetrazione delle banche esterne nelle due circoscrizioni. Nel Mezzogiorno, nel 1971, la quota di sportelli di banche aventi la loro sede legale nel resto del Paese era all'incirca pari al 16%. Alla stessa data, le banche del Mezzogiorno detenevano un modesto 1,7% degli sportelli attivi nel Centro-Nord. Ancora una volta, queste quote sono rimaste sostanzialmente invariate fino al 1986, testimoniando la coesistenza in Italia di due sistemi bancari territorialmente separati e funzionalmente autonomi. Tale struttura bipartita si è dissolta negli anni Novanta con la messa in liquidazione e l'acquisizione del Banco di Napoli, del Banco di Sicilia e delle principali Casse di Risparmio

PRESENZA DI BANCHE ESTERNE



Nota: i dati sugli sportelli per banca sono di fonte BANCA D'ITALIA.

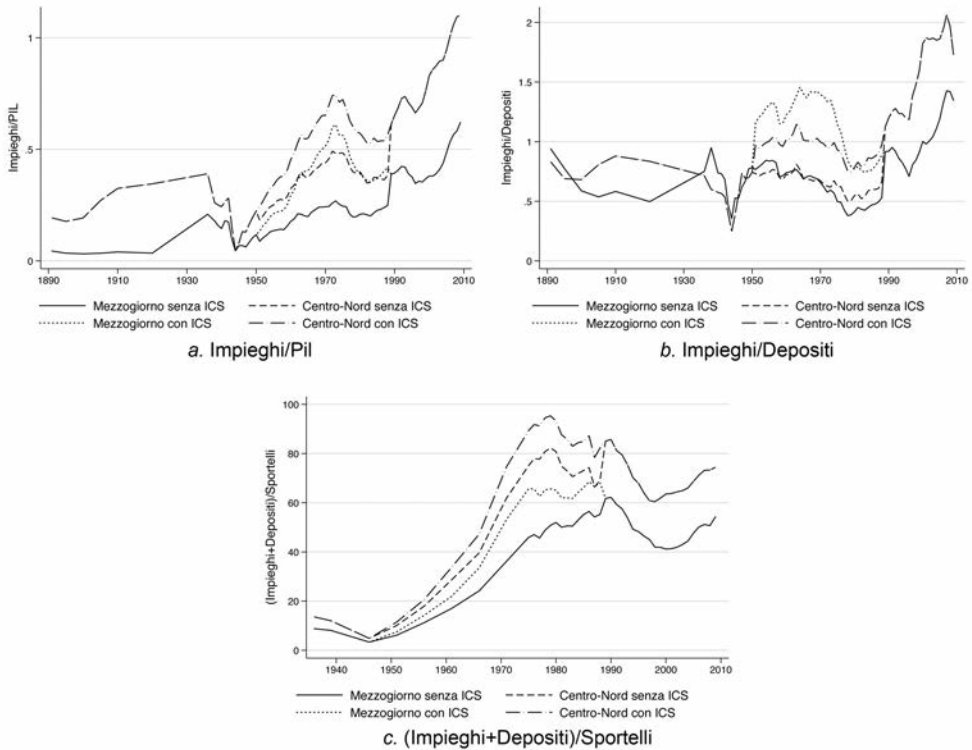
del Mezzogiorno da parte di banche del Centro-Nord¹³. Nel 2010, più del 42% degli sportelli operanti nel Mezzogiorno risulta di proprietà di banche esterne all'area e un altro 38% è riferibile a banche che hanno mantenuto la sede legale nel Mezzogiorno ma che sono entrate a far parte di gruppi bancari guidati da banche del Centro-Nord, mentre gli sportelli meridionali di banche con la sede nella regione sono ben inferiori all'1% (cfr. Graf. 5).

Una simile dinamica divergente è osservabile anche nell'attività di erogazione del credito da parte delle banche nelle due aree, mentre la loro complessiva capacità di intermediazione ha avuto andamenti sostanzialmente paralleli. Il rapporto impieghi su Pil, che, per i primi trent'anni dal dopoguerra, ha seguito un andamento medio più rapido nel Mezzogiorno che nel Centro-Nord – con un differenziale¹⁴ che nel 1973 arrivava a toccare il punto più basso, inferiore al -18% –, nel periodo successivo è cresciuto nelle regioni del Sud in maniera sensibilmente meno marcata fino a riportare nel 2009 la differenza relativa tra le due aree allo stesso livello del 1951 (vedi Graf. 6a). Allo stesso modo, il rapporto impieghi su depositi, che nel 1973 era nel Mezzogiorno del 41% più alto che nel Centro-Nord, negli ultimi trent'anni è sempre stato nelle regioni meridionali mediamente più basso: la differenza si è andata allargando nel periodo 1989-2002 quando ha toccato il punto più sfavorevole al Mezzogiorno (0,55% del Centro-Nord) per poi restringersi negli anni della recente crisi finanziaria (Graf. 6b). Infine, l'ammontare di fondi intermediati (impieghi più depositi) per sportello è rimasta nel Mezzogiorno nettamente inferiore rispetto al Centro-Nord, con un differenziale che nel 2009 è di fatto agli stessi livelli del 1971 (Graf. 6c).

¹³ Per una ricostruzione e un'interpretazione delle vicende che hanno travolto il sistema bancario meridionale negli anni Novanta, si rimanda a GIANNOLA A. (2002); MATTESINI F. e MESSORI M. (2004); ZAZZARO A. (2006).

¹⁴ Tale differenziali sono calcolati a partire dai dati riportati nei Grafici 6a, 6b e 6c considerando le differenze relative $((MZ-CN)/CN) \times 100$.

DIFFERENZIALI DI SVILUPPO FINANZIARIO NORD-SUD



Nota: A. Impieghi/Pil. Gli anni esaminati sono: 1891, 1895, 1900, 1905, 1910, 1929, 1936, 1938, 1939-2009. Fino al 1951 la serie del Pil regionale è stata ricostruita a partire dai dati sul Pil *pro-capite* regionale a prezzi 1911 riportati in DANIELE V. e MALANIMA P. (2007) e convertiti in euro e dai dati sulla popolazione media annuale per regioni della SVIMEZ (2011a); nel caso di dati mancanti sulla popolazione è stata utilizzata la popolazione osservata nell'anno immediatamente precedente, $t-x$, a quello mancante, t , aumentata per il tasso di crescita medio annuale tra l'anno immediatamente precedente e l'anno immediatamente successivo a quello mancante, $g_{t-x:t+y}$, ossia: $Pop_t = Pop_{t-x} (1 + g_{t-x:t+y})^x$. La serie degli impieghi è quella pubblicata dalla SVIMEZ (2011a) calcolata in valori del 1911 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009). Per il periodo 1951-2009 le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). Dal 1951 al 1988 gli impieghi sono calcolati sia al lordo sia al netto degli impieghi degli Istituti di Credito Speciale. **B. Impieghi/Depositi.** Gli anni esaminati sono: 1890, 1895, 1900, 1905, 1910, 1929, 1936, 1938, 1939-2009. La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. **C. (Impieghi+Depositi)/Sportelli.** Gli anni considerati sono: 1936, 1939, 1946, 1951, 1956, 1961, 1966, 1971, 1975-2009. Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a).

5. - Le alterne vicende dello sviluppo finanziario regionale

In questo paragrafo analizzeremo più in dettaglio il processo di sviluppo finanziario delle regioni italiane studiando separatamente i sei sottoperiodi individuati in precedenza per valutare in che misura la convergenza ha seguito dinamiche differenziate nel corso del tempo, parallelamente ai mutamenti delle politiche del credito e di intervento regionale richiamate nel paragrafo 2.

5.1 1890-1936

Alla fine dell'Ottocento l'Italia appare come un Paese nell'insieme finanziariamente arretrato con un rapporto impieghi/Pil che in media si attesta a un modesto 13,7%, con solo il Lazio che raggiunge quota 60%; nelle regioni del Sud tale rapporto supera appena il 4% (in Calabria gli impieghi sono il 2,3% del Pil regionale mentre in Sardegna rappresentano un misero 4,3‰). Vista l'arretratezza finanziaria del Paese, gli elevati valori del rapporto impieghi/depositi, che nelle regioni del Sud è in media pari a 0,93, vanno allora letti come un indicatore dell'incapacità dell'economia di produrre risparmio e delle banche di farlo affluire nel circuito finanziario in favore delle attività produttive, più che della capacità delle banche di mobilitarlo localmente.

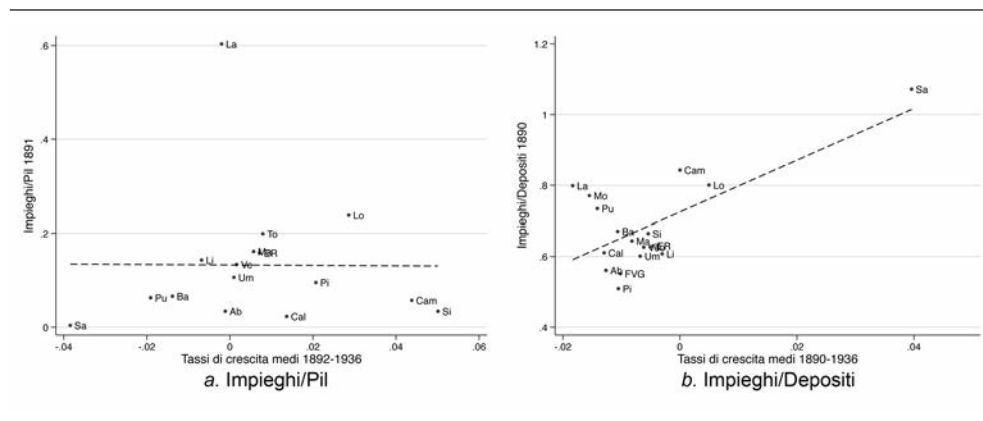
Quarant'anni dopo, nel 1936, gli impieghi bancari in Italia sono quasi triplicati rispetto al Pil attestandosi al 34,5% e la crescita dei depositi è ancora più forte facendo scendere il rapporto impieghi/depositi da 0,84 a 0,72. Tuttavia, come mostrano i Grafici 3 e 7, durante la fase del *take off* dell'economia italiana, tra la fine dell'Ottocento e lo scoppio della prima guerra mondiale, e anche successivamente fino alle crisi bancarie degli anni Venti e Trenta, la dinamica dello sviluppo finanziario in Italia è chiaramente divergente. Gli impieghi in Lombardia crescono a tassi molto elevati portando la regione a essere nel 1936 il vero centro finanziario dell'Italia con un rapporto impieghi/Pil pari all'84,8%. Anche in Piemonte la dinamica degli impieghi è sostenuta (superiore al 2% annuale) ma a partire da un livello iniziale di sviluppo finanziario ben inferiore a quello lombardo. Nel Sud, la Sicilia e la Campania vivono un vero e proprio boom finanziario con gli impieghi che crescono, rispettivamente, del 9% e del 3,7% all'anno portando le due regioni al terzo e quarto posto nella graduatoria del rapporto impieghi/Pil (30,7% e 39,5%). Nelle altre regioni del Sud, però, la dinamica dello sviluppo finanziario è di segno completamente opposto, con Abruzzo, Puglia, Basilicata e Sardegna che vedono evolvere il rapporto impieghi/Pil a tassi negativi e che nel 1936 si assesta su livelli che raramente superano il 3% (in Sardegna si ar-

riva addirittura a un irrisorio 0,7‰). Nel complesso, la variabilità del rapporto impieghi/Pil tra le regioni italiane aumenta in maniera sensibile nel periodo considerato, passando la deviazione standard da 0,11 a 0,23 (cfr. Graff. 3a e 7a).

Ancor più divergente e soprattutto altalenante è l'andamento del rapporto tra impieghi e depositi che in Italia passa dallo 0,84 del 1890, allo 0,67 del 1900, allo 0,85 del 1910 allo 0,72 del 1936. Quasi tutte le regioni evidenziano nel periodo la stessa dinamica negativa (Graf. 7b). È interessante notare però come il processo abbia un andamento opposto tra le due aree del Paese: infatti, posto 100 il rapporto tra impieghi e depositi nel Centro-Nord il corrispondente indicatore nel Mezzogiorno passa dall'1,13 del 1890 allo 0,6 del 1920 per attestarsi nel 1936 a 1,05 (Graf. 7b). Sebbene divergente nei livelli, però, la variabilità di questo indicatore tra le regioni tende a ridursi costantemente nel periodo con la deviazione standard che tocca valori simili a quelli che si riproporranno solo negli anni Settanta.

GRAF. 7

CONVERGENZA BETA 1890-1936



Nota: **A. Impieghi/Pil.** Il primo anno analizzato è 1891. Sono escluse: Valle d'Aosta, Trentino Alto Adige, Friuli Venezia Giulia e Molise. La serie del Pil regionale è stata ricostruita a partire dai dati sul Pil *pro-capite* regionale a prezzi 1911 riportati in DANIELE V. e MALANIMA P. (2007) e convertiti in euro e dai dati sulla popolazione media annuale per regioni della SVIMEZ (2011a); nel caso di dati mancanti sulla popolazione è stata utilizzata la popolazione osservata nell'anno immediatamente precedente, $t-x$, a quello mancante, t , aumentata per il tasso di crescita medio annuale tra l'anno immediatamente precedente e l'anno immediatamente successivo a quello mancante, $g_{t-x:t+y}$, ossia: $Pop_t = Pop_{t-x} (1 + g_{t-x:t+y})^x$. La serie degli impieghi è quella pubblicata dalla SVIMEZ (2011a) calcolata in valori del 1911 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009). **B. Impieghi/Depositi.** Gli anni esaminati sono: 1890, 1895, 1900, 1905, 1910, 1929, 1936, 1938, 1939-2009. La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. Fino al 1920 sono escluse la Valle d'Aosta e il Trentino Alto Adige e per il 1936 è esclusa la Valle d'Aosta. **C. (Impieghi+Depositi)/Sportelli.** Gli anni considerati sono: 1936, 1939, 1946, 1951, 1956, 1961, 1966, 1971, 1975-2009. Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a); fino al 1966 è escluso il Molise, mentre per il 1936 è esclusa anche la Valle d'Aosta.

5.2. 1937-1950

Con l'adozione dei provvedimenti di riforma in campo bancario nel 1936 la geografia dell'industria bancaria assume la conformazione che manterrà sostanzialmente inalterata fino al 1986, con il 21% degli sportelli bancari localizzati nelle regioni del Sud, il 32% nelle regioni del Centro e il 47% in quelle del Nord (il 17% nella sola Lombardia). Il processo di consolidamento del sistema bancario attuato nel ventennio precedente e la svolta in senso dirigista sancita dal nuovo assetto legislativo si traduce in un deciso aumento della presenza delle banche e del volume degli impieghi per sportello nel Mezzogiorno cui si contrappone una contrazione del medesimo indicatore nel Centro-Nord. Tutti gli indicatori dello sviluppo finanziario evidenziano così un andamento convergente con le regioni finanziariamente più sviluppate che crescono a tassi fortemente negativi (Graf. 8a). Questo processo è già evidente alla vigilia dell'ingresso dell'Italia nel secondo conflitto mondiale e prosegue durante il periodo della guerra, che naturalmente vede un forte rallentamento delle attività di intermediazione finanziaria da parte delle banche (nel 1943 gli impieghi sono il 13,6% del Prodotto interno lordo e nel 1944 il 4,6%), per avere una leggera inversione di tendenza solo nell'immediato dopoguerra.

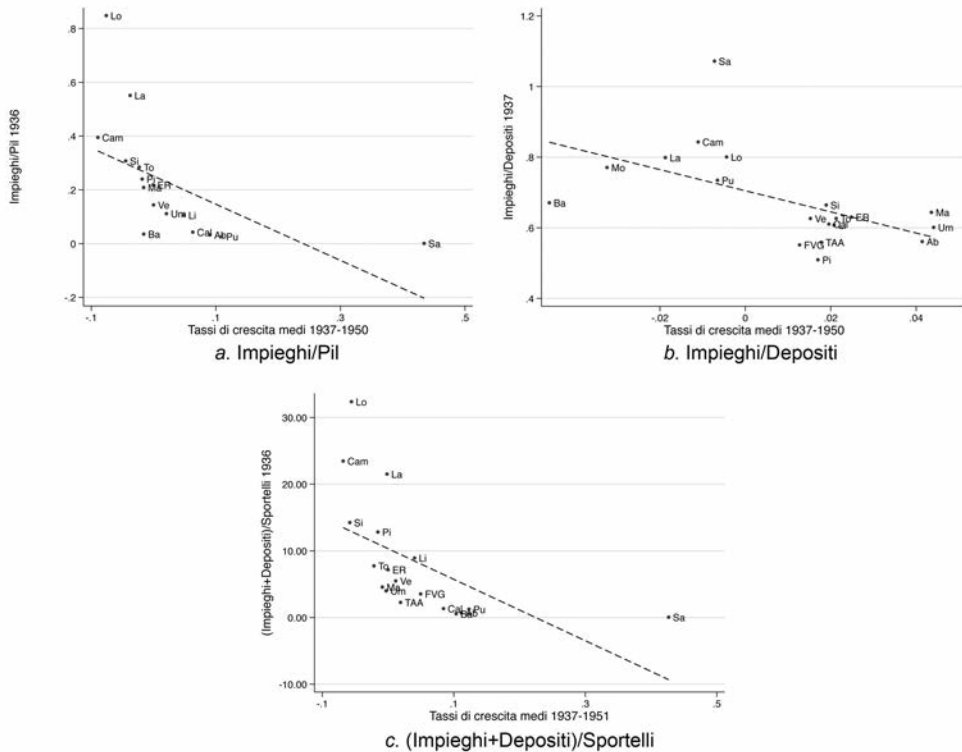
Nel 1938 il rapporto impieghi/Pil è nel Mezzogiorno pari al 68,8% di quello del Centro-Nord (nel 1920 era il 10,1%). In Lombardia, ad esempio, tale rapporto si riduce in media ogni anno del 7,6% fino a toccare nel 1950 quota 28,1%. Anche in Campania e Sicilia, le regioni del Sud finanziariamente più sviluppate, si assiste a un forte rallentamento degli impieghi bancari, mentre sono le regioni periferiche del Sud, Abruzzo, Puglia, Calabria e Sardegna a vedere crescere il rapporto impieghi/Pil. Anche la variabilità del fenomeno tra le regioni italiane ha una forte riduzione, con la deviazione standard che passa da 0,23 nel 1936, a 0,06 nel 1938, a 0,07 nel 1950.

Lo stesso andamento è evidenziato dal rapporto impieghi/depositi che si mantiene su livelli medi piuttosto contenuti: negli anni di pace raggiunge a stento il 70% e in quelli di guerra è ben al di sotto del 60%. Se si escludono gli anni immediatamente successivi alla conclusione del conflitto mondiale, il rapporto impieghi/depositi è nel Mezzogiorno sempre al di sopra dei valori espressi nelle regioni del Centro-Nord: nel 1938 nelle due aree gli impieghi erogati localmente sono pari, rispettivamente al 95% e al 60% dei depositi raccolti localmente, nel 1946 queste percentuali scendono drasticamente, ma in misura relativamente molto maggiore nel Mezzogiorno, attestandosi al 52,5% e al 53,3%, per poi risalire nel 1950 a 78,3% e 74,3% riportando il livello di attivazione del risparmio nelle regioni del Sud al di sopra di quello delle regioni del Centro-Nord (Graf. 8b).

Infine, mentre nelle regioni periferiche del Sud la produttività degli sportelli bancari regionali evidenzia tassi di crescita annuali a due cifre, in Piemonte, Lombardia, Lazio, Campania e Sicilia essa tende a ridursi nel tempo (Graf. 8c).

GRAF. 8

CONVERGENZA BETA 1937-1950



Nota: **A. Impieghi/Pil.** Sono escluse: Valle d'Aosta, Trentino Alto Adige, Friuli Venezia Giulia e Molise. La serie del Pil regionale è stata ricostruita a partire dai dati sul Pil *pro-capite* regionale a prezzi 1911 riportati in DANIELE V. e MALANIMA P. (2007) e convertiti in euro e dai dati sulla popolazione media annuale per regioni della SVIMEZ (2011a); nel caso di dati mancanti è stata utilizzata la popolazione osservata nell'anno immediatamente precedente, $t-x$, aumentata per il tasso di crescita medio annuale tra questo e l'anno immediatamente successivo a quello mancante, $g_{t-x,t}$, ossia: $Pop_t = Pop_{t-x} (1 + g_{t-x,t})^x$. La serie degli impieghi è quella pubblicata dalla SVIMEZ (2011a) calcolata in valori del 1911 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009). Per il periodo 1951-2009 le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). **B. Impieghi/Depositi.** La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. È esclusa la Valle d'Aosta. **C. (Impieghi+Depositi)/Sportelli.** Gli anni considerati sono: 1936, 1939, 1946, 1951, 1956, 1961, 1966, 1971, 1975-2009. Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a). Sono escluse Molise e Valle d'Aosta.

5.3 1951-1973

Strettamente correlati al cambiamento d'impostazione della politica economica e al ruolo crescente assunto nell'agenda politica nazionale dai problemi del riequilibrio territoriale Nord-Sud, si registrano dei lievi ma significativi mutamenti nell'articolazione territoriale del sistema bancario. Se si guarda al periodo tra il 1951, anno di avvio dell'intervento straordinario, e i primi anni Settanta, si osserva anzitutto un costante aumento della quota di sportelli che operano nel Mezzogiorno che dal 19,4% del totale nazionale passa nel 1974 al 23,9% (Tav. 1), riducendo così anche di poco il *gap* nel grado di bancarizzazione tra le due aree. Questo incremento è comune a tutte le regioni meridionali anche se non in maniera uniforme: in Sicilia, Basilicata e Molise il numero di sportelli più che raddoppia (da, rispettivamente, 503, 48 e 18 nel 1951 a 1032, 113, e 36 nel 1974), in Puglia l'aumento è dell'84%, in Campania e in Abruzzo del 64%. Più modesto è invece il numero dei nuovi insediamenti in Sardegna e Calabria, tanto che la quota di sportelli su base nazionale in queste regioni resta costante (Sardegna) o addirittura diminuisce (Calabria). Nel complesso, fatto 100 il rapporto tra il numero di sportelli e gli abitanti nel Centro-Nord, il tasso di bancarizzazione relativa delle regioni meridionali passa dal 40,6% del 1951 al 57,1% del 1973.

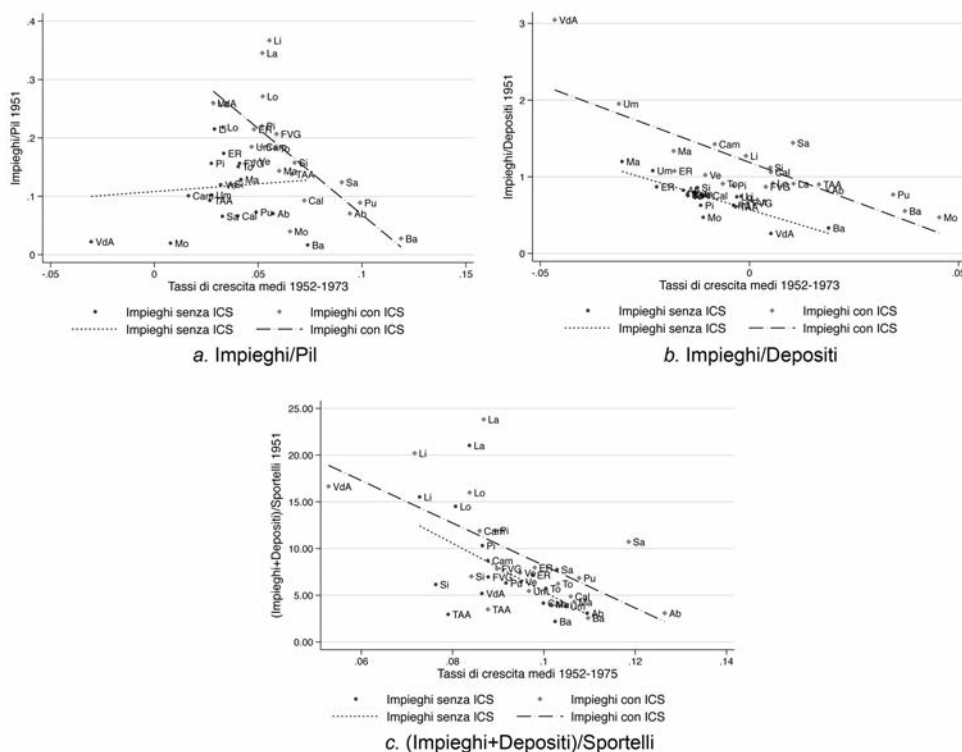
Passando a considerare i flussi finanziari, un primo chiaro riscontro della forte convergenza nello sviluppo finanziario regionale che si accompagna alle politiche attive per il Mezzogiorno si ritrova nella crescita del rapporto impieghi/Pil. Il Grafico 9a riporta l'andamento del rapporto impieghi/Pil sia nel caso in cui gli impieghi sono calcolati includendo solo quelli erogati dalle banche ordinarie sia nel caso in cui a questi si aggiungono gli impieghi attribuibili agli Istituti di Credito Speciale, statisticamente disponibili a partire dal 1951¹⁵. Il rapporto tra gli impieghi delle banche ordinarie e il Pil regionale sembra mostrare una lieve tendenza divergente dovuta essenzialmente ai tassi di crescita negativi o prossimi allo zero della Valle d'Aosta e del Molise (se si escludono queste due regioni la linea di tendenza torna ad avere una inclinazione negativa non dissimile a quella indicata con la linea tratteggiata che include gli impieghi degli Ics). Anche senza contare queste due regioni, però, gli impieghi al lordo di quelli erogati dagli Ics mostrano in tutte le regioni tassi di crescita più elevati di quelli delle sole banche ordinarie. Ciò è particolarmente vero per le regioni del Sud che, a motivo del sensibile aumento degli impieghi degli Ics nell'area, che tra il 1951 e il 1973 pas-

¹⁵ Al fine di garantire l'omogeneità della serie per l'intero sotto-periodo, i tassi di crescita sono stati calcolati a partire dal 1952.

sano dal 19% del totale nazionale al 30%, si collocano tutte nella parte bassa del grafico. Così in Basilicata, per esempio, il rapporto (impieghi con Ics)/Pil più che decuplica, passando dal 2,7% del 1951 al 32,6% del 1973, mentre in Puglia e Sardegna tale rapporto cresce, rispettivamente, di otto e sette volte.

GRAF. 9

CONVERGENZA BETA 1951-1973



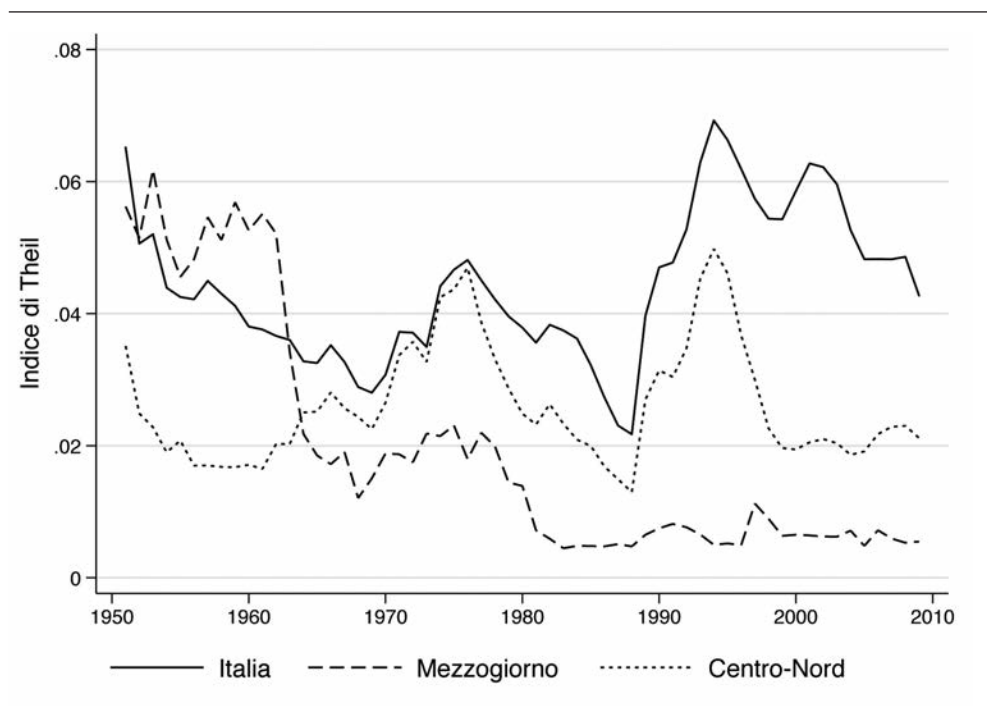
Nota. **A. Impieghi/Pil.** Le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). Gli impieghi sono calcolati sia al lordo sia al netto degli impieghi degli Istituti di Credito Speciale. **B. Impieghi/Depositi.** La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. **C. (Impieghi+Depositi)/Sportelli.** Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a).

È interessante però notare che all'andamento convergente dei tassi di crescita medi regionali del rapporto impieghi/Pil si accompagna un andamento divergente della variabilità regionale di tale indicatore che in questo periodo tende chiaramente a aumentare, non solo a livello nazionale (Graf. 3a), ma anche a livello

delle due grandi ripartizioni Mezzogiorno e Centro-Nord, indicando che alla convergenza nei livelli di sviluppo finanziario tra regioni non si accompagna una maggiore omogeneità territoriale nell'attività di impiego delle banche. Se però si tiene conto anche della dimensione delle regioni attraverso l'indice di entropia di Theil¹⁶, ciò che emerge è una complessiva riduzione delle ineguaglianze nei livelli di sviluppo finanziario indotta dalla evoluzione delle regioni del Mezzogiorno che mentre a inizio periodo apparivano tra loro molto più dissimili di quanto non fossero, tra loro, le regioni del Centro-Nord, nel 1973 evidenziano una distribuzione degli impieghi per regione sostanzialmente coincidente con quella del prodotto interno lordo (cfr. Graf. 10).

GRAF. 10

IMPIEGHI/PIL: INDICE DI THEIL



¹⁶ L'indice di Theil è dato da: $T = \sum_{i=1}^n imp_i \ln (imp_i / Pil_i)$, dove imp_i e Pil_i rappresentano rispettivamente le quote degli impieghi e del Pil attribuibili alla regione i -esima.

Coerentemente con il forte impegno pubblico per gli investimenti nel Sud e l'iniezione di consistenti risorse finanziarie esterne in questa parte del Paese, il rapporto (impieghi con Ics)/depositi, che nel 1951 era a quota 1,16, un livello del 19,5% più alto di quello prevalente nelle regioni del Centro-Nord, si mantiene in questo periodo stabilmente molto al di sopra dell'unità attestandosi nel 1973 a 1,35 (contro lo 0,95 delle altre regioni del Paese). Se si escludono invece gli ICS, sono le regioni del Centro-Nord che evidenziano una relativa maggiore capacità di attivazione del risparmio prodotto localmente, anche se nel complesso la dinamica del rapporto (impieghi senza Ics)/depositi è chiaramente negativa nella maggioranza delle regioni italiane (Graf. 9b). Conseguentemente, nel periodo indicato, si osserva una leggera convergenza nella capacità delle banche di attivazione del risparmio locale, ma soprattutto si osserva una drastica riduzione della variabilità del fenomeno tra le regioni, con la deviazione standard che si riduce a meno della metà, in modo particolare nel Centro-Nord. Se si guarda alle singole regioni, gli incrementi più significativi del rapporto impieghi/depositi si hanno in Puglia e Basilicata (che passano rispettivamente da 0,77 e 0,55, nel 1951, a 1,61 e 1,24, nel 1973) mentre le riduzioni più consistenti si verificano in Veneto, Emilia-Romagna e Umbria (oltre che in Valle d'Aosta).

Infine, si osserva un incremento del livello di produttività del sistema bancario nelle regioni meridionali che può essere colto attraverso l'andamento del rapporto tra la somma degli impieghi e dei depositi sul numero degli sportelli. Gli impieghi e i depositi intermediati per sportello nel Mezzogiorno aumentano in misura sensibile rispetto a quanto si verifica nel resto del Paese: si passa infatti dal 65,2% del 1951 al 71,3% del 1971, con gli incrementi più significati nelle regioni Campania e Sardegna (Graf. 9c).

5.4 1974-1989

Coerentemente con la nuova stagione di politiche per il Mezzogiorno rivolte più al sostegno della domanda che allo sviluppo dell'offerta, un dato che caratterizza questa fase è la forte riduzione degli impieghi degli Istituti di Credito Speciale (ICS) nelle regioni del Sud che scendono ininterrottamente dal 30% del totale nazionale nei primi anni '70 al 23% negli ultimi anni del decennio successivo, senza trovare un'adeguata compensazione nei flussi di credito a breve termine la cui quota nel Mezzogiorno rispetto al totale nazionale cresce nel periodo di poco più di 2 punti percentuali, da 15,3% a 17,7%. Il rallentamento dell'attività del credito mobiliare comporta un riassetto della struttura finanziaria delle imprese, all'interno della quale i finanziamenti bancari a breve termine tornano a

essere la voce più consistente (alla fine del 1973 rappresentavano il 46% degli impieghi complessivamente erogati nell'area, per diventare il 60% nel 1988). Tutto ciò favorisce assetti del sistema produttivo che necessitano di una minore utilizzazione di finanziamenti a lungo termine: in particolare, nel corso di questa fase le imprese tendono a ridimensionare la capacità produttiva a disposizione attraverso l'eliminazione, mediante cessione ad altre imprese, dei beni capitali economicamente più obsoleti (Giannola e Lopes, 2011).

La progressiva perdita di efficacia delle politiche attive nel Mezzogiorno e il crescente restringimento della sua base produttiva hanno un corrispettivo in ambito finanziario con la crescita relativa del rapporto tra impieghi/Pil al Centro-Nord rispetto a quello meridionale: se all'inizio del periodo tale rapporto nelle regioni centro-settentrionali eccedeva quello meridionale nella misura del 22%, nel 1988 lo scarto tocca il 35,6%.

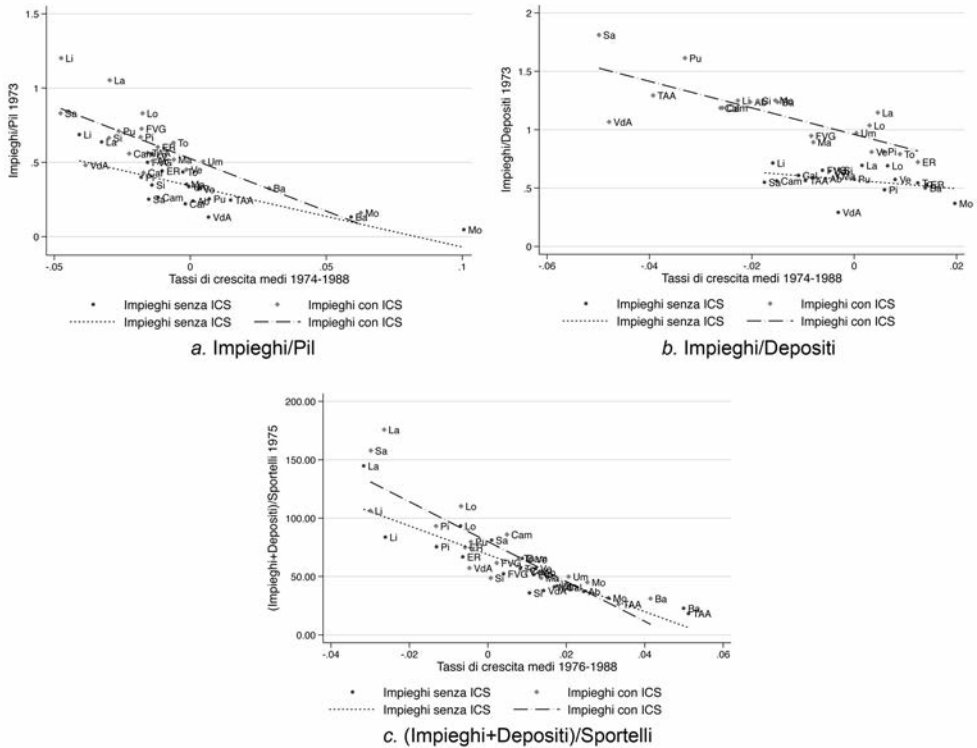
Ciò costituisce un fattore di deterioramento del quadro macroeconomico soprattutto per ciò che concerne le attività produttive in un momento in cui il progressivo ridimensionamento della grande impresa, con le sue ricadute negative in termini di investimenti e occupazione, non viene compensato adeguatamente dalle imprese locali di minori dimensioni che scontano una serie di criticità di carattere economico-finanziario, gestionale e tecnologico che ne condizionano il finanziamento degli investimenti attraverso il credito bancario e ne impediscono la crescita¹⁷. Inoltre, il drastico ridimensionamento delle politiche attive di industrializzazione in favore di politiche a sostegno dei redditi e il contenimento dei flussi finanziari intermediati dal sistema bancario determina una brusca impennata del residuo fiscale (la differenza tra le entrate raccolte e le spesa effettuate dal settore pubblico in una determinata area territoriale) nelle regioni del Mezzogiorno¹⁸.

Come emerge dal Grafico 11a, il processo di convergenza finanziaria prosegue in una fase di sostanziale contrazione del rapporto impieghi/Pil che solo risulta più sostenuta nelle regioni finanziariamente più sviluppate. In una fase di complessivo rallentamento dell'attività di impiego delle banche, la variabilità del rapporto impieghi/Pil tra le regioni italiane tende a ridursi in modo particolarmente pronunciato proprio nelle regioni del Sud (Graf. 3a).

¹⁷ Si vedano le verifiche empiriche riportate in GIANNOLA A. e LOPES A. (1992) e in GIANNOLA A. e SARNO D. (2004).

¹⁸ Secondo una recente ricostruzione effettuata da GIANNOLA A., PETRAGLIA C. e SCALERA D. (2012), nel Mezzogiorno il residuo fiscale *pro-capite*, calcolato in euro 2010, passa da valori medi di -866, per il periodo 1962-1965, a -1.759, nel periodo 1971-1973, a -3.381, nel periodo 1983-1985.

CONVERGENZA BETA 1973-1988



Nota: **A. Impieghi/Pil.** Le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). Gli impieghi sono calcolati sia al lordo sia al netto degli impieghi degli Istituti di Credito Speciale. **B. Impieghi/Depositi.** La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. **C. (Impieghi+Depositi)/Sportelli.** Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a).

Il differente andamento dei depositi e degli impieghi evidenziato in precedenza si traduce in una corrispondente riduzione relativa del rapporto impieghi/depositi nel Mezzogiorno che per la prima volta dal secondo dopoguerra scende al di sotto dell'analogo rapporto per il Centro-Nord attestandosi nel 1988 all'80% (all'inizio del periodo nel Mezzogiorno il rapporto impieghi/depositi era del 41% più alto che nel Centro-Nord). Alla vigilia della soppressione dell'intervento straordinario il Mezzogiorno si configura sempre più come un'area sussidiata nella quale l'intermediazione finanziaria svolge un'importante funzione riallocativa dei flussi finanziari attraverso un'azione che vede le regioni meridionali sempre più come

un'area di raccolta di fondi che saranno poi orientati verso gli impieghi produttivi nel resto del Paese (Graf. 11*b*).

Infine, l'esame dell'andamento relativo del rapporto (Impieghi+depositi)/sportelli permette di cogliere altri aspetti che il processo di integrazione dipendente del Mezzogiorno assume in questi anni quando il sostegno dei redditi e dei consumi è prevalentemente assicurato dall'azione redistributiva della spesa pubblica; ciò consente al sistema bancario di mantenere e di espandere la raccolta che, come detto, troverà poi impiego in altre aree del Paese. Va osservato che per le regioni meridionali, relativamente al Centro-Nord, il rapporto tra depositi e sportelli tende ad aumentare: posto pari a 100 il rapporto depositi per sportello del Centro-Nord, l'analogo rapporto per il Mezzogiorno passa dal 66% del 1975 all'87% del 1988¹⁹. Pertanto, se il venir meno di politiche attive di promozione e di ampliamento della base industriale impedisce alle regioni meridionali uno sviluppo finanziario paragonabile a quello sperimentato fino ai primi anni '70, la presenza di un massiccio processo di redistribuzione assistenziale intermediato dal sistema finanziario operante nell'area consente a quest'ultimo di mantenersi in certa misura "allineato" agli standard di produttività non certo entusiasmanti che caratterizzano le strutture bancarie nel resto del Paese (Graf. 11*c*).

5.5 1989-1996

La questione concernente l'individuazione di un'efficace politica di intervento negli anni '80, tale da consentire alle regioni meridionali di rompere il circolo vizioso della dipendenza, è fondamentalmente elusa fidando sulle potenzialità dello sviluppo autopropulsivo il cui fallimento induce a identificare gradualmente ma inesorabilmente il Mezzogiorno come un sistema sociale e politico dove hanno origine, si sedimentano e si concentrano storture e vizi capitali della società italiana: sprechi, inefficienza, clientelismo, criminalità. Ciò porterà nel 1992 alla precipitosa chiusura dell'intervento straordinario²⁰.

La difficoltà per le aziende di credito meridionali di trovare adeguate occasioni di sviluppo della propria attività di intermediazione è infatti evidenziata dall'analisi dell'andamento del rapporto impieghi/depositi: per tutti gli anni '90 esso tende a aumentare progressivamente nelle regioni centro-settentrionali rispetto all'analogo indicatore calcolato per il Mezzogiorno (cfr. Graf. 6*a*). Se, da un lato, in questi anni si realizza nel Mezzogiorno una crescita della raccolta in linea con

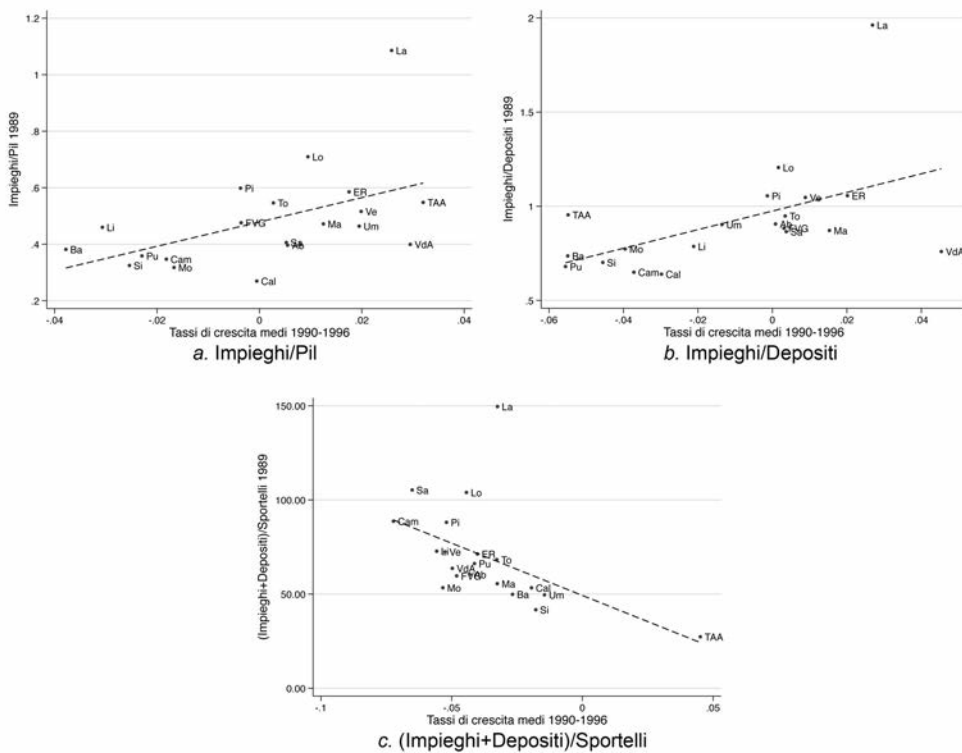
¹⁹ Si rinvia ancora a SVIMEZ (2011*a*).

²⁰ Si vedano su questo punto le considerazioni svolte in GIANNOLA A. e PETRAGLIA C. (2007).

quella delle altre aree del Paese, la quota del credito erogato nell'area tende a contrarsi progressivamente. Questo fenomeno determina l'interruzione del processo di convergenza dello sviluppo finanziario a esclusione dell'indicatore di produttività degli sportelli (cfr. Graf. 12)²¹.

GRAF. 12

CONVERGENZA BETA 1989-1996



Nota: **A. Impieghi/Pil.** Le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). **B. Impieghi/Depositi.** La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. **C. (Impieghi+Depositi)/Sportelli.** Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a).

²¹ I tassi di crescita sono stati calcolati a partire dal 1990 in maniera da mantenere l'omogeneità della serie per l'intero periodo.

In particolare, per quanto riguarda l'andamento del rapporto impieghi/Pil, dal Grafico 6a si evince che a fronte di un rapporto che, nel 1992, era pari al 42% nel Mezzogiorno e al 73% nel Centro-Nord, si scende rispettivamente al 35% e al 66% nel 1996; Campania, Calabria e Sicilia sono le regioni in cui il processo di divaricazione dei flussi finanziari si profila in misura più accentuata. La variabilità di questo indicatore di sviluppo finanziario resta sostanzialmente invariata nel periodo, anche se l'indice di entropia di Theil evidenzia una disomogeneità crescente delle regioni centro-settentrionali rispetto a quelle meridionali (cfr. Graff. 6a e 10) tra il 1989 e il 1996.

Uguualmente, tra il 1989 e il 1996 il rapporto impieghi/Pil cresce ovunque nel Mezzogiorno a tassi negativi o (è il caso dell'Abruzzo e della Sardegna) resta sostanzialmente invariato. L'unico indicatore di sviluppo finanziario che in questo periodo evidenzia una dinamica convergente, sia pure in un periodo di complessiva contrazione, è quello della produttività degli sportelli, (impieghi+depositi)/sportelli, in linea con gli sforzi di consolidamento del sistema bancario italiano che interessano in misura decisamente maggiore le banche meridionali e che conducono a una evidente divaricazione nel processo di bancarizzazione tra le due aree e soprattutto a un forte ingresso delle banche esterne nell'area, impegnate in misura consistente nell'attività di raccolta (cfr. Graf. 5).

5.6 1997-2009

A partire dalla seconda metà degli anni '90, il progressivo deterioramento del settore reale dell'economia meridionale si ripercuote significativamente sul suo sistema bancario e sui suoi assetti proprietari. Il consolidamento del sistema bancario assume una forte caratterizzazione territoriale che si traduce in una penetrazione delle banche con sede legale nel Centro-Nord nei mercati meridionali.

In più occasioni in questi ultimi anni ci si è chiesti se il processo di ristrutturazione del sistema bancario meridionale e della sua rete di sportelli abbia o meno favorito il buon funzionamento del mercato del credito nel Mezzogiorno e, in particolare, se abbia agevolato l'accesso al credito delle imprese del Sud²².

Certo è che, come mostrato da numerose evidenze empiriche²³, la prossimità geografica dei centri decisionali (ancor più di quella della rete distributiva) consente di ridurre le barriere informative tra banca e cliente, tanto più quando la

²² Per i termini del dibattito si rinvia nuovamente a GIANNOLA A. (2002) e a CANNARI L., PANNETTA F. (2006).

²³ Si vedano in particolare ALESSANDRINI P., PRESBITERO A., ZAZZARO A. (2009) e ALESSANDRINI P., PRESBITERO A., ZAZZARO A. (2009).

clientela è composta da piccole imprese, opache dal punto di vista informativo; d'altro canto gli stessi recuperi di efficienza gestionale che ci si attendeva dal processo di consolidamento del sistema bancario nazionale e meridionale non sono stati quelli attesi, soprattutto per le banche e i gruppi di dimensione maggiore rispetto alle banche minori, e in particolare agli istituti di credito cooperativo, che hanno evidenziato nel corso degli ultimi dieci anni *performance* gestionali significativamente migliori²⁴.

D'altra parte l'importanza che le banche italiane attribuiscono alla prossimità alla clientela è ben evidenziata dal forte incremento nell'apertura di nuovi sportelli a cui si assiste in questo periodo in tutto il territorio nazionale, in misura preminente nelle regioni del Centro-Nord. Come in molti altri paesi, anche in Italia la maggior parte dei prestiti alle imprese è erogata da sportelli ubicati nella provincia e nel comune di residenza dell'impresa e la distanza media tra banca e cliente è di pochissimi chilometri (Bellucci *et al.*, 2010). Tuttavia, occorre rilevare che la proliferazione dei presidi locali è avvenuta di pari passo con la riduzione del numero di addetti per sportello e con l'impovertimento delle loro funzioni e dei loro ambiti operativi. Come evidenzia l'indagine condotta nel 2007 dalla Banca d'Italia sull'organizzazione dell'attività creditizia e l'impiego delle tecniche di *credit scoring* da parte delle banche italiane (Albareto *et al.*, 2008), la lunghezza della scala gerarchica è notevolmente più lunga nelle banche di grandi dimensioni, così come il decentramento decisionale è relativamente minore (anche se la delega ai direttori di filiale sull'ammontare di fido massimo concedibile è più elevata). Nella misura in cui il grado di decentramento era giudicato dagli intervistati in calo rispetto al triennio precedente per quanto riguarda le banche maggiori e sostanzialmente stabile per le banche minori, questo implica che nel periodo considerato l'autonomia concessa alle filiali di banche esterne operanti nel Mezzogiorno era davvero molto contenuta. Anche per quanto riguarda la diffusione delle tecniche di *scoring* per la concessione dei fidi, l'indagine della Banca d'Italia evidenzia una netta differenza tra banche maggiori e banche piccole e minori, differenza che appare relativamente ancora più marcata se si guarda al solo credito alle piccole imprese.

Lo spostamento delle funzioni decisionali e strategiche delle banche acquisite al vertice della struttura, la contestuale adozione di modelli quantitativi per la valutazione del rischio e la crescente standardizzazione del prodotto bancario offerto sono tutti inequivocabili segni dell'impovertimento progressivo del contenuto informativo intangibile insito del rapporto banca-impresa (Giordano e Lopes, 2009*b*).

²⁴ Si veda su questo punto GIORDANO L., LOPES A. (2009*a*).

In definitiva, e in particolare per i grandi gruppi, l'esigenza di migliorare l'efficienza allocativa si è sostanzialmente tradotta in una strategia di razionalizzazione dei processi lavorativi presso le filiali con la standardizzazione delle attività di vendita e l'accentramento di funzioni decisionali presso le sedi centrali. Tale svuotamento della prossimità fisica mediante la standardizzazione dell'*output* e la crescente distanza funzionale tra direzione strategica e unità operativa si è realizzata meccanicamente e con particolare intensità nei confronti della (piccola e media) clientela meridionale, come prevedibile conseguenza delle fusioni ed acquisizioni "esterne" che hanno fortemente ridimensionato la presenza di banche meridionali indipendenti. Tutto ciò induce a una lettura del consolidamento del sistema bancario italiano più problematica e meno rassicurante di quella proposta da altri in questi anni (ad esempio, cfr. Cannari e Panetta, 2006), secondo la quale le fusioni e le acquisizioni delle banche meridionali hanno consentito alle banche "esterne" che ne sono state protagoniste attive di sfruttare la rete geografica di distribuzione delle prime per la collocazione di prodotti finanziari non differenziati e di attingere alla raccolta del risparmio locale. Questo ha rappresentato spesso il principale se non esclusivo "valore" dell'operazione.

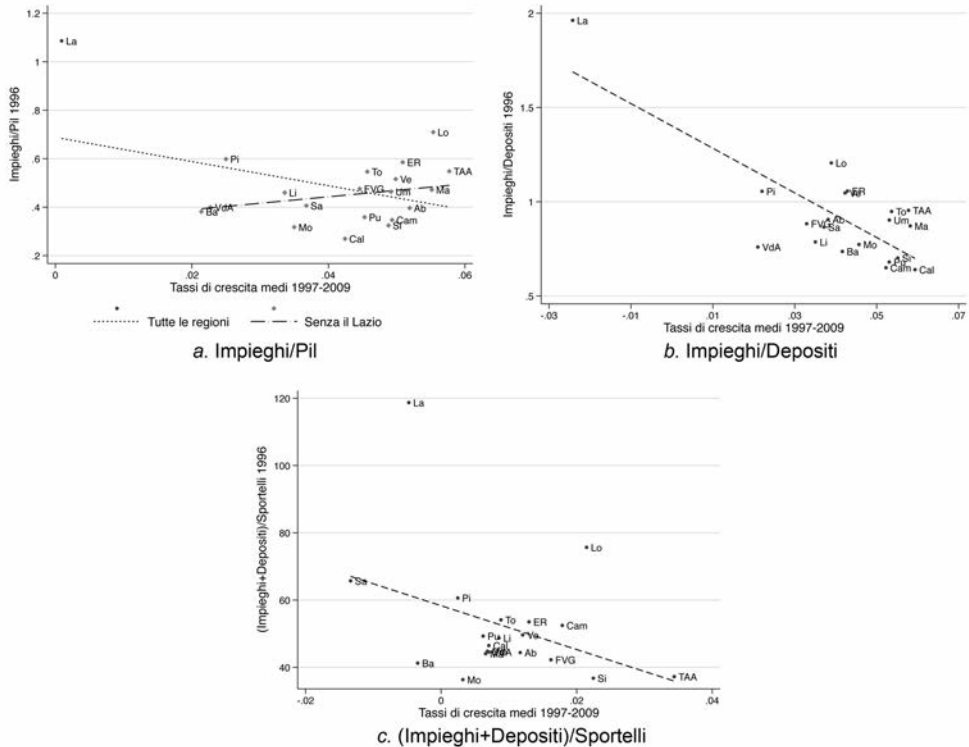
Sebbene si assista a una lieve ripresa del processo di convergenza finanziaria, con tassi di crescita positivi in quasi tutte le regioni ma leggermente più elevati per quelle che nel 1989 si ponevano su livelli di sviluppo finanziario inferiori (Graff. 13*b* e 13*c*), il Mezzogiorno nel suo complesso continua nel periodo ad accumulare ritardi. Una chiara indicazione al riguardo è data dal rapporto impieghi/depositi che, raggiunto il suo minimo assoluto nel 1996 (0,71), continua a divergere nelle regioni del Sud dai valori prevalenti nelle regioni del Centro-Nord arrivando, nel 2001, a un differenziale di quasi 90 punti percentuali (0,98 al Sud e 1,87 al Centro-Nord). Successivamente, si osserva un progressivo riavvicinamento tra le due aree e soprattutto un incremento assoluto del rapporto impieghi/depositi in entrambe le aree in linea con l'aumentato ricorso a forme di raccolta indiretta (Graf. 6*b*).

A risultati più netti si giunge se si considera il rapporto impieghi/Pil. Anche in questo caso, ciò che si nota è che a partire dal 1996 il divario a sfavore del Mezzogiorno tende ad accentuarsi progressivamente fino al 2001, quando il rapporto risulta pari appena al 44% di quello prevalente nelle altre regioni del Paese. Il modesto recupero cui si assiste negli anni successivi non consente di ritornare ai valori (assoluti e relativi) toccati all'inizio degli anni Settanta. A conferma di ciò, sia che si guardi alla *convergenza beta* sia che si guardi alla *convergenza sigma*, se si esclude dal campione la regione Lazio, l'andamento del rapporto impieghi/Pil

mostra un andamento divergente tra le regioni con la linea di tendenza che assume una pendenza positiva (Graf. 13a).

GRAF. 13

CONVERGENZA BETA 1997-2009



Note. **A. Impieghi/Pil.** Le serie del Pil e degli impieghi sono in euro correnti (SVIMEZ, 2011a). La linea di tendenza tratteggiata è ricavata escludendo la regione Lazio. **B. Impieghi/Depositi.** La serie dei depositi è quella pubblicata dalla SVIMEZ (2011a) calcolata in euro correnti. **C. (Impieghi+Depositi)/Sportelli.** Gli impieghi e i depositi sono calcolati in valori del 2008 utilizzando i coefficienti di conversione annuali pubblicati dall'ISTAT (2009), mentre gli sportelli regionali sono tratti da SVIMEZ (2011a).

6. - Gli anni della crisi (2008-2011): alcune considerazioni conclusive

In definitiva, alla vigilia della crisi finanziaria globale che si è abbattuta sull'economia italiana nell'autunno 2008 con il fallimento di *Lehman Brothers* e che ha avuto una drammatica accelerazione nell'autunno 2011 con la crisi dei debiti sovrani nell'area euro, si aveva più di un motivo per ritenere che il processo di

riorganizzazione del sistema bancario, dal punto di vista delle regioni più sfavorite, non fosse stato in grado di mettere a disposizione del sistema imprenditoriale locale risorse necessarie in misura adeguata a fronteggiare le sfide della riconversione di larghi strati dell'apparato industriale, che in altre aree del Paese stava faticosamente dando i suoi primi frutti. La crisi ha poi colpito le regioni meridionali in misura relativamente maggiore rispetto ad altre aree del Paese e anche le difficoltà ad uscirne appaiono più gravi. Tra il 2008 e il 2011 il Mezzogiorno registra un tasso di variazione cumulato del Pil reale di segno negativo pari al -6,1% contro un -4,3% registrato dalle regioni del Centro-Nord (Svimez, 2011*b* e Banca d'Italia, 2012). Il divario tra le due aree del Paese in termini di prodotto *pro capite* subisce dunque un nuovo allargamento passando dal 59,8% nel 2003 al 58% nel 2011. Le esportazioni meridionali calano di circa l'11% nel corso del secondo semestre del 2008 per poi precipitare del 35% nel primo semestre del 2009 (nello stesso periodo il dato nazionale è del -4% e -24,2%). Allo stesso modo drammatica è la caduta dell'occupazione.

Dopo la deludente fase della NP si assiste a un ulteriore disimpegno delle politiche attive in favore del Mezzogiorno, che dal 2008 si traduce in un sostanziale azzeramento. Tutti gli interventi di incentivazione che si sarebbero dovuti attivare sono rimasti non operativi: le zone franche urbane, i nuovi contratti di programma, i contratti di localizzazione, i contratti di sviluppo. Nel contempo, si sono definitivamente esaurite le risorse finanziarie per i crediti d'imposta a favore dell'occupazione e dei nuovi investimenti e si è determinata la riallocazione di una parte consistente delle risorse Fas (non meno di 20 miliardi di euro), originariamente destinate per almeno l'85% alla spesa in conto capitale per lo sviluppo delle regioni meridionali, a copertura delle azioni anticrisi che sono state portate avanti in altre aree. A fare le spese della crisi e della crescente avversione per le politiche di incentivazione è stata, ancor più della politica industriale, la politica di riequilibrio territoriale, che ha finito per pagare pesantemente l'inevitabile contrazione delle risorse disponibili²⁵. Ancora una volta, dunque, sembra ripetersi la scelta politica compiuta in occasione delle crisi petrolifere degli anni '70, e già al-

²⁵ Nel 2008 gli aiuti di Stato rappresentavano nell'Ue a 27 lo 0,5% del Pil: in Germania il livello degli aiuti è il doppio rispetto all'Italia (0,6% del Pil nazionale contro lo 0,3%), e anche Francia e Spagna riservano agli aiuti una maggiore attenzione concedendo rispettivamente lo 0,5% e lo 0,4% del proprio Pil. Inferiore è anche la quota di spesa destinata al superamento degli squilibri territoriali, appena il 18,3%, contro il 22,8% della Germania, il 40,8% della Francia, il 39,9% della Spagna e, addirittura, il 25,8% della Ue a 27. Si rinvia per maggiori dettagli a SVIMEZ (2009, 2010, 2011*b*).

l'epoca aspramente criticata (Svimez, 1980), di far ricadere gli oneri del finanziamento della lotta alla crisi sulla spesa per gli investimenti nel Mezzogiorno.

Se si guarda al mercato del credito, dopo anni di crescita del volume degli impieghi, il rallentamento dell'attività creditizia è apparso subito evidente per trasformarsi in una contrazione dopo il fallimento di *Lehman Brothers*. Da tutte le indagini campionarie condotte sulle imprese arriva la conferma che si è in presenza di un inasprimento delle condizioni di accesso al credito per le imprese e di un forte aumento dei fenomeni di razionamento del credito (Panetta e Signoretti, 2010). Tale aumento è stato più marcato nelle province dove il sistema bancario locale è prevalentemente formato da sportelli bancari di banche nazionali, localizzate in regioni distanti (Barboni e Rossi, 2012; Presbitero *et al.*, 2012) e per le imprese localizzate nel Mezzogiorno (Albareto e Finaldi Russo, 2012; Costa *et al.*, 2012).

Nel nuovo equilibrio post-crisi – anche in considerazione dei più stringenti requisiti prudenziali imposti dalla progressiva introduzione degli accordi di Basilea 3 – una significativa espansione del credito richiederà un adeguato aumento dei mezzi patrimoniali di elevata qualità e un attento contenimento della rischiosità dei nuovi crediti. Questo rischia di avere ancora una volta delle ricadute negative sullo sviluppo finanziario delle aree deboli e sull'industria e le imprese locali.

BIBLIOGRAFIA

- A'HEARN B., «Could Southern Italians Cooperate? Banche Popolari in the Mezzogiorno», *The Journal of Economic History*, vol. 60(1), 2000, pages 67-93.
- ALBARETO G. - BENVENUTI M. - MOCETTI S. - PAGNINI M. - ROSSI P., «L'organizzazione dell'attività creditizia e l'utilizzo di tecniche di scoring nel sistema bancario italiano: risultati di un'indagine campionaria», Banca d'Italia, *Questioni di Economia e Finanza*, n. 12, aprile, 2008.
- ALBARETO P. - FINALDI RUSSO P., «Fragilità finanziaria e prospettive di crescita: il razionamento del credito alle imprese durante la crisi», Banca d'Italia, *Questioni di Economia e Finanza*, n. 127, luglio, 2012.
- ALBARETO G. - TRAPANESE M., «La politica monetaria negli anni cinquanta», in COTULA F. (a cura di), *Stabilità e sviluppo negli anni cinquanta*, vol. 3, *Politica bancaria struttura del sistema finanziario, e sviluppo negli anni cinquanta*, Roma-Bari, Editori Laterza, 1999.
- ALESSANDRINI P. - CROCI M. - ZAZZARO A., «The Geography of Banking Power: The Role of Functional Distances», BNL, *Quarterly Review*, vol. 58, 2005, pages 29-167.
- ALESSANDRINI P. - PRESBITERO A. - ZAZZARO A., «Banks, Distances and Firms' Financing Constraints», *Review of Finance*, vol. 13(2), 2009a, pages 261-307.
- , -, -, «Bank Size or Distance: What Hampers Innovation Adoption by SMEs», *Journal of Economic Geography*, no. 10(6), 2009b, pages 845-881.
- ANTZOULATOS A. - PANOPOULOU E. - TSOUMAS C., «Do Financial Systems Converge?», *Review of International Economics*, vol. 19, 2010, pages 122-136.
- BAGELLA M., *Gli istituti di credito speciali e il mercato finanziario (1947-1962)*, Milano, Franco Angeli, 1986.
- BALDUCCI R. - MARCONI M. - MARINELLI M.L. - NICCOLI A., «L'evoluzione degli intermediari e dei mercati finanziari: 1950-1972», in VICARELLI F. (a cura di), *Capitale industriale e capitale finanziario: il caso italiano*, Bologna, il Mulino, 1979.
- BANCA D'ITALIA, *Economie regionali*, n. 2, giugno, 2012.
- BARBONI G. - ROSSI C., *Does Your Neighbour Know You Better? Local Banks and Credit Tightening in the Financial Crisis*, Banca d'Italia, Convegno su «Le trasformazioni dei sistemi produttivi locali», Bologna, gennaio, 2012.
- BARCA F., *Italia frenata. Paradossi e lezioni della politica per lo sviluppo*, Roma, Donzelli, 2006.
- BARRO R.J. - SALA-I-MARTIN X., «Convergence», *Journal of Political Economy*, vol. 100 (2), 1992, pages 223-251.
- BECK T.H.L. - DEMIRGÜC-KUNT A. - LEVINE R., «Financial Institutions and Markets Across Countries and Over Time. Data and Analysis», World Bank, *Policy Research Working Paper*, no. 4943, 2009.

- BELLUCCI A. - BORISOV A. - ZAZZARO A., «Does Gender Matter in Bank-Firm Relationships? Evidence from Small Business Lending», *Journal of Banking and Finance*, vol. 34, 2010, pages 2968-2984.
- BIANCO M. - GERALI A. - MASSARO R., «Financial Systems Across Developed Economies: Convergence or Path Dependence?», *Research in Economics*, vol. 51, 1997, pages 303-331.
- BISCAINI A.M. - COTULA F. - CIOCCA P.L., «Le strutture finanziarie: aspetti quantitativi di lungo periodo (1870-1979)», in VICARELLI F. (a cura di), *Capitale industriale e capitale finanziario: il caso italiano*, Bologna, il Mulino, 1979.
- BONGINI P. - FERRI G., *Il sistema bancario meridionale*, Roma-Bari, Laterza, 2005.
- BUSETTA P. - SACCO S., «Razionamento, concentrazione e segmentazione nel mercato creditizio del Mezzogiorno», in ALESSANDRINI P. (a cura di), *Il sistema finanziario italiano tra globalizzazione e localismo*, Bologna, il Mulino, 2001.
- CANNARI L. - PANETTA F., *Il Sistema finanziario e il Mezzogiorno*, Bari, Cacucci editore, 2006.
- CIOCCA P., *Interesse e profitto. Saggi sul sistema creditizio*, il Mulino, Bologna, 1982.
- .-, *Ricchi per sempre? Una storia economica d'Italia (1796-2005)*, Torino, Bollati-Boringhieri, 2007.
- COLOMBO L. - TURATI G., «La dimensione territoriale nei processi di concentrazione dell'industria bancaria italiana», *Studi Economici*, n. 84, 2004, pp. 59-89.
- CONFALONIERI A., *Banca e industria in Italia 1894-1906*, Milano, Banca Commerciale Italiana, 1976.
- CONIGLIANI C., *La concentrazione bancaria in Italia*, Bologna, il Mulino, 1990.
- CONTI G., «Banche e imprese medie e piccole nella periferia economica italiana (1890-1939)», in CESARINI F. - FERRI G. - GIARDINO M. (a cura di), *Credito e sviluppo*, Bologna, il Mulino, 1997.
- CONTI G. - FERRI G., «Banche locali e sviluppo economico decentrato», in BARCA F. (a cura di), *Storia del capitalismo italiano dal dopoguerra a oggi*, Roma, Donzelli, 1997.
- COSTA S. - MALGARINI M. - MARGANI P., *Access to Credit for Italian Firms: New Evidence from the ISTAT Business Confidence Surveys*, paper presentato alla 53ª RSA della Società Italiana degli Economisti, Matera, 2012.
- CROCE P., «Il credito industriale nel Mezzogiorno: il caso dell'Isveimer», in COTULA F. (a cura di), *Stabilità. In e sviluppo negli anni cinquanta*. vol. 3. *Politica bancaria struttura del sistema finanziario e sviluppo negli anni cinquanta*, Roma-Bari, Editori Laterza, 1999.
- DANIELE V. - MALANIMA P., «Il prodotto delle regioni e il divario Nord-Sud in Italia (1861-2004)», *Rivista di Politica Economica*, vol. 97, marzo-aprile, 2007, pp. 267-315.
- D'ANTONE L., «Da ente transitorio a ente permanente», in CASTRONOVO V. (a cura di), *Storia dell'IRI. I. Dalle origini al dopoguerra (1933-1948)*, Roma-Bari, Editori Laterza, 2011.

- DAL BIANCO S., «Reassessing Cross-Regional Convergence in Italy Trough Distribution Dynamics», *Rivista Italiana degli Economisti*, vol. 16, 2011, pages 431-476.
- DE BONIS R. - FARABULLINI F. - NUZZO G., «Prestiti e raccolta delle banche: un'analisi degli andamenti territoriali», in DE BONIS R. - ROTONDI Z. - SAVONA P. (a cura di), *Sviluppo, rischio e conti con l'esterno delle regioni italiane*, Bari, Editori Laterza, 2010.
- DE BONIS R. - FARABULLINI F. - ROCCELLI M. - SALVIO A., «Nuove serie storiche sull'attività di banche e altre istituzioni finanziarie dal 1861 al 2011: che cosa ci dicono?», Banca d'Italia, *Quaderni di Storia Economica*, n. 26, 2012.
- DEL MONTE A., «La nuova politica per il Mezzogiorno: dalla centralizzazione allo sviluppo locale incentivato», *Economia e Politica Industriale*, n. 25, 1998, pp. 45-71.
- DEL MONTE A. - GIANNOLA A., *Istituzioni economiche e Mezzogiorno: Analisi delle politiche di sviluppo*, Roma, Carocci editore, 1997.
- DELLA TORRE G., «Strutture finanziarie e crescita economica in Italia (1861-1981)», in CONTI G. - LA FRANCESCA S. (a cura di), *Banche e reti di banche nell'Italia post-unitaria*, Bologna, il Mulino, 2000.
- DELLA TORRE G. - COCCIA M. - DE LEONARDIS V. - SCHISANI M.C., «La crescita del sistema finanziario italiano dopo l'Unificazione politica, 1861-1914: *financial deepening* e/o errori statistici e metodologici?», *Rivista di Storia Economica*, n. 2, 2008, pp. 135-174.
- DELL'OREFICE A., «La politica industriale del fascismo», in FAUSTO D. (a cura di), *Intervento pubblico e politica economica fascista*, Milano, Franco Angeli, 2007.
- DI GIACINTO V. - ESPOSITO L., «Convergence of Financial Structures in Europe: An Application of Factorial Matrices Analysis», in DE BONIS R. - POZZOLO A.F. (eds.), *The Financial Systems of Industrial Countries. Evidence from Financial Accounts*, Heidelberg, Springer, 2011.
- FEDERCASSE, *L'evoluzione della struttura distributiva delle banche italiane dal 2001 al 2006*, Roma, 2007.
- FERRI G. - MESSORI M., «Bank-Firm Relationships and Allocative Efficiency in North-eastern and Central Italy and in the South», *Journal of Banking and Finance*, vol. 24, 2000, pages 1067-1095.
- FIASCHI D. - GIANMOENA L. - PARENTI A., «The Dynamics of Labour Productivity Across Italian Provinces: Convergence and Polarization», *Rivista Italiana degli Economisti*, vol. 16, 2011, pages 241-280.
- FRANZINI M. - GIUNTA A., «Ripensare le politiche per il Mezzogiorno», *Meridiana - Rivista di Storia e Scienze Sociali*, n. 61, 2008, pp. 189-222.
- GALLI G., *Il Sistema Finanziario nel Mezzogiorno*, numero speciale dei Contributi all'analisi economica, Roma, Banca d'Italia, 1990.
- GIANNOLA A., «Le imprese e lo sviluppo: problemi e prospettive nel Mezzogiorno», *Rassegna Economica*, vol. LXII, n. 1, 1998, pp. 11-48.
- .-, *Il credito difficile*, L'Ancora del Mediterraneo, Napoli, 2002.

- GIANNOLA A., «Quando le banche si incontrano. Dall'interbancario al mercato interno dei capitali», *Rivista Economica del Mezzogiorno*, vol. XXIII, n. 1-2, 2009, pp. 9-26.
- GIANNOLA A. - IMBRIANI C. - LOPES A., «Le politiche di lungo termine per il superamento del dualismo nord-sud. Problemi e prospettive», in SCAZZIERI R. - QUADRIO-CURZIO A. (a cura di), *Dinamica economica strutturale*, Bologna, il Mulino, 1990.
- GIANNOLA A. - LOPES A., «Politiche di intervento, sviluppo economico del mezzogiorno e debito pubblico», in ENTE EINAUDI (a cura di), *Il disavanzo pubblico in Italia: natura strutturale e politiche di rientro*, Bologna, il Mulino, 1992.
- , -, -, «Vigilanza, efficienza, mercato; sviluppo e squilibri del sistema creditizio italiano», *Rivista Italiana degli Economisti*, vol. 1, 1996, pp. 25-54.
- , -, -, «Banca, sistema produttivo e dualismo in Italia; continuità e mutamenti strutturali. Una riflessione di lungo periodo», *Rivista Economica del Mezzogiorno*, vol. XXV, n. 3, 2011, pp. 517-560.
- GIANNOLA A. - PETRAGLIA C., «Politiche dell'offerta, della domanda e programmazione dello sviluppo. Il dualismo "dimenticato"», *Rivista Economica del Mezzogiorno*, vol. XXI, n. 1, 2007, pp. 13-42.
- GIANNOLA A. - PETRAGLIA C. - SCALERA D., *Regional Fiscal Flows and Regional Policy in Italy: A Long-Run Perspective*, mimeo, 2012.
- GIANNOLA A. - SARNO D., «Strutture finanziarie, vincoli di liquidità, performance e crescita delle imprese. Un'analisi territoriale», in BAGELLA M. (a cura di), *Finanza e crescita. Quali vincoli, quali rischi?*, Bologna, il Mulino, 2004.
- GIORDANO L. - LOPES A., «Measuring Efficiency of the Banking System in a Dualistic Economy. Evidence from the Italian Case», in SILIPO D. (ed.), *The Banking System in the Italian Economy*, Heidelberg, Springer Verlag, 2009a.
- , -, -, «Reti bancarie, credito e sistema produttivo meridionale», *Rivista Economica del Mezzogiorno*, vol. XXIII, n. 4, 2009b, pp. 827-868.
- GOBBI G. - SETTE E., *Relationship Lending in a Financial Turmoil*, mimeo, 2012.
- GRAZIANI A., *Lo sviluppo dell'economia italiana*, Torino, Bollati Boringhieri, 1998.
- GUIO L. - SAPIENZA P. - ZINGALES L. «The Role of Social Capital in Financial Development», *American Economic Review*, no. 94(3), 2004, pages 526-556.
- ISTAT, *Il valore della moneta in Italia dal 1861 al 2009*, Roma, 2009.
- LANZARONE G., *Il sistema bancario italiano*, Torino, Einaudi, 1946.
- LEVINE R., «Finance and Growth: Theory and Evidence», in AGHION S. - DURLAUF P. (eds.), *Handbook of Economic Growth*, Amsterdam, Elsevier Science, 2005.
- LUCCHETTI R. - PAPI L. - ZAZZARO A., «Banks' Inefficiency and Economic Growth: A Micro-macro Approach», *Scottish Journal of Political Economy*, vol. 48, 2001, pages 400-424.

- MALANIMA P., «Lo sviluppo ineguale in Italia 1861-2010», in SERIO M. (a cura di), *La prospettiva del meridionalismo liberale. Politica, istituzioni, economia, storia*, Soveria Mannelli, Rubbettino, 2012.
- MARANI U., *Finanziamenti e investimenti industriali in Italia (1966 - 1976)*, Torino, Boringhieri, 1980.
- MATTESINI F. - MESSORI M., *L'Evoluzione del sistema bancario meridionale: Problemi aperti e possibili soluzioni*, Bologna, il Mulino, 2004.
- MESSORI M. - TAMBORINI R. - ZAZZARO A., (a cura di), *Il sistema bancario italiano. Le occasioni degli anni Novanta e le sfide dell'euro*, Roma, Carocci editore, 2003.
- NUZZO G. - OLIVERIO E., «Banche locali e banche del Centro Nord nel Mezzogiorno: evoluzione temporale e confronto con altre aree in ritardo di sviluppo», Banca d'Italia, *L'integrazione economica tra il Mezzogiorno e il Centro Nord*, dicembre, 2011.
- ONADO M., «La lunga rincorsa: la costruzione del sistema finanziario», in CIOCCA P. - TONIOLO G. (a cura di), *Storia economica d'Italia*, vol. 3.2, *I vincoli e le opportunità*, Roma-Bari, Editori Laterza, 2003.
- PACI R. - PIGLIARU F., «Structural Change and Convergence: An Italian Regional Perspective», *Structural Change and Economic Dynamics*, vol. 8, 1997, pages 297-318.
- PANETTA F. - SIGNORETTI F.M., «Domanda e offerta di credito in Italia durante la crisi finanziaria», Banca d'Italia, *Questioni di Economia e Finanza*, n. 63, 2010.
- PILUSO G., «Gli istituti di credito speciale», in AMATORI F. (a cura di), *Storia d'Italia, Annali, 15, L'industria*, Torino, Einaudi, 1999.
- POLSI A., *Alle origini del capitalismo italiano. Stato, banche e banchieri dopo l'Unità*, Torino, Einaudi, 1993.
- PRESBITERO A. - UDELL F.G.F. - ZAZZARO A., «The Home Bias and the Credit Crunch: A Regional Perspective», *MoFiR, Working Paper*, no. 60, March, 2012, in corso di pubblicazione in *Journal of Money, Credit and Banking*.
- RODANO G., *Il credito all'economia. Raffaele Mattioli alla Banca Commerciale Italiana*, Milano-Napoli, Ricciardi, 1983.
- ROSSI N., *Mediterraneo del Nord. Un'altra idea del Mezzogiorno*, Roma-Bari, Laterza, 2005.
- SCALERA D. - ZAZZARO A., «Group Reputation and Persistent (or Permanent) Discrimination in Credit Markets», *Journal of Multinational Financial Management*, vol. 11, 2001, pages 483-496.
- .-, -.-, «L'economia del Mezzogiorno, la nuova politica regionale, crisi globale e federalismo fiscale», in GUERRA M.C., ZANARDI A. (cura di), *La finanza pubblica italiana*, Bologna, il Mulino, 2010.
- SVIMEZ, *Rapporto sull'economia del Mezzogiorno 1979*, Roma, Svimez, 1980.
- .-, *Rapporto 2006 sull'economia del Mezzogiorno*, Bologna, il Mulino, 2006.
- .-, *Rapporto 2009 sull'economia del Mezzogiorno*, Bologna, il Mulino, 2009.

- SVIMEZ, *Rapporto 2010 sull'economia del Mezzogiorno*, Bologna, il Mulino, 2010.
- .-, *150 anni di statistiche: nord e sud (1861-2011)*, Bologna, il Mulino, 2011a.
- .-, *Rapporto 2011 sull'economia del Mezzogiorno*, Bologna, il Mulino, 2011b.
- TROIANI M. - PASQUINI C. - RIGON A. - LANDI A., «I servizi finanziari alle piccole e medie imprese nell'esperienza delle banche europee», *Bancaria*, n. 11, 1999, pp. 64-79.
- VAONA A., «Regional Evidence on Financial Development, Finance Term Structure and Growth», *Empirical Economics*, Springer, vol. 34(1), 2008, pages 185-201.
- VAONA A. - PATUELLI R., «New Empirical Evidence on Local Financial Development and Growth», *Letters in Spatial and Resource Sciences*, vol. 1(2-3), 2008, pages 147-157.
- VIESTI G., *Mezzogiorno a tradimento. Il Nord, il Sud e la politica che non c'è*, Roma-Bari, Editori Laterza, 2009.
- ZAGO A. - DONGILI P., «Credit Quality and Technical Efficiency in Banking», *Empirical Economics*, Springer, vol. 40(2), 2011, pages 537-558.
- ZAZZARO A., «Nuova programmazione e sviluppo del Mezzogiorno. C'è ancora spazio per l'impresa pubblica?», in CARILLO M.R., ZAZZARO A. (a cura di), *Istituzioni capitale umano e sviluppo del Mezzogiorno*, Napoli, ESI, 2001.
- .-, «La scomparsa dei centri decisionali dal sistema bancario meridionale», *Rivista di Politica Economica*, vol. 96(2), 2006, pp. 31-60.

Un commento alla riforma Monti - Fornero

Marco Leonardi* Massimo Pallini**
Università di Milano

This paper gives an overview of the recent labor market reform in Italy. We focus on the trade-off underlying the spirit of the reform: to increase the use of open-ended contracts the reform reduced the opportunity of (mis)using of temporary contracts and concurrently reduced the costs of individual firings. Furthermore the unemployment subsidy system and the temporary short-time work scheme (Cassa Integrazione) was vastly overhauled. We provide the details of the debate that followed some of the most controversial points of the reform and show some preliminary evidence of the potential problems facing its implementation.

[JEL Classification: J08; J38; J42; J58].

Keywords: labour market reforms; temporary contracts; individual dismissal regulations; welfare benefits.

* <marco.leonardi@unimi.it>; Dipartimento di Economia, Management e Metodi Quantitativi (DEMM).

** <massimo.pallini@unimi.it>; Dipartimento di Studi Sociali e Politici.

1. - Introduzione

Una valutazione complessiva della riforma Monti-Fornero sarà possibile solo quando le nuove disposizioni saranno a regime e i dati raccolti dal sistema permanente di monitoraggio e valutazione istituito presso il Ministero del lavoro e delle politiche sociali saranno resi noti. La prima novità della riforma è proprio l'istituzione del sistema di monitoraggio i cui dati potranno potenzialmente aprire una via promettente per la ricerca in economia del lavoro e per la valutazione delle politiche pubbliche. Si tratta di una novità positiva in quanto di solito le leggi italiane, contrariamente a quelle di molti paesi soprattutto di cultura anglosassone, non istituiscono dei sistemi di valutazione di sé stesse.

Al momento, in assenza di dati certi, commentiamo i principi ispiratori della riforma e cerchiamo di mettere in luce i vari aspetti del dibattito che si è snodato attorno ai quattro cardini o "pilastri" principali della stessa. Per ciascuno dei pilastri forniamo anche dell'evidenza basata su dati preliminari che non vuole essere una valutazione ma soltanto un commento informato dalla presenza di un riscontro empirico. Spesso faremo riferimento ai dati della regione Veneto, una delle poche regioni che ad oggi produce delle elaborazioni puntuali ed accurate delle fonti amministrative in suo possesso. In questo modo intendiamo mettere in luce anche i potenziali aspetti da approfondire una volta che i dati saranno a disposizione.

Insolitamente per una legge dello Stato, e probabilmente ai fini di influenzare la giurisprudenza, la riforma Monti - Fornero esplicita le linee direttrici che la ispirano:

- (1) ribadire e valorizzare il contratto di lavoro a tempo indeterminato come "contratto dominante" ovvero forma comune del rapporto di lavoro;
- (2) ridistribuire più equamente le tutele dell'impiego, riconducendo nell'alveo di usi propri i margini di flessibilità progressivamente introdotti negli ultimi vent'anni;
- (3) adeguare la disciplina del licenziamento individuale per alcuni specifici motivi oggettivi alle esigenze dettate dal mutato contesto di riferimento;
- (4) rendere più efficiente, coerente ed equo l'assetto degli ammortizzatori sociali e delle politiche attive a contorno.

I pilastri della riforma sono da considerarsi in solido e non singolarmente: allo sforzo di riduzione degli abusi nell'utilizzo del lavoro autonomo fa da contro-altare una potenziale facilitazione del licenziamento individuale in modo che le imprese possano far ricorso al contratto a tempo indeterminato invece che alla partita Iva. La contestuale variazione della disciplina del licenziamento (di cui al punto 3) e delle regole di assunzione con contratto a termine o a partita Iva (di cui ai

punti 1 e 2) è la chiave di volta per comprendere il senso di questa riforma. E per comprendere come la radicalità delle misure inizialmente proposte sia stata progressivamente ridotta in misura proporzionale su entrambi i versanti delle assunzioni attraverso partita Iva e sul licenziamento di tipo individuale.

Da una parte, la proposta originale di riforma aveva previsto in un primo momento l'esclusione della possibilità di reintegro nel caso di licenziamento individuale per motivi economici e poi invece la sua reintroduzione ma solo per alcuni specifici motivi oggettivi "manifestamente insussistenti". Dall'altra parte, le misure inizialmente molto severe che avrebbero dovuto impedire l'assunzione attraverso partita Iva fasulla invece che a contratto a tempo indeterminato, sono state ammorbidite attraverso il decreto del 17 agosto 2012. Le resistenze politiche al cambiamento sono state molto intense e di segno opposto (i sindacati per quanto riguarda l'articolo 18, i datori di lavoro per quanto riguarda l'utilizzo dei contratti a termine e le partite Iva). Nonostante tutto noi riteniamo che la riforma abbia affermato delle direzioni di cambiamento importanti i cui effetti saranno sicuramente visibili (e valutabili) nei prossimi anni.

In questo articolo analizziamo il complesso della riforma utilizzando come schema espositivo i quattro pilastri di cui sopra. Riteniamo che l'unica direttrice mancante di questa riforma, su cui si potrà ancora porre rimedio in un futuro prossimo, sia la necessità di ridurre il contenzioso giudiziario in caso di licenziamento.

2. - Tipologie contrattuali

È noto come il principale problema del mercato del lavoro italiano sia la precarietà dei contratti di lavoro soprattutto per le nuove generazioni. Meno noto è il fatto che l'anomalia dell'Italia nel confronto europeo non riguarda i contratti a termine di tipo dipendente, gli apprendisti o il lavoro interinale quanto piuttosto l'utilizzo massiccio di forme di lavoro autonomo che nascondono una reale dipendenza economica. La comprensione di quanti fra coloro che rientrano negli autonomi sono in realtà "dipendenti economicamente" da un'unica impresa non è un compito facile e dipende da come si intende definire un lavoratore "dipendente economicamente"¹. Tuttavia l'inclusione di una stima ragionevole del nu-

¹ In un'indagine condotta dall'Isfol è stato chiesto ai lavoratori autonomi se essi hanno un unico datore di lavoro, se hanno una sede esclusiva di lavoro e se il regime orario è fissato in modo ./.

mero di questi lavoratori (poco più di un milione) nel gruppo dei dipendenti a termine aumenterebbe in modo significativo l'incidenza del lavoro temporaneo sul totale dell'occupazione portando l'Italia ben al di sopra degli altri Paesi europei, portando la quota di contratti a termine dal 13% al 19% della popolazione occupata tra i 15 e i 64 anni e dal 50% al 70% della popolazione occupata tra i 15 e i 24 anni (Cappellari *et al.*, 2012*b*).

La riforma indica specificatamente tra i propri obiettivi quello di introdurre nuovi criteri di qualificazione del lavoro subordinato così da poter definire più agevolmente l'ambito di applicazione della relativa disciplina legale e contrastare l'abuso del ricorso da parte delle imprese a contratti di lavoro autonomo in modo continuativo e inserito organicamente nei processi di produzione al solo scopo di sottrarsi all'applicazione delle tutele legali del lavoro subordinato. Questo fenomeno di "fuga" dal lavoro subordinato è cresciuto esponenzialmente dagli anni ottanta ed ha ora raggiunto una dimensione nel mercato del lavoro difficilmente governabile, eleggendo i contratti di lavoro autonomo a modalità prevalente di accesso dei giovani al mondo del lavoro godendo, di conseguenza, tutele e garanzie sideralmente lontane da quelle applicabili al lavoro subordinato "standard".

Per porre rimedio a questo problema la riforma si propone di incentivare l'utilizzo da parte delle imprese del lavoro subordinato per un verso rendendo più oneroso e problematico per le imprese ricorrere a contratti di lavoro autonomo per assolvere attività che strutturalmente e stabilmente sono parti del loro ciclo di produzione, per altro verso rendendo marginalmente più flessibile l'assunzione e il licenziamento di un dipendente con un contratto a tempo indeterminato. In fondo questa era la strategia che aveva originariamente ispirato anche il libro bianco di Marco Biagi e che aveva poi trovato una attuazione inadeguata dopo la sua morte nel D.lgs. n. 276 del 2003, comunemente denominato "Legge Biagi" (Leonardi e Pallini, 2006)².

unilaterale dal proprio principale. Coloro che rispondono positivamente alle tre domande sopra citate sono solamente 100.000, ma si passa a circa due milioni di individui se consideriamo chi risponde affermativamente ad almeno una delle tre domande. In considerazione di ciò, la stima migliore del gruppo degli autonomi cosiddetti "dipendenti economicamente da un'impresa" potrebbe aggirarsi sul milione di unità, in maggioranza coloro che lavorano in maniera esclusiva presso la sede del committente (ANASTASIA B., 2010*a*).

² Parzialmente diverso è il principio ispiratore dei due diversi progetti di "contratto unico" a firma dei professori ICHINO P. (2009) e BOERI T. e GARIBALDI P. (2010) che, ritenendo molto difficile dal punto di vista politico la variazione delle regole di licenziamento per i contratti in essere, si concentrano invece sui contratti per i soli nuovi assunti: questi ultimi sarebbero soggetti ad una diversa disciplina del licenziamento per un periodo di tempo di tre anni o indefinito a seconda dei progetti.

Per quanto riguarda le assunzioni la riforma del ministro Fornero mira ad elevare il contratto di apprendistato, che è una tipologia di contratto di lavoro subordinato, a modalità “prevalente” per l’ingresso nel mercato del lavoro. Si prevede infatti l’abrogazione del contratto di inserimento. L’apprendistato è stato già oggetto di una legge di riforma significativa allo scadere del Governo Berlusconi, la legge n. 167 del 2011. Questa legge qualifica il contratto di apprendistato non più come contratto a termine, ma come contratto a tempo indeterminato; questa qualificazione ha una forte valenza simbolica e prospettica, sebbene il datore di lavoro abbia la facoltà di recedere liberamente dal rapporto allo scadere del periodo di formazione dell’apprendista. Quest’ultimo gode di tutte le tutele del lavoratore subordinato ma il suo costo per l’impresa è significativamente più ridotto perché gli si può applicare il trattamento retributivo previsto dai contratti collettivi per professionalità inferiori inquadrate sino a 2 livelli più bassi ed è soggetto ad un onere di contribuzione previdenziale più contenuto per l’impresa³.

La riforma interviene a modifica della legge n. 167 del 2011 cercando di limitare le possibilità di abusi di questo contratto da parte delle imprese. A tal fine si prevede:

- (a) una durata minima del contratto di apprendistato, fissata in almeno sei mesi, fatta salva la possibilità di durate inferiori per attività stagionali e le eccezioni previste nel D.lgs. n. 167 del 2011;
- (b) l’introduzione di un meccanismo in base al quale l’assunzione di nuovi apprendisti è collegata alla percentuale di stabilizzazioni effettuate nell’ultimo triennio (50%) con l’esclusione dal computo della citata percentuale dei rapporti cessati durante il periodo di prova, per dimissioni o per licenziamento per giusta causa.

La riforma interviene inoltre sulla disciplina del contratto a termine sotto due profili solo in apparenza in contraddizione. Per un verso la riforma consente di utilizzare liberamente il contratto di lavoro come un lungo periodo di prova del lavoratore per un anno. Prevede, infatti, che il primo contratto a termine stipulato

³ L’apprendistato è un contratto speciale a causa mista (formazione e lavoro) utilizzato non solo per i giovani che abbandonano gli studi ma anche per l’inserimento al lavoro operaio e impiegatizio di lavoratori fino ai 34 anni. Gli sgravi contributivi sono costosi per la finanza pubblica (2,3 miliardi di euro nel 2010) e sono accompagnati da altri incentivi normativi, contrattuali e fiscali. Nel periodo di apprendistato l’accreditamento dei contributi per il lavoratore (contribuzione del 5,84% del salario) è pieno e se il rapporto di lavoro viene stabilizzato il diritto agli sgravi prosegue per ulteriori 12 mesi.

tra un lavoratore ed un'azienda (anche con contratto di somministrazione) per una durata massima di un anno, non prorogabile, non deve essere giustificato dalla sussistenza di requisiti tecnici, produttivi, organizzativi o sostitutivi, la carenza dei quali, come noto, comporta la trasformazione dei contratti a termine in contratti a tempo indeterminato.

Per altro verso la riforma, in sintonia con gli obiettivi della direttiva europea 99/70, cerca di impedire la reiterazione di contratti a termine di un lavoratore con la stessa azienda per esser in realtà assegnato ad esigenze organizzative o produttive permanenti di quest'ultima. A tal fine si allunga in misura rilevante l'intervallo che deve necessariamente intercorrere tra due contratti a termine stipulati con il medesimo lavoratore: l'intervallo minimo da osservare dopo un contratto a termine di durata inferiore a 6 mesi è stato elevato da 10 a 60 giorni, mentre dopo un contratto a termine di durata superiore l'intervallo obbligatorio è stato elevato da 20 a 90 giorni. Inoltre la riforma tenta di rendere meno attrattivo l'utilizzo continuativo di contratti a termine incrementandone il costo con l'imposizione di un contributivo previdenziale aggiuntivo (pari all'1,4%). L'impresa può però recuperare tale maggiorazione contributiva se il rapporto di lavoro viene poi trasformato in uno a tempo indeterminato.

Come detto, dunque, questi interventi mirano a rendere da un lato più costoso l'uso del contratto a tempo determinato e più problematica la sua reiterazione con lo stesso lavoratore, d'altro lato più facile e conveniente l'uso del contratto di apprendistato. Per questi aspetti la proposta di riforma appare condivisibile.

Al riguardo Cappellari *et al.* (2012a) hanno mostrato come le precedenti riforme del contratto a tempo determinato abbiano generato incertezza negli imprenditori e compromesso almeno negli anni iniziali dopo la riforma Biagi del 2003 l'utilizzo efficiente della nuova tipologia di contratto a tempo determinato. La riforma Biagi del 2003 applicata nei contratti collettivi nel corso del 2005-2006 prevedeva l'abolizione della causali specifiche e la sostituzione con un unico "causalone" per regioni tecnico-organizzative. Il timore degli imprenditori che la genericità della norma potesse portare a maggior contenzioso giudiziale e infine alla trasformazione coatta del contratto in un tempo indeterminato, ne ha rallentato l'utilizzo a favore dei contratti di co.co.co. o di agenzie interinali. La stessa ricerca mostra che nello stesso periodo 2005-2006 le imprese delle regioni che adottavano i regolamenti del nuovo contratto di apprendistato (che estendeva a 30 anni l'età ammissibile e consentiva la formazione interna all'impresa) sostituivano i co.co.co. con gli apprendisti migliorando la produttività. Si mette così in luce un punto essenziale per il funzionamento del contratto di apprendistato che

riguarda appunto le regole applicative regionali e dei contratti collettivi nazionali e che quindi hanno poco a che fare con la responsabilità diretta del Ministero del Lavoro.

Il dibattito sulla parte della riforma che riguarda l'apprendistato si è concentrato sull'adeguatezza di questo tipo di contratto per l'inserimento dei giovani nel mercato del lavoro. La riforma ha preferito privilegiare un tipo di contratto già esistente, l'apprendistato appunto, invece di aggiungere un nuovo tipo di contratto come avrebbero voluto i proponenti del contratto unico di inserimento (Boeri e Garibaldi, 2010 o Ichino, 2009). Questa scelta ha ovviamente dei "pro" e dei "contro" che potranno essere valutati solo con i dati effettivi delle assunzioni.

Una delle critiche più frequenti risiede nelle difficoltà di diffusione che ha avuto il contratto di apprendistato già in passato. L'utilizzo di questo tipo di contratto ha un *trend* decrescente da anni: ad esempio il numero di contratti di apprendistato nella regione Veneto, una delle regioni con la maggior diffusione di manifattura mostra un *trend* decrescente di circa il 5% all'anno, già negli anni precedenti alla riforma (Veneto Lavoro, 2012). Un'altra critica frequente riguarda la difficoltà di stipulare contratti di apprendistato per i laureati proprio mentre la disoccupazione tra i laureati è maggiore. Quest'ultima critica non pare particolarmente convincente. È sicuramente vero che la disoccupazione relativa dei laureati rispetto ai diplomati è molto aumentata dopo la crisi – complice anche una riforma dell'università che ha raddoppiato in pochi anni il numero dei laureati (Bosio e Leonardi, 2010) – ma è anche vero che negli ultimi anni sono aumentati i posti di lavoro poco qualificati che sarebbero i più adatti per il contratto di apprendistato. I dati Istat 2012 mostrano un aumento dei posti di lavoro poco qualificato di 220mila unità dal 2008 al 2012, unici posti di lavoro in crescita evidente nel periodo. Infine il contratto di apprendistato oggi può essere stipulato anche ai laureati, quindi le ragioni per il scarso uso di questo contratto devono probabilmente essere ricercate altrove.

In conclusione i pro dell'apprendistato sono da vedersi nella scelta di puntare su un contratto già presente nell'ordinamento italiano e di tutti i paesi europei, che è stato allargato (dal Testo Unico precedente alla riforma) anche oltre i limiti di età per il reinserimento dei lavoratori messi in mobilità. I contro risiedono potenzialmente nella difficoltà, già sperimentata più volte in passato, nell'uso del contratto di apprendistato da parte delle imprese. Per diventare il canale di accesso privilegiato dei giovani nel mercato del lavoro è necessario che venga utilizzato in modo massiccio, superando le difficoltà di comprovare l'avvenuta formazione e le lentezze delle leggi regionali di applicazione. I sostenitori del contratto unico

lamentano che il contratto unico avrebbe avuto rispetto all'apprendistato un ambito di applicazione più ampio in termini di età e un costo di licenziamento crescente nel tempo in modo da costituire un incentivo all'assunzione di giovani. L'incentivo all'uso del contratto di apprendistato è invece dato dal ridotto costo dei contributi pensionistici e sociali, manca tuttavia allo stato delle cose quel sistema integrato di passaggio dalle scuole superiori e dalle università tecniche all'esperienza pratica in azienda che costituisce il successo dell'apprendistato nei paesi di cultura tedesca (Cappellari e Leonardi, 2012).

3. - Redistribuire più equamente le tutele dell'impiego

La riforma interviene in modo efficace sulla disciplina del lavoro a progetto così da inibire che possano essere utilizzati in modo continuativo contratti di lavoro autonomo per assolvere mansioni che soddisfano esigenze organizzative e produttive permanenti dell'impresa (anche qualora i lavoratori non siano soggetti a incisivi obblighi di subordinazione nell'esecuzione delle loro prestazioni). La riforma specifica che il lavoro autonomo può essere utilizzato in modo coordinato con l'organizzazione d'impresa esclusivamente se è strumentale al compimento di un progetto (e non più anche di un mero programma o fase) che realizzi un risultato finale separato e distinguibile dall'obiettivo complessivo dell'attività aziendale. I caratteri del progetto e il risultato atteso dallo stesso debbono essere specificatamente indicati in forma scritta nel contratto quale condizione di legittimità dello stesso. La mancata individuazione del progetto o la non rispondenza di questo ai caratteri stringenti dettati dalla legge comporta di diritto la trasformazione in un rapporto di lavoro subordinato alle dipendenze del committente. Inoltre per agevolare i lavoratori a progetto nel soddisfare l'onere della prova di esser stati in concreto utilizzati nelle forme della subordinazione è dettata una presunzione (che ammette prova contraria a carico del committente) secondo cui i collaboratori a progetto che provino di prestare la loro attività con modalità analoghe a quelle dei dipendenti del committente debbono anch'essi essere qualificati come lavoratori subordinati alle dipendenze di quest'ultimo. Infine la riforma – a differenza della disciplina precedente – garantisce sia che il loro rapporto di lavoro a progetto non possa essere interrotto dal committente prima della scadenza, salvo che non vi sia una giusta causa o una loro comprovata inidoneità professionale, sia che questi collaboratori abbiano diritto ad un compenso minimo sufficiente e rispettoso dei minimi dettati dai contratti collettivi per i lavoratori a

progetto o, in assenza di specifiche previsioni in tal senso, di quelli previsti per i lavoratori subordinati.

Nel disegno di riforma, ben cogliendo che i contratti di lavoro autonomo sono ormai percepiti dalle imprese come perfettamente fungibili e alla compressione da parte del legislatore delle possibilità di ricorrere ad uno automaticamente le imprese incrementano il ricorso agli altri, si limita anche la possibilità di ricorrere a prestazioni di lavoro autonomo con partita Iva. Si introduce una presunzione di diritto secondo cui tali rapporti di collaborazione coordinata e continuativa qualora ricorrano almeno due dei seguenti presupposti:

- a) che la durata della collaborazione superiore a otto mesi nell'arco di un anno per due anni consecutivi;
- b) che il ricavo dei corrispettivi percepiti dal collaboratore nell'arco dello stesso anno solare superi la misura dell'80 % dei corrispettivi annui complessivamente percepiti dal collaboratore nell'arco di 2 anni solari consecutivi;
- c) che il prestatore abbia la disponibilità di una postazione di lavoro presso il committente. Tale presunzione vale fino a prova contraria da parte del committente, il quale può dimostrare la genuinità del rapporto di lavoro autonomo. La configurazione del rapporto come collaborazione coordinata e continuativa implica l'applicazione di tutte le norme che disciplinano tale contratto, incluse quelle relative al regime previdenziale e all'eventuale trasformazione della collaborazione in un rapporto di lavoro subordinato a tempo indeterminato qualora sia stata instaurata senza l'individuazione di uno specifico progetto.

La riforma intende annullare il vantaggio di contenimento del costo del lavoro con l'utilizzare il lavoro autonomo. Pertanto, da un lato, come già sottolineato, impone l'erogazione di un compenso sufficiente al collaboratore; d'altro lato si prevede un progressivo incremento della contribuzione previdenziale alla Gestione separata dell'Inps dei lavoratori autonomi sino a raggiungere l'aliquota del 33% a decorrere dall'anno 2018. Certamente la parificazione delle aliquote contributive non rende di per sé in misura significativa meno attraente il lavoro autonomo rispetto al lavoro subordinato considerato la maggiore flessibilità gestionale e in uscita che il primo è ancora in grado di offrire.

La riforma semplifica le modalità di ricorso al lavoro c.d. accessorio (con pagamento mediante *voucher*), consentendolo per ogni tipologia di attività entro il limite dei 5.000 euro annuali complessivi ma escludendolo in favore di committenti imprenditori commerciali e professionisti. Anche il ricorso alla associazione in partecipazione è ridotto al massimo a tre associati per ogni azienda e con la

prova che l'associato partecipi agli utili e ne possa controllare la loro determinazione. Vengono infine eliminate le modalità di lavoro intermittente con cui il lavoratore deve garantire la sua disponibilità senza ricevere alcun compenso se non viene chiamato.

Alla fine i meccanismi di qualificazione del rapporto di lavoro hanno una portata sistemica: non solo disegnano una nuova nozione del lavoro a progetto, ma – specularmente – incidono anche sulla stessa nozione di lavoro subordinato o meglio sull'ambito di applicazione della disciplina legale dettata per quest'ultimo contratto di lavoro. Nell'ambito di applicazione del regime legale del lavoro subordinato devono essere ricondotti sia il lavoro eterodiretto in senso proprio, sia il lavoro autonomo prestatato in modo continuativo e coordinato con l'organizzazione aziendale. Questo nuovo e più ampio ambito di applicazione del regime legale del lavoro subordinato risulta coerente con la nozione di subordinazione elaborata dalla Corte di giustizia europea secondo cui il lavoratore autonomo deve atteggiarsi come una "impresa individuale" capace di vendere a terzi un bene o un servizio.

Al momento non è possibile quantificare gli effetti della riforma sull'utilizzo di co.co.co e partite Iva, ci limitiamo a notare che l'Italia costituisce un'eccezione in sede europea in quanto circa il 25% per cento dell'occupazione totale è costituita da lavoro autonomo o parasubordinato a fronte di un 15% degli altri paesi europei. Di circa i 5 milioni di partite Iva e co.co.co italiani un quarto lavora in mono-committenza. Nonostante le difficoltà a stabilire i criteri che determinano la "finta" partita Iva e quindi a quantificarne esattamente il numero (come abbiamo visto le stime si aggirano su cifre superiori al milione di individui), tuttavia sembra evidente che c'è un ampio margine in cui la riforma può potenzialmente incidere (Ranci, 2012).

A tal proposito tuttavia una critica frequente è che la stretta sulle partite Iva rischia di aumentare la disoccupazione. I difensori della riforma sostengono invece che molti di quei contratti sarebbero convertiti in contratti a tempo indeterminato. Una prima idea di quel che è successo in Veneto si ha guardando non alle partite Iva ma ai contratti a chiamata, anch'essi resi più difficile dalla riforma. Anastasia (2012) sostiene che in Veneto il numero di contratti a chiamata ha subito un crollo di 30.000 unità circa dal luglio 2012, circa 10.000 di questi contratti sarebbero stati rinnovati dalle medesime imprese sotto forma di contratti a tempo determinato oppure di contratti a tempo indeterminato *part-time*. Ovviamente in queste cifre si possono vedere entrambi gli effetti della riforma Fornero, ma tutte le interpretazioni sono necessariamente affrettate. Da un lato si può so-

stenere che la riforma distrugge posti di lavoro, dall'altro si può invece sostenere che la riforma fa esattamente quello che intendeva fare. Infatti le nuove norme hanno introdotto l'obbligo di comunicare la durata di ogni attivazione effettiva del lavoratore a chiamata, con l'esplicito intento di contrastare l'utilizzo distorto del contratto a chiamata come schermo di forme di lavoro nero di varia intensità. L'effetto sembra essere stato proprio quello di trasformare molti di questi contratti in contratti più stabili nella medesima impresa e contestualmente eliminare quei contratti che non potevano soddisfare i nuovi requisiti.

4. - La nuova disciplina del licenziamento individuale

Come già rilevato, a fronte delle stringenti limitazioni all'utilizzo di lavoro autonomo in modo continuativo in seno all'organizzazione d'impresa il disegno di legge cerca di realizzare una sorta di compensazione virtuosa inserendo margini di flessibilità in entrata e in uscita nel tipo contrattuale "standard" di lavoro subordinato a tempo indeterminato. Proprio le previsioni della riforma che investono la flessibilità in uscita costituiscono il punto più delicato che ha già registrato un cambiamento di rotta estremamente significativo: dalla originaria esclusione della possibilità della tutela reintegratoria avverso il licenziamento illegittimo alla reintroduzione della stessa seppur solo per casi specifici.

La riforma non interviene sulle ragioni giustificatrici del licenziamento, che rimangono disciplinate dall'art. 3 della legge n. 604 del 1966, né sul regime di tutela applicabile nelle piccole aziende con meno di 15 lavoratori. L'intervento è limitato alle forme di tutela e al procedimento stragiudiziale e giudiziale per farle valere nelle aziende di medio-grandi dimensioni.

La reintegrazione, con il connesso diritto a tutte le retribuzioni non percepite dal momento del licenziamento sino alla effettiva reintegrazione, viene mantenuta nei casi di licenziamenti cc.dd. discriminatori, adottati cioè per ragioni tassative che la legge qualifica espressamente quali causa di nullità degli atti datoriali (per discriminazione per genere, età, opinioni politiche, appartenenza sindacale, disabilità, etc.).

Se il licenziamento è adottato per motivi disciplinari adducendo condotte infedeli o inadempienti del lavoratore, la tutela reintegratoria trova applicazione soltanto se la condotta contestata al lavoratore risulti insussistente o sia sanzionata secondo le previsioni del contratto collettivo o del codice disciplinare applicabile con una sanzione "conservativa", che cioè non prevede l'estinzione del rapporto

di lavoro. In tal caso, però, il lavoratore reintegrato avrà diritto anche ad un risarcimento del danno nella misura massima di 12 mensilità, e non pari a tutte le retribuzioni perse dal momento del licenziamento; la contribuzione previdenziale deve essere versata per tutto questo periodo dal datore di lavoro. Questa tutela trova applicazione anche avverso i licenziamenti adottati per giustificato motivo oggettivo (cc.dd. licenziamenti economici) quando il fatto e le circostanze che il datore di lavoro ha addotto a giustificazione del licenziamento risultano non vere e insussistenti oppure quando nei licenziamenti collettivi vengano violati i criteri di scelta dai lavoratori da licenziare⁴.

In effetti in tutte le altre ipotesi in cui il licenziamento per motivi disciplinari o il licenziamento per motivi economici sia ritenuto illegittimo la riforma prevede inequivocabilmente che il datore sia sanzionato soltanto con la condanna al pagamento di una indennità in favore del lavoratore da un minimo di 12 ad un massimo di 24 mensilità che determinerà equitativamente il giudice alla luce dell'anzianità del lavoratore, delle dimensioni aziendali, delle ragioni addotte per il licenziamento, della condotta delle parti.

I vizi formali o procedimentali, che non investono cioè la veridicità o fondatezza o proporzionalità delle circostanze addotte a giustificazione del licenziamento, comportano sia per i licenziamenti individuali sia per quelli collettivi l'obbligo per il datore di lavoro di corrispondere al lavoratore una indennità determinata equitativamente dal giudice da un minimo di 6 ad un massimo di 12 mensilità di retribuzione.

La riforma dunque ha inteso compensare le restrizioni poste all'utilizzo del lavoro autonomo importando maggior flessibilità in uscita nella disciplina del lavoro subordinato. La tutela reintegratoria non costituisce più l'unica sanzione che il giudice deve adottare in caso di illegittimità del licenziamento. Laddove il giudice ravvisi che il fatto addotto a giustificazione del licenziamento è effettivamente sussistente ma non perfezioni i requisiti di giusta causa o giustificato motivo oggettivo che legittimano il licenziamento potrà adottare una condanna economica da 12 a 24 mensilità. Si pensi, ad esempio, alla ipotesi in cui il giudice ravvisi una mancanza disciplinare ma rilevi un difetto di proporzionalità nel provvedimento

⁴ A riprova delle intenzioni del legislatore, Monti nella lettera in inglese al *Wall Street Journal* del 6 aprile 2012 ha parlato di reintegro solo nel caso di motivi economici *inexistent*, il che significa che anche se il giudice ritiene che il motivo economico sia *unjustified* dovrebbe ordinare il risarcimento e non il reintegro, il reintegro si avrebbe solo nei casi di motivi economici palesemente *inexistent*.

di licenziamento oppure alla ipotesi in cui la scelta organizzativa dell'impresa ha comportato la soppressione delle mansioni assegnate al lavoratore licenziato, ma il datore non soddisfi l'onere probatorio della non riallocabilità dello stesso lavoratore in altre mansioni.

Non si può non rilevare che l'ambiguità e complessità del testo del nuovo articolo 18 della legge n. 300 del 1970 introdotto dalla riforma lascia al giudice spazi di discrezionalità valutativa molto ampi, che rendono – per entrambe le parti – difficilmente prevedibile l'esito del giudizio non più soltanto con riguardo all'accoglimento o al rigetto della impugnazione del licenziamento, ma anche con riguardo al tipo di tutela – reintegratoria o obbligatoria – che verrà concessa nel primo caso e alla quantificazione dell'eventuale risarcimento economico concesso. Sino ad ora non giungono indicazioni chiarificatrici dalle prime pronunce giurisprudenziali, che invece denotano un certo disorientamento da parte dei giudici del lavoro chiamati ora a decidere secondo parametri valutativi assai diversi rispetto al passato. Non stupisce quindi che emergano orientamenti conservativi e di stretta continuità con il sistema precedente che finiscono di fatto per vanificare qualsiasi portata innovatrice della riforma. Ad esempio ha alimentato il dibattito la posizione espressa dal Tribunale di Bologna con la sentenza del 15 ottobre 2012 (Trib. Bologna, sez. lavoro, dott. Marchesini, proc. n. 2631 del 2012) che ha rinvenuto l'applicabilità della tutela reintegratoria non soltanto quando risulti inveritiero il fatto "materiale" addotto a giustificazione del licenziamento, ma anche quando non si realizzi il fatto "giuridico" idoneo a perfezionare la giusta causa o il giustificato motivo di licenziamento, quando cioè, pur nella veridicità delle condizioni fattuali, non ricorrano tutti i presupposti che sul piano giuridico legittimano il licenziamento. Secondo questa lettura ogni licenziamento che non sia supportato da una giusta causa o da un giustificato motivo si fonderebbe necessariamente su un fatto "giuridico" inesistente, comportando l'applicazione della sanzione della reintegrazione. In tal modo di fatto si finisce per annullare ogni portata della riforma perché non vi sarebbe alcuno spazio praticabile per applicare la sanzione economica avverso il licenziamento illegittimo.

Nonostante gli elevati margini di incertezza interpretativa che residuano a causa della scarsa qualità tecnico-giuridica della nuova norma, occorre registrare una indiscutibile efficacia di quest'ultima nel perseguire in concreto il fine di velocizzare celere la definizione delle controversie avverso i licenziamenti e di "monetizzare" la tutela offerta al lavoratore. La nuova norma ha infatti introdotto un rito processuale che può contare su un canale preferenziale per i tempi di trattazione, ma soprattutto sta dando buona prova sul campo la procedura obbligatoria

di conciliazione preventiva che deve essere attivata dalle aziende prima di adottare un licenziamento per motivi economici. Sebbene al momento ancor manchino dei dati certi e significativi, gli operatori processuali registrano tuttavia che l'estrema incertezza dell'esito del giudizio cui sono esposti sia il datore di lavoro sia il lavoratore licenziato incentiva il raggiungimento di accordi volontari in misura più consistente di quanto avvenisse nel regime precedente.

Nonostante questo *trend* incoraggiante, risulterebbe a nostro avviso più razionale ed efficace l'introdurre una indennità di licenziamento che si atteggi nei termini di un *severance pay*; una indennità cioè che il datore deve pagare in ogni caso al momento del licenziamento, così come funziona in molti paesi europei (tra cui ad es. Francia, Germania, UK). Potrebbe ipotizzarsi un sistema analogo a quello introdotto nell'ordinamento tedesco dalla riforma Hartz approvata dal governo socialdemocratico di Schroeder: prevedere che l'azienda laddove intenda adottare un licenziamento per giustificato motivo debba corrispondere al lavoratore una indennità in cifra variabile a seconda dell'anzianità di servizio; se il lavoratore accetta la corresponsione della indennità, decade automaticamente dal diritto di impugnare il licenziamento, altrimenti può rifiutare la corresponsione della indennità e adire il giudice del lavoro per cercare di ottenere la reintegrazione o un risarcimento più elevato laddove sia fiducioso di poter comprovare l'insussistenza della ragione posta a fondamento del licenziamento o la natura discriminatoria di questo.

Questo meccanismo per un verso responsabilizzerebbe l'impresa imponendole un costo certo in caso di licenziamento per giustificato motivo oggettivo, costringendola a verificare seriamente se il licenziamento del lavoratore compensa quel costo; per altro verso responsabilizzerebbe il lavoratore per cui non è più indifferente decidere di impugnare o meno il licenziamento e dunque valuterà attentamente se affrontare o meno il giudizio. Occorre rilevare che in Germania questo sistema non sta funzionando al meglio e il tasso di contenzioso giudiziale è rimasto elevato. Ciò però è addebitabile alla scelta di quantificare l'indennità ad un importo così (0,5 mensilità per anno di anzianità) che la perdita non costituisce per il lavoratore un sufficiente deterrente dal ricorrere in giudizio (Goerke e Pannenberg, 2010). Al contrario l'importo dell'indennità dovrebbe essere determinato dalla legge in un importo assai prossimo alla soglia massima che l'impresa sarebbe disposta a pagare per risolvere il rapporto così da assolvere al meglio sia la funzione di indurre le imprese a valutare bene gli effettivi vantaggi conseguenti al licenziamento del lavoratore sia specularmente quella di disincentivare il lavoratore ad andare in giudizio rischiando di perdere un importo elevato o di ottenere un risarcimento assai prossimo a questo.

Il sistema più rapido per stabilire una cifra di base per gli arbitrati per legge ci sembra sia quello di monitorare le conciliazioni che avvengono ora nelle Direzioni Territoriali del Lavoro e, sulla base delle cifre pattuite volontariamente tra le parti, stabilire in futuro un cifra base delle conciliazioni come indennità automatica⁵.

Il dibattito sull'opportunità di cambiare le norme sul licenziamento individuale ha visto su parti contrapposte chi sosteneva che i licenziamenti individuali sono così pochi nelle grandi imprese che non valeva la pena di cambiare la legge e chi invece sosteneva che l'articolo 18 allontanava gli investimenti stranieri (quindi le grandi imprese) dall'Italia. Noi aggiungiamo una considerazione sulla base della letteratura economica esistente che non riguarda le grandi aziende (che certamente non hanno particolare bisogno dei licenziamenti individuali) bensì la moltitudine di imprese che stanno poco sopra i 15 dipendenti (indicativamente da 16 a 50): mentre pare vero che la discontinuità dell'articolo 18 per le imprese sopra i 15 dipendenti non sembra avere effetti sostanziali sulla dimensione d'impresa (Schiavardi e Torrini, 2004; Borgarello *et al.*, 2008), tuttavia la legge sul licenziamento individuale sembra introdurre almeno due distorsioni rilevanti nelle scelte delle imprese vicino alla soglia dei 15 dipendenti.

La prima distorsione riguarda salari inferiori che compenserebbero la maggiore sicurezza del posto di lavoro nelle imprese poco sopra i 15 dipendenti, la seconda distorsione riguarda la proporzione maggiore di contratti a termine impiegati nelle imprese sopra la soglia al fine di aggirare la norma dell'articolo 18. Leonardi e Pica (2012) analizzando la riforma del 1990 che ha incrementato la "severance pay" nelle imprese sotto i 15 dipendenti, mostrano che i nuovi assunti in queste imprese hanno ricevuto in media salari inferiori del 3% presumibilmente a compensazione della maggiore protezione del posto di lavoro. Hijzen *et al.* (2012) mostrano evidenza preliminare che la composizione della forza lavoro delle imprese poco sopra i 15 dipendenti è spostata a favore dei contratti temporanei che permettono in qualche modo di evitare il raggiungimento della soglia di applicazione dell'articolo 18.

⁵ Al momento non sono disponibili dati certi né sulla proporzione di licenziamenti individuali che si concludono con conciliazioni né delle cifre pattuite, pare però che in alcune direzioni territoriali la proporzione sia ben superiore al 50% e le cifre si aggirino sulle 12 mensilità. Nel caso di proporzioni di conciliazioni così elevate si può certamente parlare di un successo della nuova norma che ci avvicina molto al caso tedesco in cui le conciliazioni sono circa il 60% dei casi (VENN D., 2009).

Queste due distorsioni e i dati sull'incidenza dei licenziamenti individuali nelle imprese sotto i 15 dipendenti inducono a pensare che sia stato giusto intervenire per rendere marginalmente più facile il licenziamento individuale nelle imprese sopra i 15 dipendenti, in modo che esse possano fare più licenziamenti individuali e meno licenziamenti collettivi. Utilizzando ancora una volta i dati del Veneto elaborati da Veneto-Lavoro si possono notare due fenomeni interessanti: in primo luogo la maggior parte delle imprese in Veneto ha proceduto a ridurre la forza lavoro durante la crisi per meno di 5 unità (quindi sotto il numero minimo necessario per il licenziamento collettivo) e in secondo luogo i licenziamenti individuali in Veneto nel 2012 sono stati di 26mila unità contro i 10mila licenziamenti collettivi.

Ad oggi i licenziamenti individuali avvengono in misura preponderante nelle imprese sotto i 15 dipendenti ma queste cifre fanno ritenere che anche le imprese di poco sopra la soglia dei 15 dipendenti abbiano bisogno di procedure più rapide per ridurre la forza lavoro di poche unità in caso di crisi. Le imprese sopra i 15 dipendenti oggi ricorrono quasi esclusivamente ai licenziamenti collettivi e potrebbero aver beneficio dalla riforma che facilita l'aggiustamento di manodopera sui piccoli numeri del licenziamento individuale invece che sui grandi numeri del licenziamento collettivo. Potenzialmente questa riforma potrebbe dare un incentivo alle imprese poco sopra i 15 dipendenti a ridurre le procedure di licenziamento collettivo e sostituirle con un maggior numero di licenziamenti individuali (per il quale, dopo la riforma, il lavoratore viene comunque compensato con la procedura di conciliazione). Questo fenomeno, se si realizzerà, potrebbe avere conseguenze non indifferenti – e un effetto non chiaro a priori sul numero totale di licenziamenti – in quanto le imprese tra i 15 e i 50 dipendenti in Italia occupano circa il 20% del totale degli occupati.

5. - Il nuovo assetto degli ammortizzatori sociali e delle politiche attive

Il sistema degli ammortizzatori italiani, pensato per tutela avverso il rischio di perdita del reddito da lavoro i lavoratori subordinati standard impiegati nella grande impresa, rileva oggi tutta la sua inadeguatezza ad offrire una protezione alla infinitamente variegata utenza dei lavoratori disoccupati. Spesso chi ha un contratto parasubordinato e perde il lavoro è escluso da molte protezioni pur a fronte di un rischio di disoccupazione quasi cinque volte superiore ai lavoratori subordinati; cassa integrazione guadagni ordinaria e straordinaria e indennità di

mobilità sono destinate soltanto ai lavoratori impiegati nelle aziende industriali con più di 15 dipendenti e in quelle commerciali con più di 50 dipendenti.

La riforma intende distinguere i sussidi di disoccupazione dalle integrazioni salariali in caso di riduzione delle ore lavorate in costanza di rapporto introducendo un trattamento di disoccupazione di applicazione universalistica, così come è in altri sistemi continentali che pure utilizzano strumenti simili alla cassa integrazione (come la *Kurzarbeit* tedesca). Dopo una breve descrizione di come funzionano ad oggi i sussidi di disoccupazione e il sistema della Cig, analizziamo come la riforma Fornero affronta i punti (i) e (ii).

Ad oggi in Italia esistono due tipi di indennità: a requisiti ordinari e a requisiti ridotti. L'indennità ordinaria di disoccupazione è prevista per tutti i lavoratori dipendenti in condizione di disoccupazione involontaria che abbiano versato almeno un contributo settimanale all'assicurazione contro la disoccupazione almeno due anni prima dell'inizio del periodo di disoccupazione (requisito assicurativo), e che abbiano versato almeno 52 contributi settimanali a tale assicurazione nel biennio che precede l'inizio della disoccupazione (requisito contributivo).

La misura dell'indennità è pari al 60% della retribuzione precedente per i primi sei mesi indennizzati, al 50% per ulteriori due e al 40% per altri quattro. La durata massima del beneficio è di 8 mesi o di 12 mesi per gli *over 50*.

L'accesso all'indennità ordinaria presenta diversi problemi di accesso per i lavoratori atipici: innanzitutto apprendisti e lavoratori parasubordinati ne sono formalmente esclusi; poi il requisito assicurativo e quello contributivo tendono ad escludere dalla prestazione i lavoratori da poco entrati sul mercato del lavoro e quelli con carriere discontinue. Esiste inoltre un minimo contributivo per l'accredito di un'intera settimana presso l'assicurazione contro la disoccupazione che tende invece ad escludere i lavoratori con retribuzioni complessive più basse per esempio i *part-time*.

Berton, Richiardi e Sacchi (2009) sulla base dei dati Inps stimano che i requisiti imposti per l'accesso escludono dall'indennità di disoccupazione ordinaria circa il 30% dei lavoratori a tempo indeterminato *part-time*, e circa il 60% dei titolari di un contratto di formazione e lavoro, dei lavoratori con contratto a tempo determinato e dei lavoratori interinali.

Molti lavoratori atipici sono esclusi anche dall'indennità di disoccupazione a requisiti ridotti malgrado l'accesso richieda solo che siano state prestate almeno 78 giornate di lavoro nell'anno per il quale la prestazione è richiesta. Contrariamente all'indennità ordinaria da cui sono esclusi, possono accedervi anche gli ap-

prendisti, tuttavia mancano l'accesso all'indennità il 78,9% degli apprendisti, il 50% dei titolari di contratti di formazione e lavoro e degli interinali e il 40% dei lavoratori a tempo determinato⁶.

Infine, per i lavoratori parasubordinati esiste solo una sorta di indennità di disoccupazione per i soli casi di fine lavoro, da liquidarsi in un'unica soluzione pari al 10 per cento del reddito percepito l'anno precedente (stabilita dalla legge n. 2 del 28 gennaio 2009).

L'articolo 22 del progetto di riforma istituisce presso l'Inps, a decorrere dal 1° gennaio 2013, l'Assicurazione Sociale per l'Impiego (ASpI). L'ambito di applicazione della nuova forma di sostegno viene esteso agli apprendisti e ai soci lavoratori di cooperativa che abbiano stabilito, in ragione del rapporto associativo, un rapporto di lavoro in forma subordinata.

Per i co.co.co. che sono tuttora esclusi dall'ASpI, si rafforza il meccanismo di protezione "una tantum" in favore dei collaboratori coordinati e continuativi, in quanto esclusi dall'ambito di applicazione dell'ASpI. In tal senso, per poter risultare destinatari dell'indennità, i collaboratori dovranno soddisfare congiuntamente le seguenti quattro condizioni: aver operato, nell'anno precedente, in regime di mono-committenza; aver conseguito, nell'anno precedente, un reddito complessivo valido ai fini fiscali non superiore a 20.000 euro; che risulti accreditato, con riguardo all'anno di riferimento, presso la Gestione separata dell'Inps almeno un numero di mensilità; aver avuto un periodo di disoccupazione ininterrotto di almeno due mesi nell'anno precedente; che risultino accreditate nell'anno precedente almeno quattro mensilità presso la medesima Gestione separata dell'Inps.

I requisiti di eleggibilità dell'ASpI non cambiano rispetto alla vecchia indennità ordinaria di disoccupazione, ma migliorano la generosità e la durata. La nuova indennità mensile risulta essere pari al 75% della retribuzione mensile nei casi in cui quest'ultima non superi i 1.180 euro (annualmente rivalutata sulla base dell'indice Istat dei prezzi al consumo); lo stesso articolo stabilisce che qualora la retribuzione mensile sia superiore al predetto importo, l'indennità verrà determinata aggiungendo al 75% dell'importo di euro 1.180 una somma pari al 25% del dif-

⁶ Le prestazioni di malattia e maternità sono state recentemente estese ai lavoratori atipici che tuttavia hanno spesso gravi difficoltà ad accedere alle prestazioni per requisiti d'accesso restrittivi e hanno prestazioni inadeguate. Mentre per i lavoratori subordinati tali prestazioni non sono sottoposte a requisiti contributivi, lavoratori parasubordinati sono esclusi in buona parte dall'accesso agli schemi per loro formalmente previsti a causa dell'operare di requisiti contributivi e del meccanismo di accredito dei contributi versati.

ferenziale tra la retribuzione mensile e il predetto importo. All'importo viene applicata una riduzione nella misura del 15% dopo i primi sei mesi di fruizione e di un ulteriore 15% dopo il dodicesimo mese di fruizione.

La durata massima dell'ASpI è di 12 mesi tranne che per i lavoratori di età pari o superiore ai 55 anni per i quali si può arrivare a 18 mesi. Questa estensione della durata per i lavoratori anziani è quanto più necessaria in quanto viene contestualmente eliminata l'indennità di mobilità che spesso accompagnava i lavoratori anziani fino alla pensione. Da questo punto di vista 18 mesi sono anche pochi, in quanto è sicuramente condivisibile fare dei sussidi di disoccupazione lo strumento principale di aiuto al reddito nei casi di perdita del lavoro, bisogna anche ricordare però che da una parte non c'è un effettivo ostacolo al licenziamento dei lavoratori più anziani (vedi sopra) e soprattutto in Italia a differenza di molti altri paesi, all'esaurirsi del sussidio di disoccupazione non esiste un sussidio di povertà per chi non è riuscito a trovare una nuova occupazione.

Essenzialmente, con la riforma la platea degli aventi diritto all'indennità ordinaria si amplia seppur di poco (soci cooperative e apprendisti) e si rafforza l'"una tantum" per i co.co.co.; vengono estese le durate e viene incrementata la misura della prestazione, tuttavia per i parasubordinati non cambia nulla. Poiché i requisiti contributivi e assicurativi per l'accesso ai sussidi non sono cambiati, molti atipici (lavoratori a termine e in somministrazione, oltre ai *part-time*) continuano ad essere esclusi dall'ASpI, per questi ultimi, però, c'è una novità nella mini-ASpI.

La mini-ASpI viene pagata per una durata massima pari alla metà delle settimane contribuite nei 12 mesi precedenti la disoccupazione, e prevede una prestazione di entità pari all'ASpI, per un massimo di 6 mesi (in presenza di una contribuzione di 52 settimane nell'ultimo anno). Per accedervi il disoccupato deve aver contribuito per almeno 13 settimane nei 12 mesi precedenti la disoccupazione, tuttavia non esiste un requisito assicurativo, cioè non è richiesto che il disoccupato abbia lavorato per un certo periodo di tempo prima della disoccupazione. Questo amplia notevolmente la platea di quanti hanno accesso effettivo alle indennità di disoccupazione, principalmente fra gli atipici. Sacchi (2012a e 2012b) stima che, rispetto al precedente sistema di sussidi, gli esclusi tra i lavoratori a tempo determinato passerebbero dal 40% al 20%, tra i somministrati dalla metà a un quarto, tra i lavoratori a tempo indeterminato *part-time* dal 20% al 5%. In totale l'ampliamento della platea dei potenziali beneficiari sarebbe di più di 1,5 milioni di persone che erano escluse dal precedente sistema. I lavoratori esclusi dai benefici del nuovo regime ASpI e Mini-ASpI rimarrebbero solo poco più di 500mila su un totale di 13,6 milioni di occupati dipendenti. L'amplia-

mento della platea di lavoratori coperti da un sussidio di disoccupazione è, insieme al riordino della Cig con l'abolizione dell'indennità di mobilità, la maggiore novità della riforma per quanto riguarda gli ammortizzatori sociali.

Ad oggi la cassa integrazione guadagni ordinaria (Cig) è concessa – *a*) per situazioni aziendali dovute ad eventi transitori e non imputabili all'imprenditore o agli operai, *b*) ovvero determinate da situazioni temporanee di mercato – solo alle aziende industriali e, per causali sostanzialmente analoghe, alle imprese edili e agricole, quest'ultime governate da gestioni speciali. La Cigs è invece concessa per "ristrutturazione, riorganizzazione o conversione aziendale" ed anche per "crisi aziendale" soltanto alle aziende industriali, incluse quelle edili, con più di 15 dipendenti, alle aziende commerciali con più di 200 dipendenti, alle cooperative di produzione lavoro, alle imprese appaltatrici di servizi di pulizia, alle imprese artigiane che subiscono l'influsso gestionale prevalente di un'altra impresa ammessa al trattamento. Possono invece godere dell'indennità di mobilità soltanto i lavoratori espulsi a seguito di licenziamenti collettivi nelle aziende che hanno accesso alla Cigs.

Con l'avvento della crisi la legge n. 2 del 2009 si è limitata ad estendere in deroga la fruizione di cassa integrazione guadagni straordinaria (Cigs) e indennità di mobilità alle imprese commerciale con più di 50 dipendenti, alle agenzie di viaggio e turismo con più di 50 dipendenti, alle imprese di vigilanza con più di 15 dipendenti. In seguito la cassa in deroga è stata estesa, nei limiti delle risorse economiche rese disponibili⁷, a tutte le piccole imprese industriali, commerciali e artigianali che assorbono gran parte della forza di lavoro italiana. Questo provvedimento, pur reso necessario dall'urgenza della crisi, ha reso ancor più evidenti i limiti della CIG italiana: il nostro sistema di ammortizzatori sociali è "balcanizzato" ed è difficilmente compatibile con la disciplina comunitaria degli aiuti di stato in quanto limita il suo ambito di applicazione ad alcuni settori e alle imprese in possesso di determinati requisiti dimensionali e attribuisce all'autorità amministrativa ampi margini di discrezionalità nella concessione del trattamento alle imprese. Inoltre c'è l'ovvio problema che chi oggi riceve la cassa in deroga non ha mai versato i relativi contributi.

Occorre al riguardo rammentare che la Corte di giustizia ha ricompreso tra gli aiuti di Stato vietati dall'art. 87 del Trattato comunitario ogni misura che si

⁷ Anno 2009: € 289 milioni; Anno 2010: € 304 milioni; Anno 2011: € 304 milioni; dall'Anno 2012: € 54 milioni.

traduca in una spesa pubblica o in una mancata entrata per l'erario pubblico sotto qualsiasi forma si realizzi⁸, indipendentemente da quale sia la natura del soggetto o dell'ente deputato all'erogazione dell'aiuto stesso, rilevando esclusivamente l'effetto che si produce in favore delle imprese beneficiarie e la riconducibilità in via diretta o indiretta a un intervento "pubblico". Osservando questi canoni il diritto comunitario ritiene inammissibili gli aiuti volti "al mantenimento dell'occupazione", equiparandoli sostanzialmente agli aiuti "al funzionamento delle imprese"⁹. Si tratta di una vera e propria messa al bando degli aiuti "assistenziali", quelli cioè che si limitano a rallentare il declino di imprese in crisi irreversibili o di interi settori non più competitivi sui mercati internazionali e a posticipare semplicemente i problemi occupazionali che ne conseguono. Il diritto comunitario, invece, non pone limitazioni di sorta ed anzi promuove aiuti economici e l'offerta di servizi ai lavoratori una volta espulsi dal ciclo produttivo od anche in caso di sospensioni temporanee dall'attività, laddove tali ammortizzatori siano indifferenziati per settori o per tipologie di imprese e tanto più non condizionati a valutazioni discrezionali della pubblica amministrazione.

Da questo punto di vista è condivisibile l'intento della riforma Fornero di omogeneizzare i trattamenti disponibili tra imprese diverse estendendo la cassa integrazione (e i relativi obblighi contributivi) a tutte le imprese su base volontaria, e su base obbligatoria per quelle con più di 15 dipendenti. Altrettanto giusto eliminare dalla Cigs la causale di procedura concorsuale con cessazione di attività (abolita a partire dal 2016) e l'indennità di mobilità: si volta pagina rispetto a decenni di uso distorto della Cassa integrazione guadagni nelle crisi occupazionali aziendali. In generale, l'utilizzo di Cigo e Cigs continua ad essere uno strumento prezioso per le imprese in crisi: esse pagano un contributo speciale pari al massimo all'8% del salario orario, mentre i contributi sociali sono presi in carico dallo stato attraverso il meccanismo della contribuzione figurativa. Tuttavia lo spirito della Cassa torna ad essere quello di fornire una rete di aiuto durante crisi e ristrutturazioni transitorie, non quello di mantenere in vita aziende decotte e bloccare la ricollocazione dei lavoratori in altre attività.

⁸ v. per tutte Corte di giustizia, sent. 2 luglio 1974, C-173/73, *Italia c. Commissione*, in *Racc.*, 1974, pp. I-709; Corte di giustizia, sent. 5 ottobre 1999, C-251/97, *Francia c. Commissione*, in *Racc.*, 1999, pp. I-6639; Corte di giustizia, sent. 8 novembre 2001, C-143/99, *Adria-Wien Pipeline e Wietersdorfer & Peggauer Zementwerke*, in *Racc.*, 2001, pp. I-8365; Corte di giustizia, sent. 12 dicembre 2002, C-5/01, *Regno del Belgio c. Commissione*, in *RIDL*, 2003, II, p. 709.

⁹ Cfr. Regolamento CE n. 2204/2002, art. 9, co. 5.

Per la stessa ragione è meno lodevole aver mantenuto in vita l'indennità speciale per i lavoratori agricoli e nell'edilizia, che servono oggi per lo più a integrare i salari di chi già lavora, piuttosto che ad aiutare chi ha perso il lavoro e ne sta cercando un altro.

Una novità da segnalare è che l'estensione della Cig e Cigs ai lavoratori dei settori non coperti non avviene direttamente in capo all'Inps ma attraverso fondi di solidarietà bilaterali costituiti nell'ambito di accordi e contratti collettivi (e in via suppletiva per via di un decreto).

L'istituzione dei fondi è obbligatoria per tutti i settori non coperti dalla normativa in materia di integrazione salariale per le imprese che occupano mediamente più di 15 dipendenti. Il testo prevede che i fondi di cui all'articolo in esame possono avere, in maniera aggiuntiva, anche le seguenti finalità:

- a) assicurare ai lavoratori una tutela in caso di cessazione dal rapporto di lavoro, integrativa rispetto all'assicurazione sociale per l'impiego;
- b) prevedere assegni straordinari per il sostegno al reddito, riconosciuti nel quadro dei processi di agevolazione all'esodo, a lavoratori che raggiungano i requisiti previsti per il pensionamento di vecchiaia o anticipato nei successivi cinque anni;
- c) contribuire al finanziamento di programmi formativi di riconversione o riqualificazione professionale, anche in concorso con gli appositi fondi nazionali o comunitari.

A sfavore della scelta di estendere la Cig attraverso la costituzione di enti bilaterali invece che direttamente presso l'Inps c'è la consuetudine di lottizzazione degli stessi e il fatto che settori economicamente deboli avranno fondi poco generosi, a favore c'è la responsabilizzazione delle aziende e dei sindacati del settore. L'articolo 45 detta la disciplina finanziaria dei fondi, prevedendo in particolare: l'obbligo di bilancio in pareggio e la impossibilità di erogare prestazioni in carenza di disponibilità finanziarie.

Infine la riforma ha ben presente la tradizionale debolezza italiana nelle politiche attive e dei servizi per la ricollocazione dei disoccupati e (meritoriamente) prevede l'istituzione presso il Ministero del Lavoro e delle Politiche Sociali di un sistema permanente di monitoraggio e valutazione. Tuttavia sul punto dei servizi all'impiego non sono state fatte scelte nette, anche per l'incertezza attuale sul destino delle Province. Il nostro sistema dei servizi è di competenza regionale e provinciale ed è separato dalla gestione dei sussidi di disoccupazione. Questo rende più difficile condizionare l'erogazione dei sussidi alla ricerca di un lavoro da parte del beneficiario e rende la performance del servizio molto dipendente dall'effi-

cienza dei singoli enti locali. Alcuni paesi, tra questi Germania, Regno Unito e più recentemente la Francia, hanno optato invece per l'integrazione dei servizi reali con la funzione di erogazione dei trattamenti economici in caso di disoccupazione. L'esperienza mostra che questa scelta facilita la programmazione, migliora la gestione delle risorse e promuove i comportamenti attivi dei beneficiari dei sussidi. L'alternativa alla separazione delle competenze in materia di politiche del lavoro potrebbe essere un organismo unico di livello nazionale articolato a livello territoriale, che faccia le funzioni di raccordo tra domanda e offerta di lavoro oggi svolte dai centri pubblici per l'impiego delle Province.

6. - Conclusioni

L'equilibrio complessivo della riforma è fondato sull'azione contestuale su due fronti: ad un indebolimento della protezione dell'art. 18 corrisponde un serio tentativo di evitare gli abusi troppo frequenti nell'utilizzo del lavoro autonomo. Nella versione iniziale della riforma entrambi i fronti dell'azione erano caratterizzati da novità molto incisive sia per quanto riguarda la disciplina del licenziamento individuale sia per quanto riguarda le regole di utilizzo del lavoro autonomo. Nel corso delle trattative politiche e sindacali entrambi i fronti della riforma sono stati ammorbiditi. A questo punto il successo della riforma si misurerà su quanto la volontà del legislatore, in termini di norme di licenziamento e di riassorbimento del lavoro precario, si trasformerà in pratica giurisprudenziale. A questo proposito è importante che il linguaggio della legislazione sia chiaro e il ricorso ai giudici non eccessivo. Per quanto riguarda la semplificazione del ricorso in sede giudiziale abbiamo proposto nel corso del testo una norma che prende ad esempio il caso tedesco, migliorandolo ed adattandolo ai parametri italiani.

Molte lezioni si possono trarre dagli eventi che si sono succeduti nel corso – tutto sommato breve – della stesura e dell'approvazione della riforma. Il cammino della riforma è stato assai difficile in quanto si è deciso di applicare le nuove norme sui licenziamenti a tutti i contratti e non solo ai contratti dei neo-assunti. Questo ha generato apprensione nell'opinione pubblica e tensione nella politica. Riteniamo tuttavia, e non solo per ragioni di equità, che il capitale politico speso per ridurre la protezione dell'articolo 18 su tutti i contratti sia stato ben speso. A nostro parere sia il progetto di Boeri e Garibaldi (2010) sia il progetto di Ichino (2009) sono di più facile applicazione proprio oggi dopo questa riforma dell'articolo 18 piuttosto che in assenza della riforma stessa. Entrambi i progetti riguar-

dano solo i nuovi assunti: quello di Boeri e Garibaldi prevede un contratto unico di inserimento cui si applica la tutela reale dell'articolo 18 solo dopo tre anni; quello di Ichino prevede l'introduzione di contratti con minore tutela contro il licenziamento individuale per ragioni economiche (ma molto maggiore tutela risarcitoria e di riqualificazione al lavoro) esclusivamente per i nuovi assunti e la graduale sostituzione nel corso del tempo dei vecchi contratti protetti dall'art. 18. L'attuazione di entrambi i progetti è più facile in presenza di un articolo 18 depotenziato: prima della riforma, infatti, nessun imprenditore avrebbe agevolmente lasciato passare i tre anni di un contratto di inserimento per farlo trasformare automaticamente in un contratto a tempo indeterminato protetto dall'articolo 18, e nessun lavoratore avrebbe volentieri cambiato posto di lavoro se cambiando lavoro e contratto avrebbe nel contempo rinunciato alla tutela reale¹⁰. Rimangono anche da investigare alcune riserve di tipo giuridico sulla coesistenza di contratti diversi (con diritti in caso di licenziamento assai diversi) tra lavoratori di una stessa impresa in mansioni equivalenti.

È da notar infine che per la prima volta finalmente si istituisce un'unica assicurazione contro la disoccupazione uguale per tutti i lavoratori dipendenti e si riconduce la Cassa integrazione guadagni alla sua funzione originaria: quella di tenere i lavoratori legati all'impresa nelle situazioni di difficoltà temporanea, non quella di sostenere con il denaro pubblico imprese senza nessuna prospettiva. Anche da questo punto di vista le resistenze sono state molteplici e le richieste di rimandare l'applicazione delle nuove regole in ragione della crisi economica pressanti. Ciò nonostante questa parte della riforma ha retto per ora in modo sostanziale alle pressioni e i nuovi principi dovrebbero essere applicati da subito.

Una critica alla riforma riguarda l'interazione con la riforma delle pensioni, avendo quest'ultima in un breve lasso di tempo allungato la permanenza dei lavoratori anziani sul posto di lavoro. In effetti nell'ultimo anno stiamo assistendo al prolungamento in attività dei lavoratori al di sopra dei 55 anni (Cnel, 2012) ed ad una contemporanea riduzione delle assunzioni tra i giovani. L'effetto di sostituzione tra giovani e anziani è tuttavia presumibilmente un effetto di breve periodo dovuto al rapido aumento dell'offerta di lavoro degli ultra 55enni, è infatti dimostrato in diversi lavori che nel lungo periodo non esiste relazione alcuna tra il tasso

¹⁰ Non esistono studi a conoscenza degli autori sul potenziale effetto di *lock in* dei lavoratori dovuto alla riduzione dei cambi volontari di posto di lavoro. Gli effetti sulla riallocazione delle risorse ai settori più produttivi sono potenzialmente tanto più seri in un paese come l'Italia dove la mobilità è volontaria più che involontaria, i.e. i flussi dentro e fuori la disoccupazione sono limitati e invece prevale proprio lo spostamento dei lavoratori tra posti di lavoro.

di occupazione dei giovani e quello degli anziani (Galasso e Leonardi, 2011).

In generale la critica prevalente alla riforma riguarda il *timing*. I molti contestatori sostengono che sarebbe stato meglio aspettare la fine della crisi per attuare una riforma che in qualche modo facilita i licenziamenti e rende più difficile l'abuso del lavoro autonomo rischiando di diminuire in questo modo l'occupazione. In realtà però le riforme del mercato del lavoro sono già state rinviate molte volte e la presenza di un governo dalla larga maggioranza con un vincolo europeo da rispettare costituiva una combinazione difficilmente ripetibile. È in qualche modo inevitabile inoltre che una riforma sia pensata per risolvere problemi strutturali del mercato del lavoro (l'eccesso di contratti precari e di lavoro autonomo falso, la rigidità delle procedure di licenziamento individuale rispetto a quelle collettive, l'esclusione di milioni di lavoratori dagli ammortizzatori sociali) e non per fronteggiare una crisi occupazionale che, sebbene lunga come quella presente, non è destinata a durare all'infinito. È opinione comune che le riforme, soprattutto quelle che riguardano i licenziamenti, non abbiano un effetto sul numero totale degli occupati quanto piuttosto sui flussi di assunzioni e licenziamenti. Per le carenze temporanee di domanda di lavoro in periodi di crisi esiste uno strumento come la cassa integrazione, a sua volta compreso nella riforma e allargato ad una platea molto più ampia di imprese.

In conclusione molti dei principi contenuti nella riforma sono apprezzabili e vanno portati avanti senza tentennamenti, il successo del cambiamento dipenderà molto dalla volontà delle parti: aziende, sindacati, giudici del lavoro.

BIBLIOGRAFIA

- ANASTASIA B., «Sulla trappola della precarietà: quali indicazioni dalle ricerche empiriche e dalle statistiche disponibili?», in DELL'ARINGA C. - TREU T., *Giovani senza Futuro? Proposte per una Nuova Politica*, Bologna, il Mulino, 2012a.
- .-, «Effetto Fornero sul lavoro intermittente», in *www.lavoce.info*, <http://www.lavoce.info/articoli/lavoro/pagina1003439.html>, 2012b.
- BERTON F. - RICHIARDI M. - SACCHI S., *Flex-Insecurity. Perché in Italia la flessibilità diventa precarietà*, Bologna, il Mulino, 2009.
- BOERI T. - GARIBALDI P., «Tutti i vantaggi del contratto unico», in *www.lavoce.info* <http://archivio.lavoce.info/articoli/pagina1001663-351.html> e disegno di legge <http://www.senato.it/service/PDF/PDFServer/BGT/00459782.pdf>, 2010.
- BORGARELLO A. - GARIBALDI P. - PACELLI L., «Employment Protection Legislation and the Size of Firms», *Giornale degli Economisti*, no. 63, 2004.
- BOSIO G. - LEONARDI M., «The Impact of Bologna Process on the Graduate Labour Market: Demand and Supply», *Giornale degli Economisti*, no. 69, 2011.
- CAPPELLARI L. - DELL'ARINGA C. - LEONARDI M., «Flexible Employment, Job Flows and Labor Productivity», *Economic Journal*, no. 122, 2012a.
- .-,-.-, «Temporary Employment in Italy», *CESifo DICE Report*, 2012b.
- CAPPELLARI L. - LEONARDI M., «A favore di un sistema di vocational tertiary education in Italia», in DELL'ARINGA C. - TREU T., *Giovani senza futuro? Proposte per una nuova politica*, Bologna, il Mulino, 2012.
- CNEL, *Rapporto sul mercato del lavoro 2011-2012*, Roma, 2012.
- GALASSO V. - LEONARDI M., «Lavoro e pensioni», in GOLINI A. - ROSINA A., *Il secolo degli anziani. Com'è destinata a cambiare la società italiana*, Bologna, il Mulino, 2011.
- HIJZEN A. - MONDAUTO L. - SCARPETTA S., *The Perverse Effects of Job Security Provisions on Job Security in Italy: Results from a Regression Discontinuity Design*, OECD, mimeo, 2012.
- ICHINO P., «Progetto Flexsecurity», in *www.pietroichino.it* <http://www.pietroichino.it/?p=1079> e disegno di legge <http://www.pietroichino.it/wp-content/uploads/2009/09/dll-1481.pdf>, 2009.
- GOERKE L. - PANNENBERG M., «Take it or Go to Court - The Impact of Sec. 1a of the German Protection against Dismissal Act on Severance Payments», *CESifo Working Paper Series*, no. 3120, CESifo Group Munich, 2010.
- LEONARDI M. - PALLINI M., «Quale riforma per la legge Biagi», in *www.lavoce.info* <http://archivio.lavoce.info/articoli/pagina2150-351.html>, 2006.
- LEONARDI M. - PICA G. «Who Pays for It? The Heterogeneous Wage Effects of Employment Protection Legislation», in stampa, *Economic Journal*, 2012.
- PALLINI M. (a cura di), *Il lavoro a progetto in Italia e in Europa*, Bologna, il Mulino, 2006.

- RANCI C. (a cura di), *Partite Iva. Il lavoro autonomo nella crisi italiana*, Collana “Studi e Ricerche”, Bologna, il Mulino, 2012.
- SACCHI S., «Il mondo è cambiato ma non per gli ammortizzatori sociali - parte I», in *www.NelMerito.com*, http://www.nelmerito.com/index.php?option=com_content&task=view&id=1673&Itemid=1, 2012a.
- .-, *Riformare il lavoro (e gli ammortizzatori sociali)*, Bologna, il Mulino, n. 5/12, 2012b.
- SCHIVARDI F. - TORRINI R., «Identifying the Effects of Firing Restrictions Through Size-Contingent Differences in Regulation», *Labour Economics*, no. 15, 2008.
- VENETO LAVORO, «Il mercato del lavoro veneto nel terzo trimestre 2012», http://www.venetolavoro.it/portal/html/v4/osservatorio/nota/Bussola_amm_novembre_2012.pdf, 2012.
- VENN D., «Legislation, Collective Bargaining and Enforcement: Updating the OECD Employment Protection Indicators», OECD, *Social, Employment and Migration*, no. 89, 2009.

SAGGI

L'economia non osservata fra evasione e crimine: una rivisitazione del *Currency Demand Approach* con una applicazione al contesto italiano

Guerino Ardizzi*

Banca d'Italia, Roma

Massimiliano Piacenza***

Università di Torino

Carmelo Petraglia**

Università della Basilicata

Gilberto Turati****

Università di Torino

This study contributes to the debate on estimating the size of underground economy, by proposing a reinterpretation of the Currency Demand Approach (CDA). Three main innovations are introduced. First, a direct measure of cash transactions is taken as dependent variable in the money demand equation. Second, indicators of detected tax evasion substitute difficult-to-find proxies for all the relevant determinants of noncompliance. Third, a control for cash used in criminal activities is introduced. The application of this 'modified-CDA' to Italian provinces reveals that the provinces in Centre-North exhibit a higher incidence on GDP of both the underground economy and illegal production.

[JEL Classification: E26; E41; H26; K42; O17].

Keywords: non-observed economy; underground production; crime, currency demand approach.

* <guerino.ardizzi@bancaditalia.it>, Servizio Supervisione sui Mercati e sul Sistema dei Pagamenti.

** <carmelo.petraglia@unibas.it>, Dipartimento di Matematica, Informatica ed Economia.

*** <massimiliano.piacenza@unito.it>, Dipartimento di Scienze Economico-Sociali e Matematico-Statistiche.

**** <gilberto.turati@unito.it>, Dipartimento di Scienze Economico-Sociali e Matematico-Statistiche.

Gli autori desiderano ringraziare anonimi *referee*, Fabio Bagliano, Fabio Berton, Raul Caruso, Domenico De Palo, Emma Galli, Alessandra Sanelli, Paolo Sestito, Roberta Zizza e i partecipanti al *Workshop* "Alle radici del crimine" (Università Cattolica, Milano, maggio 2011), al *Lunch Seminar* tenuto presso la Banca d'Italia (Roma, giugno 2011) e alla XXIII Conferenza SIEP (Pavia, settembre 2011), per i commenti forniti su versioni preliminari di questo lavoro. Gli autori sono gli unici responsabili di eventuali errori residui.

1. - Introduzione

Una recente ricerca ha mostrato che nelle economie dove maggiore è l'utilizzo di strumenti di pagamento elettronici (es. carte di pagamento e bonifici via Internet), minore è l'incidenza dell'economia sommersa (Schneider, 2009). In base a tale studio, l'Italia detiene due primati non invidiabili tra i paesi dell'Europa continentale: quello della quota di transazioni che vengono regolate con pagamenti in contanti (un indice di arretratezza finanziaria)¹ e quello dell'incidenza dell'economia sommersa sul Pil (349 miliardi di euro l'anno, pari al 24% del Pil, mentre nella UE a 27 si raggiungono i 2.000 miliardi di euro, circa il 18% del Pil).

Questa evidenza suggerisce l'esistenza di una relazione tra il ricorso massiccio al contante come strumento di pagamento e l'elevata diffusione di scambi non registrati fra le transazioni "regolari" che è stata sfruttata da tempo dalla letteratura economica. Tale relazione rappresenta infatti l'ipotesi fondamentale su cui poggia l'approccio monetario alla misurazione dell'economia sommersa. Il *Currency Demand Approach* (CDA) tradizionale – una metodologia introdotta originariamente da Cagan (1958) e in seguito affinata da Tanzi (1980, 1983) – si basa sulla stima econometrica della domanda di contanti in eccesso rispetto alla preferenza per la liquidità *standard*, ovvero riconducibile alla decisione di effettuare transazioni irregolari così da evitare gli adempimenti fiscali e contributivi previsti dall'ordinamento vigente. La dimensione dell'economia sommersa viene quindi stimata applicando la velocità di circolazione della moneta rilevata nell'economia ufficiale alla domanda di contanti per transazioni irregolari così ottenuta.

Sebbene largamente adottato in letteratura, il CDA è stato oggetto di severe critiche (si veda, in particolare, Schneider e Enste, 2000 e 2002):

(1) la necessità di ricorrere alla teoria quantitativa della moneta (equazione di Fisher) e di assumere la medesima velocità di circolazione nell'economia ufficiale e in quella sommersa;

¹ Il numero di pagamenti per abitante con strumenti diversi dal contante in Italia è di 66 unità, a fronte di 170 in media nell'Area Euro (BANCA D'ITALIA, 2009, p. 277). Con riferimento alle operazioni in contante, non vi sono criteri esaustivi per misurare un fenomeno che per definizione è "anonimo". Si può stimare (DE PLOEY W., DENECKER O. e KEHRMANN T., 2005) che nel nostro paese il numero dei pagamenti in contante ancora oggi costituisce circa il 90% di quello complessivo (negli altri paesi industrializzati tale cifra è intorno all'80%). In termini di importo, gli acquisti in contante rappresentano una quota inferiore sebbene rilevante, pari a circa il 50% della spesa complessiva delle famiglie italiane (BANCA D'ITALIA, 2008).

- (2) l'ipotesi che la decisione di partecipare al settore sommerso sia determinata soltanto dall'eccessivo carico fiscale (imposte e contributi);
- (3) il fatto di non tenere conto nella stima della produzione sommersa della presenza dell'economia criminale, ovvero di attività illegali tipicamente regolate in contanti – quali il traffico di stupefacenti e la prostituzione – che rientrano nella sfera più ampia dell'economia *non osservata* (Oecd, 2002) e che, se trascurate, possono generare distorsioni nelle valutazioni del sommerso basate sul CDA².

La terza critica merita un'attenzione particolare, dal momento che il tentativo di isolare nell'economia non osservata complessiva la componente connessa al *crimine* (produzione *illegale*) da quella attribuibile invece all'*evasione* (produzione *sommersa*) rappresenta un aspetto poco indagato dalla letteratura empirica, soprattutto a causa delle difficoltà di delimitare i confini dell'analisi e la carenza di informazioni affidabili³. Tuttavia, quasi tutti gli studiosi sono concordi nel classificare transazioni illegali come il traffico di stupefacenti e la prostituzione fra le principali attività criminali che concorrono a formare l'economia non osservata⁴, in accordo alla definizione proposta, fra gli altri, da Smith (1994), secondo la quale «*the non-observed economy includes any market-based production of goods and services, whether legal or illegal, that escapes detection in the official estimates of GDP*». Al di là delle distorsioni nella stima del sommerso dovute alla specificazione incompleta del modello di domanda di contanti (si noti che nella maggior parte dei contesti esiste una correlazione positiva fra diffusione dell'evasione e attività criminali; cfr. OECD, 2002), pervenire ad una valutazione separata delle due componenti acquista una valenza particolare nel caso italiano. In generale, è noto che le scelte degli individui coinvolti nei due tipi di attività sono influenzate da differenti motivazioni economiche e da diversi meccanismi di incentivo (che includono il ruolo svolto dalle azioni di deterrenza), sebbene tali fattori comportamentali non vengano

² In base alla classificazione ufficiale utilizzata a livello internazionale (cfr. *Handbook for Measurement of the Non-observed Economy*, OECD, 2002) l'economia non osservata trae origine sia dalla produzione sommersa, sia dalla produzione illegale (in prevalenza legata ad attività criminali), oltre che – in misura residuale – dal settore informale e dalle inadeguatezze del sistema di rilevazione statistica.

³ Per un'analisi completa dell'economia non osservata in diversi paesi in cui si discute del contributo delle due componenti si veda lo studio di THOMAS J.J. (1992). Un'applicazione recente che tiene conto della presenza delle attività criminali e utilizza il CDA nella versione tradizionale per la stima del sommerso è quella di FERWERDA J., DELEANU I. e UNGER B. (2010).

⁴ Si veda la tassonomia originariamente proposta da LIPPERT O. e WALKER M. (1997) e successivamente ripresa e integrata da SCHNEIDER F. e ENSTE D.H. (2000, 2002) e SCHNEIDER F. (2010).

esplicitamente analizzati in questo lavoro, che ha come obiettivo principale quello di proporre una metodologia per migliorare la stima del sommerso e non l'individuazione di interventi di *policy* per il contrasto del fenomeno. Inoltre, le due componenti generano effetti molto diversi sulle finanze pubbliche, dal momento che è possibile identificare un gettito potenziale sottratto all'Erario soltanto per la produzione sommersa. Non vi è dubbio, tuttavia, che per l'Italia l'interesse principale derivi dalla profonda differenziazione delle basi imponibili e dal diverso radicamento delle attività illegali sul territorio, caratteristiche queste che rendono rilevanti almeno due questioni a cui si cercherà di fornire una risposta:

- (1) la componente sommersa dell'economia non osservata si manifesta con una diversa intensità al Sud rispetto al Nord, tenuto conto dei divari significativi fra i due contesti nel grado di sviluppo?
- (2) Le attività criminali con "centri decisionali" concentrati in prevalenza al Sud determinano un contributo maggiore delle regioni meridionali alla formazione dell'economia illegale, oppure, per effetto della mobilità delle risorse della criminalità organizzata, è ragionevole attendersi che le differenze territoriali riflettano quelle osservate per l'economia regolare?

Alla luce di queste considerazioni, il presente lavoro intende contribuire al dibattito sulla misurazione della cosiddetta *underground* o *shadow economy*, proponendo una stima per l'Italia della più ampia economia non osservata basata sul contante, in cui si distingue la produzione classificabile come sommersa – in quanto collegata ad attività legali ma esercitate irregolarmente, per via della decisione di evasione – da quella riconducibile invece all'economia criminale – in quanto associata a transazioni che rientrano fra le attività illegali. Le stime si fondano su una "rivisitazione" del CDA che consente di superare alcune delle restrizioni presenti nei precedenti studi che hanno adottato questa tecnica per "misurare" l'estensione dell'economia sommersa, modificando, in particolare, la variabile dipendente dell'equazione di domanda di contanti. In tal modo, da un lato si illustra un approccio metodologico nuovo utile a indirizzare la ricerca futura sulla stima del sommerso, dall'altro si offre un contributo all'evidenza empirica esistente per il caso italiano.

La disponibilità di una base dati originale contenente informazioni monetarie e fiscali per le province italiane relative al periodo 2005-2008 consente, in particolare, di introdurre tre aspetti innovativi. In primo luogo, invece di utilizzare una misura di *stock* di liquidità come variabile dipendente nell'equazione di domanda di circolante per stimare il valore delle transazioni sommerse, qui si impiega una misura *diretta*, ovvero il *flusso* di contante prelevato da conti correnti bancari rap-

portato al totale dei pagamenti regolati con strumenti diversi dal contante; questo elimina la necessità di ricorrere alla teoria quantitativa della moneta⁵. In tal modo, vengono superate le due critiche al CDA tradizionale concernenti la necessità di definire arbitrariamente un anno base per il calcolo della velocità di circolazione della moneta e l'ipotesi che quest'ultima sia identica per l'economia ufficiale e per quella sommersa. I dati monetari cui facciamo riferimento consentono inoltre di superare il problema – sorto dopo l'introduzione della valuta unica – della difficoltà di misurazione dello *stock* di circolante effettivamente detenuto in ciascun paese dell'Eurosistema⁶. In secondo luogo, vengono utilizzate misure *dirette* di evasione fiscale rilevata, rimuovendo così la necessità di introdurre nel modello *proxy* che siano sufficientemente esaustive delle determinanti del fenomeno, di cui il livello della pressione fiscale rappresenta solo uno dei possibili fattori motivazionali (es. Ferwerda, Deleanu e Unger, 2010). L'approccio adottato, anziché cercare di identificare le variabili-causa più importanti del comportamento di evasione, mira ad individuare degli indicatori *diretti* di diffusione del fenomeno a cui associare l'utilizzo del contante, al fine di consentirne una quantificazione in termini di produzione economica sommersa. Infine, come già discusso sopra, la metodologia che proponiamo tiene anche conto degli effetti sulla domanda di contanti dovuti alla presenza di transazioni di tipo criminale (traffico di droga e prostituzione), che per la loro specifica natura vanno tenute distinte dallo svolgimento irregolare di attività legali normalmente oggetto di prelievo fiscale.

Dai risultati ottenuti emerge un valore medio della produzione sommersa e illegale in Italia sul quadriennio 2005-2008 pari, rispettivamente, al 16,5% e 10,9% del Pil, sostanzialmente in linea con le stime disponibili. La componente di sommerso è infatti compresa fra il 16,3% e il 17,5% del Pil secondo le stime ufficiali dell'Istat (Istat, 2010). Per quanto riguarda l'economia criminale, pur in assenza di statistiche ufficiali, è possibile fare riferimento al Rapporto Italia 2008

⁵ In particolare, l'adozione di un modello *standard* di domanda di moneta, in base alla nota formula di Fisher $MV = PT$, richiede di conoscere non solo il livello di *stock* di moneta M in circolazione (saldo monetari detenuti dagli operatori economici) ma anche la sua velocità di circolazione V nel regolare il valore complessivo degli scambi di beni e servizi PT (approssimato attraverso il Pil). La conoscenza del *flusso* effettivo di moneta utilizzata nelle transazioni (es. il valore dei prelievi totali di contante dai conti correnti bancari) consente di semplificare l'analisi e di avere una misura più precisa dell'eccesso di domanda di pagamenti in contanti imputabile alle transazioni che avvengono nell'ambito dell'economia non osservata (sommersa e criminale).

⁶ I contributi nazionali agli aggregati monetari pubblicati dall'Eurosistema non includono infatti la quantità di circolante (cfr. BANCA D'ITALIA, *Bollettino economico trimestrale*).

curato dall'Eurispes, che per il 2007 stima in oltre 175 miliardi di euro il volume di affari da attività illegali, corrispondente all'11,4% del Pil. Disaggregando le stime a livello territoriale, le province del Centro-Nord, in media, sembrano esibire un'incidenza maggiore rispetto alle province del Sud sia dell'economia non osservata da evasione sia di quella associata ad attività illegali. Se il risultato sulla produzione sommersa merita approfondimenti circa le ragioni che sostengono questa maggior tendenza ad evadere, quello sull'economia illegale invita a riflettere sulla capacità delle organizzazioni criminali di "esportare" traffici illeciti nelle aree dove c'è maggiore domanda pagante.

Il lavoro è organizzato come segue. Il paragrafo 2 discute brevemente il CDA – mettendo in luce come il presente contributo si proponga di reinterpretare tale metodologia al fine di superare alcuni aspetti critici che possono influenzare la precisione delle stime ottenute – e l'evidenza empirica disponibile sul caso italiano. Nel paragrafo 3 viene introdotto il modello stimato e si descrive la metodologia adottata per la valutazione dell'economia non osservata (sommersa e criminale), sia a livello nazionale che disaggregando per macro-aree geografiche. Il paragrafo 4 presenta i risultati delle stime e il paragrafo 5 conclude lo studio. Nell'Appendice viene fornita una descrizione della base dati utilizzata nell'analisi.

2. - La letteratura di riferimento: aspetti metodologici ed evidenza sul caso italiano

In questo paragrafo viene fornita una breve rassegna della letteratura sul CDA. Prima si introducono gli aspetti metodologici caratterizzanti dell'approccio monetario alla misurazione dell'economia sommersa. In particolare, vengono esposte le ipotesi principali e gli elementi di criticità del CDA, per poi illustrare gli aspetti salienti della rivisitazione qui proposta – della quale il paragrafo 3 fornisce un'esposizione analitica – al fine di superarne i limiti (par. 2.1). In secondo luogo, vengono riportati e discussi i risultati più recenti della letteratura sulla dimensione e la dinamica del sommerso nell'economia italiana (par. 2.2).

2.1 Il Currency Demand Approach

Il CDA rappresenta il più diffuso tra i cosiddetti metodi macroeconomici-indiretti per la stima dell'economia sommersa⁷. Ispirato inizialmente da Cagan

⁷ I metodi macroeconomici-indiretti usano dati aggregati di contabilità nazionale per pervenire ad una misurazione della dimensione assoluta del sommerso di un'economia. I metodi indiretti
./.

(1958), il metodo è stato poi affinato ed applicato da Tanzi (1980, 1983) all'economia statunitense e ha conosciuto un'ampia diffusione in letteratura⁸. Il metodo perviene alla stima del peso del sommerso in due stadi:

- (1) stima econometrica della componente di domanda aggregata di moneta dedicata a transazioni nel settore sommerso;
- (2) calcolo del valore delle transazioni imputabili al sommerso via *teoria quantitativa della moneta*.

Le ipotesi di fondo su cui si basa la stima al primo stadio sono che le transazioni economiche sommerse vengano effettuate in contanti – per esigenze di non tracciabilità – e che la causa principale del sommerso sia una elevata pressione fiscale, a cui la letteratura successiva a Tanzi (1980, 1983) ha affiancato altre motivazioni quali, ad esempio, la regolamentazione eccessiva dei mercati. Il metodo prevede la stima della domanda aggregata di circolante includendo tra i regressori sia variabili esplicative *standard* della preferenza per la liquidità (es. tasso di interesse sui depositi), sia alcune *proxy* delle determinanti della partecipazione al settore sommerso. La domanda di contanti associata alle transazioni sommerse viene calcolata come differenza tra la domanda di contanti stimata nel modello completo e quella che si ottiene ponendo uguale a zero il valore delle variabili esplicative del sommerso (es. aliquota fiscale media), ovvero, la domanda di contanti solo per transazioni regolari.

Nel modello di Tanzi (1980, 1983), la variabile dipendente dell'equazione di domanda di contanti è il rapporto tra circolante delle famiglie (C) e *stock* di mo-

alternativi al CDA sono rappresentati soprattutto dal cosiddetto approccio MIMIC (*multiple indicators-multiple causes*) che – seguendo essenzialmente la logica sottostante l'analisi fattoriale – assume come variabile latente o non osservabile il “sommerso”, che può essere stimato utilizzando una serie di variabili quantitative configuranti le possibili “cause”, ossia il carico fiscale, il costo della regolamentazione, la rigidità del mercato del lavoro, etc. Per maggiori dettagli su questa metodologia di stima si veda GILES D.E.A. (1999); SCHNEIDER F. e ENSTE D.H. (2002) e SCHNEIDER F. (2010). I metodi microeconomici diretti si fondano invece su indagini campionarie condotte presso imprese e individui e mirano a fornire una valutazione in genere circoscritta a determinati settori e operatori. Una discussione completa dei vantaggi e dei limiti di questi metodi è fornita in MOGENSEN G.V., KVIST H.K., KORMENDI E. e PEDERSEN S. (1995).

⁸ CAGAN P. (1958) viene considerato l'iniziatore del metodo, avendo studiato per primo la correlazione tra domanda di circolante e pressione fiscale (per l'economia statunitense negli anni 1919-1955). Per le applicazioni in campo internazionale, si veda da ultimo FERWERDA J., DELEANU I. e UNGER B. (2010). Per le applicazioni del CDA al caso italiano che verranno discusse in seguito, oltre a SCHNEIDER F. e ENSTE D.H. (2000, 2002), che analizzano un ampio gruppo di paesi OCSE fra cui anche l'Italia, cfr. BOVI M. e CASTELLUCCI L. (2001); ZIZZA R. (2002) e CHIARINI B. e MARZANO E. (2004).

neta ($M2 = \text{circolante} + \text{depositi bancari delle famiglie}$). Il rapporto $C/M2$ viene regredito sulle seguenti variabili: aliquota media d'imposta sul reddito, quota di salari pagati in contanti sul reddito nazionale, tasso di interesse sui depositi a risparmio e reddito *pro-capite*. L'aliquota d'imposta sul reddito ha segno atteso positivo e viene interpretata ipotizzando che un aumento della pressione fiscale incentivi l'evasione e provochi un incremento della domanda di contanti come strumento di regolazione dei pagamenti nell'economia sommersa. Sfruttando l'idea originaria di Feige (1979), che individua nel volume delle transazioni monetarie una misura della dimensione complessiva di un'economia, nel secondo stadio, la stima del contributo del sommerso alla formazione di tale valore viene fornita applicando la teoria quantitativa della moneta. Feige (1979) assume che esista una relazione di proporzionalità – costante nel tempo – tra volume delle transazioni e valore nominale totale di un'economia (ufficiale e non ufficiale) e che valga l'equazione di Fisher: $MV = PT$ ($M = \text{quantità di moneta in circolazione}$, $V = \text{velocità di circolazione}$; $P = \text{livello generale dei prezzi}$; $T = \text{volume delle transazioni}$). In tal modo, avendo a disposizione una serie storica di dati, se in un anno base il prodotto derivante dall'economia non ufficiale si assume pari a zero (e così anche – per definizione – il volume delle transazioni corrispondente), nello stesso anno PT sarà pari al valore nominale ufficiale. Per tutti gli anni successivi, il valore delle transazioni attribuibili al settore sommerso può essere calcolato come differenza tra il valore nominale totale e quello ufficiale. Ispirandosi a Feige (1979), nel secondo stadio del CDA, viene calcolata per un anno base la velocità di circolazione della moneta (come rapporto tra valore nominale e quantità di circolante) e si ipotizza che questa sia la stessa che vige nell'economia irregolare. L'ammontare di sommerso si ottiene quindi moltiplicando tale velocità per l'eccesso di domanda di circolante stimato dal modello.

Schneider e Enste (2000, 2002) avanzano le seguenti critiche all'utilizzo del metodo del CDA per stimare la *shadow economy*:

- (1) non tutte le transazioni nell'ambito del settore sommerso avvengono in contanti, il che può causare una sottostima del sommerso⁹;
- (2) considerare l'eccessiva pressione fiscale come unica – o principale – determinante del sommerso è limitativo; altri fattori come la regolamentazione pub-

⁹ A questa critica se ne aggiunge un'altra, non esplicitata da SCHNEIDER F. e ENSTE D.H. (2000, 2002): non tutte le transazioni effettuate in contanti rientrano nel sommerso, benché facciano parte di una più ampia economia non osservata basata sul contante; si pensi, per esempio, ad alcune attività criminali che prevedono un vero e proprio scambio, come si discute più avanti nel lavoro.

blica dei mercati (in particolare del lavoro), la fiducia nelle istituzioni e la *tax morale* dei cittadini possono incidere significativamente sulla decisione di partecipare al settore sommerso;

- (3) un aumento della domanda di circolante può essere la conseguenza di una ridotta domanda di depositi a risparmio, anche a seguito di fenomeni di “tesaurizzazione”, o di un utilizzo del contante come mezzo di pagamento internazionale, anziché di un incremento delle transazioni nel settore sommerso;
- (4) l'ipotesi di assenza di economia sommersa in un anno base non è plausibile;
- (5) l'assunzione di una stessa velocità di circolazione della moneta nell'economia ufficiale e in quella irregolare introduce una forte restrizione nel metodo di stima.

In un recente contributo, Ferwerda, Deleanu e Unger (2010) propongono di rivisitare il modello di Tanzi allo scopo di testarne la robustezza alla luce delle suddette critiche, suggerendo che il modello potrebbe conservare – in larga parte – la propria validità se applicato alla stima dell'economia sommersa nei paesi della zona Euro dal 2002 in poi, ovvero, all'indomani dell'introduzione della moneta unica. L'applicazione al caso europeo – nella visione degli autori – consentirebbe di superare le criticità elencate ai punti (2), (3) e (4). In relazione alle *proxy* da introdurre nel modello quali determinanti dell'economia sommersa, viene considerata – in modo coerente con quanto discusso nell'introduzione circa le diverse attività che ricorrono all'uso del contante per evitare la tracciabilità – una definizione più ampia della stessa che si estende a tutta l'economia non osservata, comprensiva quindi della produzione illegale (economia criminale) oltre che delle transazioni associate a comportamenti di evasione fiscale. La stima della domanda di circolante dovrebbe includere di conseguenza un set più ampio di variabili esplicative, al fine di riuscire a controllare anche l'effetto sull'utilizzo del contante derivante dalla presenza di attività illegali. Secondo gli autori, inoltre, entrambe le critiche enunciate al punto (3) non avrebbero ragion d'essere nel caso dell'euro. La prima, alla luce del fatto che la moneta unica sarebbe stata adottata diffusamente come strumento di pagamento internazionale solo dopo una lunga fase di scetticismo dei mercati. La seconda, perché non sarebbe verosimile assumere una riduzione dei depositi bancari nel periodo successivo al 2002, dato lo sviluppo delle transazioni elettroniche e di una fase non segnata da conflitti bellici in Europa nel periodo di riferimento. Infine, l'anno dell'introduzione dell'euro (2002) potrebbe essere scelto come anno base, senza che questa scelta abbia caratteristiche di arbitrarietà, vista la difficoltà che le transazioni nel settore sommerso in quel-

l'anno siano avvenute in euro, ossia in concomitanza con la procedura di “*cash changeover*”¹⁰.

L'aspetto di maggiore interesse di Ferwerda, Deleanu e Unger (2010) consiste comunque nell'estensione del modello di Tanzi su un arco temporale più lungo, con l'obiettivo di testare la robustezza della relazione di causalità tra pressione fiscale e dimensione del settore sommerso¹¹. Anziché consentire di superare le critiche avanzate da Schneider e Enste (2002), i risultati nel modello rivisitato vanno invece nella direzione di sollevare maggiori dubbi sul nesso causale tra i due fenomeni. Infatti, la relazione perde di significatività statistica a partire dall'inizio degli anni '70¹². Alla luce di tale risultato, gli autori propongono delle varianti del modello al fine di incorporare – in sostituzione dell'aliquota dell'imposta sul reddito – almeno una *proxy* alternativa tra quelle proposte dalla letteratura in materia. In particolare, le variabili considerate sono il tasso di disoccupazione, diversi indicatori di spesa pubblica e alcune misure del grado di istruzione e di disuguaglianza sociale, controllando anche per la diffusione delle attività criminali. I risultati vengono giudicati dagli stessi autori insoddisfacenti, dal momento che nessuna delle *proxy* utilizzate riesce a spiegare la diffusione dell'economia sommersa in termini di eccesso di domanda di circolante rispetto al suo utilizzo regolare. Poiché gli autori evidenziano la necessità per la ricerca futura di individuare variabili-causa con maggiore capacità esplicativa del comportamento di partecipazione al sommerso, si può dedurre che, ai fini di una buona misurazione del fenomeno, ancora più auspicabile sarebbe l'individuazione di indicatori che siano in grado di fornire una misura *diretta* del verificarsi degli episodi di evasione. Seguendo tale approccio, infatti, il problema della corretta identificazione di tutte le determinanti rilevanti del fenomeno verrebbe superato (sebbene tale identificazione sia importante in fase di definizione delle politiche di contrasto ed è quindi bene che tale aspetto

¹⁰ Gli autori non chiariscono bene questo punto. Probabilmente la motivazione sottostante alla difficoltà di effettuare transazioni irregolari in Euro era di tipo pratico, legata all'avvio nel 2002 della procedura di “*cash changeover*” che richiedeva di fatto di convertire la vecchia moneta legale presso il sistema bancario, con evidente rischio di tracciabilità, o di far emergere operazioni sospette, soprattutto in presenza di un elevato ammontare di contante depositato.

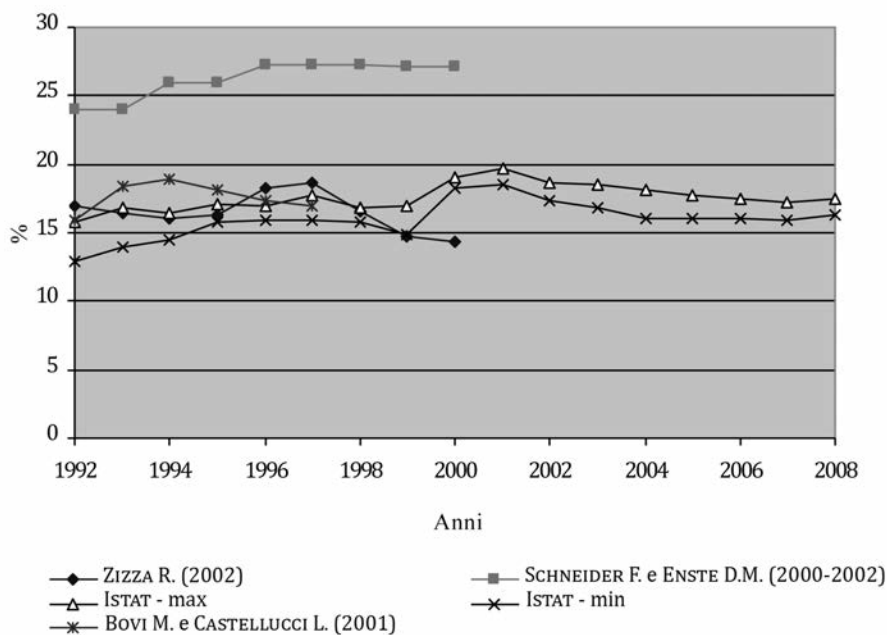
¹¹ Le stime originarie di TANZI V. (1980, 1983), che coprivano il periodo 1930-1980, vengono replicate dagli autori includendo anche gli anni successivi fino al 2006.

¹² Gli autori stimano un modello lineare in serie storica ripreso dall'equazione di TANZI V. (1983). Tuttavia, il modello OLS originario è corretto facendo riferimento alle differenze prime delle variabili (rispetto al periodo precedente) anziché ai livelli assoluti, al fine di rimuovere il problema di “radici unitarie” presente nelle serie temporali quando la variabile dipendente e quella indipendente hanno un medesimo *trend* e la relazione positiva che ne verrebbe fuori sarebbe “spuria”.

venga riconsiderato in un secondo stadio di analisi), dal momento che non si farebbe più riferimento alle motivazioni sottostanti alla scelta di evadere bensì all'esito finale di tale processo decisionale, con l'obiettivo di stimarne in modo preciso la dimensione economica attraverso lo studio della relazione fra gli episodi di evasione e l'utilizzo del contante nelle transazioni. Come sarà chiarito nel paragrafo 3, il nostro contributo va proprio in questa direzione.

GRAF. 1

STIME DELL'ECONOMIA SOMMERSA IN ITALIA
(in % del Pil)



Nota: per maggiore chiarezza espositiva il grafico si riferisce solo al periodo 1992-2008, sebbene in alcuni dei lavori citati si considerino anche anni antecedenti al 1992. La fonte delle stime Istat per gli anni 1992-1999 è ISTAT (2003), mentre la serie 2000-2008 è stata estratta da ISTAT (2010). I risultati riportati per BOVI M. e CASTELLUCCI L. (2001) riguardano solo le stime ottenute con l'approccio *standard* di azzeramento del valore delle variabili che determinano il sommerso e non anche l'approccio alternativo in cui viene utilizzato il valore minimo di tali variabili.

2.2 Le stime del sommerso disponibili per l'Italia

In questo paragrafo si discutono le stime dell'economia sommersa nel nostro paese prodotte in letteratura utilizzando il CDA e si confrontano con le stime ufficiali dell'Istat (basate però sul metodo microeconomico dell'*input di lavoro*). Il

Grafico 1 riporta tali stime per gli anni 1992-2008. Accanto agli studi condotti da Istat (2003, 2010), si riportano i risultati di Bovi e Castellucci (2001); Zizza (2002) e Schneider e Enste (2000, 2002) che rappresentano, di fatto, gli unici lavori disponibili sul tema. Di seguito vengono brevemente confrontati i principali risultati ottenuti, discutendo anche il lavoro di Chiarini e Marzano (2004) che, rispetto agli altri, ha un taglio metodologico mirante ad evidenziare alcuni aspetti critici del CDA.

Le stime ufficiali Istat. L'Istat fornisce stime relative al lavoro e al valore aggiunto generato dal sommerso economico, definito come quella parte di economia non osservata che «deriva dall'attività di produzione di beni e di servizi che, pur essendo di natura legale, sfugge all'osservazione diretta, in quanto connessa al fenomeno della frode fiscale e contributiva» (Istat, 2010, p. 1). Le stime Istat, dunque, prescindono dall'economia illegale, a causa «dell'eccessiva difficoltà nel calcolare tale aggregato e la conseguente incertezza della stima, che renderebbe poco confrontabili i dati dei vari paesi». Gli ultimi dati resi disponibili sono relativi al periodo 2000-2008. Il valore aggiunto prodotto nell'area del sommerso economico viene stimato nel 2008 tra un minimo del 16,3% (pari a circa 255 miliardi di euro) e un massimo del 17,5% del Pil (circa 275 miliardi di euro). Nel 2000, la percentuale minima e massima erano pari rispettivamente al 18,2% e al 19,1% del Pil (rispettivamente pari a circa 217 miliardi e a 228 miliardi di euro). Fra il 2000 e il 2008 l'andamento è alterno, con un picco nel 2001 (18,5-19,7%), una progressiva discesa fino al 2007 (15,9-17,2%) e una ripresa a partire dal 2008.

Rimandando alla nota metodologica dell'Istat (2010, p. 3) per maggiori dettagli, è utile ricordare che l'Istat ottiene tali stime applicando la metodologia dell'*input di lavoro* (cfr. anche Zizza, 2002; Palmieri, 2004). Come base di partenza delle stime si utilizza il confronto tra i dati sulla domanda di occupazione da parte delle imprese e quelli dell'offerta di lavoro delle famiglie. A tali stime dell'*input di lavoro* si applicano appositi coefficienti di produttività. Questa prima dimensione viene integrata con altre due riconducibili alla sottodichiarazione del fatturato e al bilanciamento tra dati della domanda e della produzione. Per ciascun anno vengono quindi fornite due valutazioni, corrispondenti agli estremi di un intervallo all'interno del quale si ritiene sia compreso il valore "esatto" del fenomeno: una ipotesi minima (valore aggiunto *certamente* ascrivibile al sommerso economico) e una massima (che include anche il sommerso statistico).

Le stime di Zizza (2002). Zizza stima la dimensione e l'evoluzione del sommerso in Italia nei periodi 1984-2000 e 1992-2000. Nel primo caso viene im-

plementato il CDA, nel secondo caso il *model approach*¹³. Nell'applicazione del CDA al caso italiano, ci si prefigge l'obiettivo di estendere il modello *à la* Tanzi, introducendo una variabile che controlli l'impatto delle attività criminali sulla domanda di contante. Alla luce della nostra discussione sulla diverse componenti dell'economia non osservata (cfr. nota 2), tale estensione è finalizzata ad ottenere risultati confrontabili con le stime Istat, le quali, come si è detto sopra, escludono dalla misurazione la produzione di tipo illegale.

La stima dell'equazione del circolante viene operata su un numero ridotto di osservazioni (68, dati trimestrali 1984-2000), utilizzando come variabile dipendente il rapporto tra *stock* di circolante e Pil (*proxy* del volume delle transazioni ufficiali) e includendo come regressori la pressione fiscale, il tasso di interesse e un indicatore di criminalità. La variabile di pressione fiscale è calcolata come rapporto tra la somma di imposte dirette e contributi sociali e il Pil. Il tasso di interesse è quello sui conti correnti passivi con più di 20 milioni di lire. L'effetto sul circolante dell'economia illegale è catturato da un indicatore di criminalità diffusa (definito dall'autrice come il rapporto di furti e rapine meno gravi sul totale dei delitti), elaborato dall'Istat sulla base di dati forniti dal Ministero dell'Interno.

Applicando in modo *standard* il CDA, la serie del valore aggiunto sommerso viene ottenuta azzerando il valore della pressione fiscale (unica determinante del sommerso) e moltiplicando l'eccesso di domanda di circolante così calcolato per la velocità di circolazione della moneta del periodo corrispondente, in questo caso stimata attribuendo arbitrariamente a ciascuno degli anni nel periodo 1984-2000 il ruolo di anno "base" senza sommerso. L'autrice sostiene che i risultati ottenuti fanno emergere una coerenza con le stime Istat per gli anni 1992-1998 e che l'inclusione dell'indicatore di criminalità, depurando le stime dagli effetti delle attività illegali, rimuove la principale causa di discrepanza tra il CDA (Schneider e Enste, 2000 e 2002, cfr. Graf. 1) e le valutazioni ufficiali. Tuttavia, a nostro avviso, la variabile adottata per "misurare" il crimine – che cattura la diffusione di reati comuni quali furti e rapine, per i quali non è individuabile una transazione fra domanda e offerta di un bene o di un servizio, quindi nemmeno un pagamento corrispondente che motivi l'impiego di circolante (Smith, 1994) – non è soddisfacente e l'effetto atteso sulla domanda di contante risulta anche ambiguo, dal

¹³ Si tratta di una versione meno sofisticata della metodologia MIMIC discussa alla nota 7, basata sulla teoria statistica delle variabili latenti, per misurare le quali si ricorre all'analisi fattoriale. Rispetto alla tecnica MIMIC, il *model approach* è un modello di analisi fattoriale di natura più esplorativa, in cui non viene fatta una distinzione *a priori* tra cause e indicatori del fenomeno studiato.

momento che il maggior rischio di subire un furto o una rapina può anche *ridurre* la propensione a utilizzare o a detenere scorte monetarie.

Le stime di Bovi e Castellucci (2001). Bovi e Castellucci applicano il CDA all'economia italiana riprendendo la specificazione originaria di Tanzi e utilizzando come variabile dipendente il rapporto tra il circolante detenuto dalle famiglie e l'aggregato *M2*. La domanda di circolante così definita viene messa in relazione con la tassazione diretta effettiva (rapporto tra entrate fiscali e valore aggiunto), con la quota dei salari pagati in settori "sensibili" (agricoltura, edilizia), con il tasso di interesse sui depositi bancari e con il Pil *pro-capite*. Una volta stimato il sommerso fiscale su base nazionale, un interessante approfondimento del lavoro consiste nel tentativo di calcolare il sommerso a livello regionale, ricorrendo tuttavia, per mancanza di disponibilità di dati, ad un approccio alternativo basato sulla stima della domanda di lavoro irregolare anziché sulla domanda di circolante.

Con riferimento alla stima con il CDA del sommerso fiscale a livello nazionale, gli autori propongono – in aggiunta all'approccio *standard*, che richiede di azzerare le variabili rilevanti per misurare la domanda di contante utilizzato per le transazioni irregolari – un modello alternativo che vincola le variabili rilevanti per il sommerso ad assumere un valore minimo, ritenuto più realistico rispetto al valore nullo. Ciò farebbe emergere un'incidenza del sommerso in Italia assai inferiore (7,9% nel 1997) rispetto a quella presentata negli altri studi esistenti sull'Italia, sebbene sia opportuno procedere con cautela nell'effettuare comparazioni in presenza di diverse ipotesi metodologiche. Per quanto riguarda i risultati a livello regionale, gli autori rilevano una maggiore incidenza del sommerso al Sud rispetto al Centro-Nord. Tali stime, senza dubbio interessanti considerata l'assenza di evidenza empirica disaggregata sul territorio, si discostano tuttavia dall'applicazione del CDA finora discussa e potrebbero non catturare in modo coerente ed esaustivo il fenomeno del sommerso (fiscale e criminale) a livello locale.

La critica di Chiarini e Marzano (2004). Il lavoro di Chiarini e Marzano si inserisce in questa letteratura con una prospettiva di carattere prettamente metodologico, proponendo una verifica della robustezza del CDA attraverso la stima di un sistema multiequazionale dinamico cointegrato (*Error Correction Model Vettoriale*), al fine di superare le ipotesi di esogeneità e stazionarietà delle variabili esplicative sottostanti ai modelli *time series* uniequazionali utilizzati sull'esempio di Tanzi. Le determinanti della domanda di circolante (rapportato anche in questo caso a *M2*) diventano quindi endogene al modello e fra queste viene incluso il Pil *pro-capite*, il livello di tassazione (diretta e contributiva, in rapporto al Pil) e l'in-

cidenza dei redditi da lavoro dipendente. I risultati che emergono appaiono molto instabili e sensibili al tipo di variabili utilizzate e al periodo iniziale considerato. Secondo gli autori, sono quindi confermate le perplessità avanzate circa la robustezza della metodologia CDA una volta rimosse alcune restrizioni sottostanti.

3. - Metodologia di analisi

Dopo aver brevemente discusso la letteratura in materia sul caso italiano, in questo paragrafo presentiamo prima le innovazioni introdotte rispetto al CDA tradizionale e a seguire le stime del modello di domanda di circolante ottenute applicando il nuovo approccio. L'aspetto cruciale è la definizione dell'utilizzo di contante attraverso un indicatore di *flusso* anziché di *stock*, sulla scorta dell'idea che per misurare in modo corretto la produzione sommersa occorra identificare *direttamente* le transazioni che rientrano nell'economia non osservata (di natura sia legale che criminale).

3.1 *Il modello di domanda di pagamenti in contanti*

Come anticipato, l'equazione adottata per stimare l'economia sommersa utilizza come variabile dipendente un indicatore di *flusso* di contanti anziché di *stock*, che è la misura a cui ha fatto finora riferimento tutta la letteratura basata sul CDA *à la* Tanzi (1980, 1983). Tale innovazione metodologica ha il pregio di fornire una valutazione *diretta* del valore delle transazioni, consentendo così di superare una delle critiche fondamentali al CDA avanzate da Schneider e Enste (2000, 2002), che riguarda l'utilizzo dell'equazione di Fisher e la necessità di ipotizzare un anno base senza sommerso e identica velocità di circolazione della moneta fra economia ufficiale e irregolare. In particolare, la variabile adottata (*CASH*) misura il rapporto fra il valore dei prelievi totali di contanti da conti correnti bancari e il valore totale dei pagamenti non in contanti e si può interpretare come la quantità domandata di pagamenti riferiti a transazioni non tracciabili per ogni euro di pagamenti tracciabili (ovvero eseguiti tramite assegno, bonifico e strumenti elettronici quali bancomat e carta di credito).

Più in dettaglio, l'utilizzo dello *stock* di banconote in circolazione per spiegare la domanda di contante da parte degli operatori economici, pur in presenza di statistiche affidabili, può risultare altamente impreciso per diversi motivi:

- a) non è agevole quantificare l'ammontare di circolante di ciascun paese utilizzato al di fuori dei confini nazionali;

- b) all'interno di ciascun paese o area geografica, una certa quantità di moneta viene detenuta per finalità diverse da quella transattiva (es. tesaurizzazione o riserva di valore);
- c) la velocità di circolazione della moneta viene ipotizzata costante rispetto ai diversi aggregati macroeconomici.

Un indicatore di domanda di circolante costruito come *stock* di banconote in circolazione rispetto al Pil o rispetto ai saldi monetari più ampi (es. $M1$ o $M2$) non tiene conto di fenomeni compensativi tra movente transattivo e non transattivo della domanda di contante, delle ricomposizioni di portafoglio (o di *buffer* di liquidità), oltre che tra contante utilizzato nell'economia regolare e in quella irregolare. A tal proposito, la Banca Centrale Europea (Bce, 2008) rileva che, in occasione del c.d. "Euro cash changeover" lo *stock* di banconote in circolazione è aumentato (rispetto al Pil, a $M1$ e a $M2$) più di quanto la precedente circolazione delle valute nazionali avrebbe suggerito. Tale aumento è peraltro proseguito negli anni successivi, in presenza di tassi di interesse moderati e attese di bassa inflazione, senza contare che una quota stimata fino al 20% delle banconote in circolazione è detenuta al di fuori dell'area dell'Euro. Diventa quindi difficile (o quantomeno fortemente distorsivo) stimare la componente di contante detenuto per regolare gli scambi nell'ambito del sommerso. L'incremento tendenziale sia del rapporto circolante/Pil sia del rapporto circolante/ $M1$ (o circolante/ $M2$) negli ultimi dieci anni, così come registrato nelle statistiche dell'Eurosistema, appare inoltre poco coerente con il contestuale incremento dei pagamenti elettronici nei paesi europei e la crescita del grado di finanziarizzazione e di innovazione tecnologica dell'economia.

Da qui la necessità di selezionare una variabile monetaria più direttamente connessa con il movente transattivo e meno influenzata dai meccanismi compensativi discussi sopra. L'indicatore utilizzato in questo lavoro, rappresentato dall'ammontare dei prelievi in rapporto ai pagamenti basati su strumenti alternativi al contante, appare più adatto per una serie di ragioni. Innanzitutto, il flusso dei prelievi consente di cogliere più da vicino il movente transattivo. È ragionevole infatti supporre che i prelievi di piccolo taglio effettuati direttamente allo sportello tradizionale o con carta bancomat tramite sportello automatico (ATM) siano sostanzialmente guidati da questa finalità: il contante viene cioè prelevato per concludere scambi, regolari o irregolari, più che per essere tesaurizzato. È altrettanto sensato attendersi che questi pagamenti al dettaglio vengano reintrodotti nel sistema bancario attraverso un unico ciclo di scambio, così che la velocità di circolazione sia vicina all'unità e sostanzialmente stabile. Oltre ai prelievi di piccolo

taglio, vi sono poi prelievi – effettuati soprattutto presso gli sportelli tradizionali – di elevato ammontare, che potrebbero ragionevolmente essere utilizzati dagli operatori economici anche per più cicli di scambio prima di venire reintrodotti nel circuito bancario. Tuttavia, qualora questa componente registri un tasso di rotazione superiore all'unità o anomalo (rispetto, ad es., a quello dei prelievi da ATM), esso impatterebbe anche sulla velocità di circolazione degli strumenti alternativi. Per esempio, se in una determinata area il contante prelevato dovesse circolare più volte tra gli operatori prima di rientrare nel circuito bancario, è evidente che – specularmente – in quella determinata area si utilizzeranno di meno gli strumenti alternativi al contante per gli scambi. L'indicatore di domanda di contante utilizzato in questo studio aumenta (o diminuisce) quindi sia all'aumentare (o al diminuire) della domanda di prelievi di contante, sia al diminuire (o all'aumentare) dell'ammontare di scambi regolati con strumenti alternativi, i quali risentono a loro volta del grado di finanziarizzazione e dell'eventuale variazione nella velocità di circolazione del contante in un determinato territorio. Tale fenomeno non è correttamente colto da grandezze monetarie di tipo *stock*, per le quali non è possibile catturare l'effetto dei flussi di pagamento intermediati. In Ardizzi, Petraglia, Piacenza e Turati (2011, 2013) si mettono a confronto nel tempo, da una parte, l'andamento del rapporto tra banconote in circolazione e *M1* e, dall'altra, il flusso di prelievi di contante rispetto all'ammontare dei pagamenti regolati con strumenti alternativi al contante. Il primo rapporto è sostanzialmente crescente e inizia a diminuire in occasione della crisi economico-finanziaria, lasciando quindi supporre una stretta connessione con la congiuntura economica. Il secondo rapporto, basato sui flussi, pur tenendo conto del ciclo economico, appare invece decrescente nel tempo, in linea con lo sviluppo dei pagamenti elettronici e dell'innovazione finanziaria.

Per quanto riguarda le determinanti di *CASH*, così come è stata definita, distinguiamo tre gruppi di fattori che individuano una componente di tipo *strutturale*, riferita cioè a transazioni che avvengono nell'economia regolare, e due componenti relative all'economia *non osservata*, ovvero imputabili a transazioni non tracciabili e motivate sia da comportamenti di *evasione fiscale* (produzione sommersa) sia da *attività criminali* (produzione illegale). Nel seguito viene fornita una descrizione dettagliata delle variabili incluse in ciascuna di queste tre componenti. L'Appendice riporta le informazioni sulle diverse fonti dei dati (Tav. 1) e le statistiche descrittive (Tav. 2), sia per l'intero campione che a livello disaggregato per macro-area geografica (Centro-Nord e Sud).

3.1.1 *La componente strutturale*

Si assume che la componente strutturale dell'equazione che definisce la domanda di pagamenti in contanti dipenda da alcune determinanti fondamentali segnalate dalla letteratura sui modelli per la domanda di circolante, non necessariamente nel contesto della stima del sommerso (es. Goodhart e Krueger, 2001). In particolare, sono stati individuati i seguenti quattro fattori:

- il livello di sviluppo economico;
- il grado di diffusione territoriale dell'offerta bancaria;
- la tecnologia dei pagamenti;
- il tasso di interesse.

Il livello di sviluppo dell'economia viene misurato dal Pil provinciale per abitante (*YPC*). Come suggerito da diversi autori (cfr. Schneider e Enste, 2000 e Schneider, 2010), il segno atteso è negativo: più elevato è il tenore di vita, minore dovrebbe essere l'utilizzo di contante nel territorio di riferimento (e più elevata la *domanda* di strumenti di pagamento evoluti). In particolare, seguendo Bovi e Castellucci (2001), con questa variabile si dovrebbe cogliere un effetto legato al grado di istruzione – generale e soprattutto “finanziaria” (*financial literacy*) – la cui rilevanza è stata evidenziata in un rapporto della World Bank (Schiff e Ozden, 2005) e ripresa poi nel recente studio di Ferwerda, Deleanu e Unger (2010): una maggiore istruzione comporta un minor ricorso al contante, dal momento che gli individui sono più informati e hanno maggiore fiducia nell'utilizzo di strumenti di pagamento più evoluti (elettronici e non)¹⁴.

La *proxy* adottata per misurare la diffusione territoriale dell'offerta bancaria è rappresentata dal numero di conti correnti bancari *pro-capite* (*BANK*). Anche in questo caso il segno atteso dell'impatto esercitato su *CASH* è negativo: la spiegazione è abbastanza intuitiva, considerando che una maggiore penetrazione nell'utilizzo dei conti correnti (che riflette il grado di “bancarizzazione” dell'economia) tende a ridurre la necessità di ricorrere al prelievo di contanti per regolare le transazioni. L'inserimento di tale controllo nel modello assume particolare importanza per il contesto italiano dove la diffusione dei conti correnti bancari è fortemente sperequata sul territorio e fra comparti produttivi, con un minor grado di “ban-

¹⁴ È bene precisare che, nel caso dell'Italia, *YPC* include sia la rilevazione delle transazioni ufficiali che una stima della produzione sommersa effettuata dall'Istat (cfr. paragrafo 2.2). È ragionevole quindi attendersi che l'impatto di tale variabile risulti attenuato (ovvero minore in valore assoluto) rispetto alle stime che si otterrebbero con un modello in cui la misura del Pil si riferisce solo all'economia regolare, a causa della presenza della componente di transazioni sommerse che spinge nella direzione di accrescere l'utilizzo del contante.

carizzazione” tipicamente osservato nelle province del Sud (la Tav. 2 riporta un numero medio di conti correnti *pro-capite* pari a 0,347 contro 0,684 per le province del Centro-Nord) e in settori come l’edilizia e l’agricoltura, già tradizionalmente arretrati (i salari sono pagati in contante settimanalmente) e ancor di più negli ultimi anni a causa del fenomeno dell’immigrazione¹⁵.

Il ruolo svolto dalla tecnologia dei pagamenti, in particolare dall’*offerta* di strumenti elettronici, è stato enfatizzato in vari studi (es. Drehmann e Goodhart, 2000; Goodhart e Krueger, 2001; Schneider, 2009) ed è considerato nel nostro modello inserendo una variabile che misura il valore in rapporto al Pil delle transazioni regolate con pagamenti elettronici (*ELECTRO*). Il segno atteso è di nuovo negativo, poiché maggiore è la quota di transazioni che passano attraverso strumenti elettronici (pagamenti tramite POS e *internet banking*), minore sarà quella che viene regolata tramite contante. Come per *BANK*, anche in questo caso il controllo consente di catturare differenze significative nella diffusione di sistemi elettronici di pagamento fra Centro-Nord e Sud (cfr. Tav. 2).

Infine, per quanto riguarda l’inclusione nel modello del tasso di interesse, l’equazione strutturale viene stimata in due versioni – nel seguito, MODELLO 1 e MODELLO 2 – a seconda della presenza o meno del costo-opportunità del denaro (qui misurato dal tasso di interesse passivo sui depositi bancari, *INT*) nella spiegazione dei comportamenti di impiego del contante per regolare le transazioni. Concettualmente, il modello trattato nel presente lavoro, diversamente da quelli utilizzati nella letteratura precedente, si focalizza sui flussi e non sugli *stock* di attività monetarie. La variabile *CASH* non è quindi la *scorta* di moneta legale detenuta dagli operatori per la quale è importante valutare il costo-opportunità rispetto alle attività fruttifere. Il tasso di interesse, infatti, che normalmente ha un effetto negativo sulla domanda di scorte monetarie, potrebbe influire solo in misura contenuta sulle decisioni riguardanti l’ammontare di contante da prelevare per regolare le transazioni, rispetto agli altri fattori strutturali discussi sopra¹⁶; tassi di interesse

¹⁵ Ringraziamo un *referee* anonimo per aver evidenziato la rilevanza di questa variabile ai fini di una corretta stima della dimensione del sommerso e della sua ripartizione fra le diverse aree territoriali.

¹⁶ Al riguardo merita evidenziare che diversi studi sul ruolo dell’innovazione tecnologica nella domanda di contante da parte delle famiglie italiane (cfr. ARDIZZI G. e TRESOLDI C., 2003; LIPPI F. e SECCHI A., 2008; ALVAREZ F. e LIPPI F., 2009) affermano che – sebbene i risultati empirici sull’elasticità della domanda di moneta al tasso di interesse siano in parte coerenti con la teoria monetaria *à la* Baumol - Tobin (segno negativo) – un miglioramento delle tecnologie transattive (es. ATM, POS, *internet banking*) può ridurre sensibilmente (se non addirittura cancellare) l’impatto del tasso di interesse sulla domanda di moneta per acquisti.

elevati potrebbero addirittura incidere positivamente sui flussi di contante scambiati, ad esempio perché spingono a ricorrere a forme alternative di reperimento dei fondi al di fuori del canale bancario. Tuttavia, non si può escludere a priori che il tasso di interesse sui depositi bancari svolga un qualche ruolo anche nella propensione al prelievo di contante rispetto all'utilizzo di strumenti di pagamento alternativi. Per tale motivo verranno stimate entrambe le specificazioni dell'equazione strutturale, con e senza l'inclusione del tasso d'interesse:

$$(1) \text{CASH}_{it} = \alpha_0 + \alpha_1 \text{YPC}_{it} + \alpha_2 \text{BANK}_{it} + \alpha_3 \text{ELECTRO}_{it} + \alpha_4 \text{INT}_{it} + \varepsilon_{it}$$

Per stimare la domanda di pagamenti in contanti imputabile a transazioni rientranti nell'economia non osservata, l'equazione strutturale (1) è integrata con due diversi gruppi di fattori esplicativi, al fine di misurare separatamente le componenti di produzione sommersa e di produzione illegale. In accordo con la letteratura precedente basata sul CDA, il segno atteso per i coefficienti di tutte le variabili aggiunte è positivo, ad indicare un eccesso di domanda di prelievi di contanti rispetto a quella necessaria a soddisfare le necessità di pagamenti determinate solo da fattori strutturali, ovvero in assenza di transazioni irregolari associate a fenomeni di evasione e di crimine.

3.1.2 *La componente di produzione sommersa (evasione)*

Data l'assenza di informazioni complete sulla pressione fiscale effettiva disaggregata a livello provinciale e tenuto conto dell'argomentazione avanzata in diversi lavori per cui esistono anche altri fattori – oltre al livello delle imposte e dei contributi – che possono influenzare la scelta di evadere (es. regolamentazione del mercato del lavoro, *tax morale* dei cittadini, efficienza della Pubblica Amministrazione, ecc.¹⁷), ai fini di una corretta stima della dimensione del sommerso si è deciso di guardare direttamente all'esito finale di tale processo decisionale (anziché cercare di individuarne tutte le motivazioni sottostanti) – aggirando così il problema dell'incompleta specificazione del modello – e di analizzare la relazione

¹⁷ Riguardo l'importanza di questi diversi aspetti nel determinare la dimensione dell'economia sommersa, si vedano i lavori di FRIEDMAN E., JOHNSON S., KAUFMANN D. e ZOIDO-LOBATÓN P. (2000); SCHNEIDER F. e ENSTE D.H. (2000, 2002); FELD L. e FREY B.S. (2007); DREHER A., KOTSOGIANNIS C. e MCCORRISTON S. (2009); TORGLER B. e SCHNEIDER F. (2009); FERWERDA J., DELEANU I. e UNGER B. (2010); DREHER A. e SCHNEIDER F. (2010) e SCHNEIDER F. (2010).

fra il numero rilevato di episodi di evasione e l'utilizzo del contante nelle transazioni.¹⁸ A tal fine vengono utilizzate le informazioni relative ad alcune attività di controllo della Guardia di Finanza mirate al contrasto dell'evasione fiscale.

In particolare, sono state definite due diverse variabili *proxy* per la diffusione di attività legate alla produzione sommersa a livello provinciale. Un primo indicatore è rappresentato dal numero di verifiche fiscali specifiche (*EVAS1*), che riguardano un controllo effettuato presso un'impresa sulla base di una qualche informazione *ex-ante* su irregolarità avvenute nell'ambito di particolari operazioni (es. pagamento stipendi) e rispetto ad una singola voce (es. imposte dirette o contributi). Tale variabile è stata standardizzata per il valore medio assunto su tutto il campione – in modo da ricavare una misura di intensità dell'evasione a livello provinciale – e pesata poi per un indice di concentrazione del Pil (calcolato come rapporto fra Pil provinciale e valore medio campionario): ciò consente di effettuare una comparazione del fenomeno fra province che tenga conto del livello relativo di sviluppo dell'attività economica, evitando così di trarre indicazioni fuorvianti circa una maggiore diffusione dell'evasione a seguito della rilevazione di un numero più elevato di verifiche rispetto alla media.¹⁹

La seconda variabile inserita nel modello riflette le irregolarità riscontrate dalla Guardia di Finanza presso gli esercizi commerciali (*EVAS2*) ed è costruita come rapporto tra il numero dei controlli con rilievi effettuati su registratori di cassa e ricevute fiscali e il numero di POS esistenti nella provincia, pesando anche in

¹⁸ Si noti che, per completezza dell'analisi, sono state anche effettuate stime di un modello contenente una variabile di pressione fiscale effettiva, misurata come rapporto tra il gettito da addizionale Irpef regionale e comunale e il reddito imponibile regionale (la rilevazione dell'informazione è quindi costante fra le diverse province all'interno di ogni regione e varia fra le regioni e negli anni). I risultati, ottenuti sia sostituendo la pressione fiscale agli indicatori di evasione rilevata dalla Guardia di Finanza – in modo da riprodurre il CDA tradizionale – sia includendo la nuova variabile in un modello contenente anche gli indicatori di evasione, non hanno tuttavia fornito evidenza a supporto dell'utilizzo di tale variabile in sostituzione o in aggiunta agli indicatori diretti di diffusione dell'evasione: il coefficiente stimato presenta sempre un segno negativo e quindi non cattura la propensione ad evadere collegata alla maggiore pressione fiscale (se così fosse l'effetto osservato dovrebbe essere positivo), bensì qualche altro fenomeno. Il sospetto è che la variabile, riflettendo la pressione fiscale *effettiva* e non quella *legale*, sia affetta da problemi di endogeneità e non risulta quindi appropriata ai fini della stima dell'economia sommersa.

¹⁹ In ARDIZZI G., PETRAGLIA C., PIACENZA M. e TURATI G. (2013) è stato utilizzato anche un indicatore alternativo del livello di sviluppo economico, basato sulla concentrazione dell'occupazione, come peso per le variabili di evasione e le stime ottenute per l'economia sommersa sono sostanzialmente confermate.

questo caso per l'indice di concentrazione provinciale del Pil per le stesse ragioni discusse sopra. L'idea che giustifica la standardizzazione rispetto al numero dei POS è che vi siano differenze significative nelle dotazione di apparecchiature tra le diverse aree del paese (in particolare, fra Centro-Nord e Sud) e questo va ad incidere sull'opportunità di evasione (minore dove il numero di POS è più elevato); la misura degli episodi di frode fiscale per unità di POS dovrebbe aiutare a controllare queste differenze nelle opportunità e a quantificare meglio l'estensione del fenomeno.

Si noti che l'inclusione nel modello di entrambe le variabili relative all'attività della Guardia di Finanza è motivata dal fatto che il primo indicatore fa riferimento ad ispezioni che possono riguardare qualunque irregolarità fiscale *presupposta* (evasione di imposte dirette, indirette o contributi) relativa a qualsiasi operazione effettuata nell'ambito di ogni tipo d'impresa (industriale, commerciale o agricola), mentre il secondo indicatore rileva con *certezza* solo le irregolarità nell'ambito delle operazioni di vendita (evasione Iva e imposte dirette) svolte dagli esercizi commerciali. Pertanto, l'utilizzo congiunto dovrebbe consentire una valutazione più esaustiva dell'estensione dell'economia sommersa.

3.1.2 *La componente di produzione illegale (crimine)*

L'indicatore di criminalità (*CRIME*) utilizzato come determinante della domanda di pagamenti in contanti connessa ad attività illegali è definito dalla quota (sul totale dei delitti denunciati dalle Forze di Polizia all'Autorità Giudiziaria) dei delitti connessi a violazioni della normativa sugli stupefacenti e sullo sfruttamento e il favoreggiamento della prostituzione, rapportata ancora una volta alla concentrazione provinciale del Pil. Diversamente dagli studi precedenti che hanno evidenziato l'importanza di controllare questo aspetto nell'analisi della domanda di circolante, dato che il crimine adotta tipicamente il contante come mezzo di pagamento (es. Rogoff, 1998; Drehmann e Goodhart, 2000; Zizza, 2002; Ferwerda, Deleanu e Unger 2010), qui non vengono considerati quei reati con possibili effetti ambigui sulla dimensione del prelievo di contanti (ad esempio, i furti e le rapine, che potrebbero anche avere un impatto negativo, a causa della maggiore rischiosità dell'ambiente in cui si svolgono le transazioni).

La selezione delle variabili idonee a stimare la componente di economia non osservata legata ad attività di natura illegale è un punto importante che merita un breve approfondimento. Il criterio che ha guidato la nostra scelta è stato quello di focalizzare l'attenzione sulle attività criminali per le quali – in accordo alla de-

finizione di economia non osservata discussa sopra (es. Smith, 1994) e a quella di economia criminale elaborata dall'OCSE (2002) – esiste una transazione “di mercato” con un mutuo consenso tra venditore e acquirente, quindi un pagamento presumibilmente in contanti; sono stati pertanto esclusi tutti quei crimini per i quali, in qualche misura, c'è un ricorso alla violenza su cose o persone (es. rapine, furti, estorsioni, ecc.) e manca quindi una transazione consensuale (e il connesso pagamento)²⁰. Tale scelta appare coerente anche alla luce del modello di domanda adottato, che sfrutta le informazioni sui flussi di approvvigionamento di contante “in chiaro” sul territorio da parte della clientela bancaria con un sottostante movente (volontario) transazionale.

Il modello completo di domanda di pagamenti in contanti – che va ad integrare l'equazione strutturale (1) per tenere conto delle componenti associate alla presenza di economia sommersa e criminale – è rappresentato dalla seguente equazione:

$$(2) \quad CASH_{it} = \alpha_0 + \alpha_1 YPC_{it} + \alpha_2 BANK_{it} + \alpha_3 ELECTRO_{it} + \alpha_4 INT_{it} \\ + \alpha_5 EVAS1_{it} + \alpha_6 EVAS2_{it} + \alpha_7 CRIME_{it} + \varepsilon_{it}$$

3.2 Procedura per il calcolo dell'economia sommersa

Il modello introdotto nel paragrafo precedente è stato stimato utilizzando come base dati un *panel* bilanciato di 91 province osservate dal 2005 al 2008, diversamente da tutti gli studi precedenti sull'Italia fondati sul CDA, che hanno condotto un'analisi su serie storica a livello nazionale. Le unità incluse nel campione rappresentano circa il 90% dell'universo delle 103 province italiane e sono quelle per le quali sono disponibili sull'intero arco temporale i dati necessari alla costruzione di tutte le variabili incluse nel modello completo (2), in particolare le informazioni relative al numero di crimini per prostituzione, che in 12 casi

²⁰ Anche il reato di riciclaggio di “denaro sporco” non è considerato nella presente analisi, dal momento che si tratta di un reato conseguente ad altre attività criminali sottostanti che amplifica in modo cumulativo l'impatto delle organizzazioni criminali sull'economia in generale, non solo nell'ambito delle attività illegali. La definizione di riciclaggio implica infatti che i redditi provenienti da un crimine debbano essere “ripuliti” attraverso il canale legale (es. transazioni bancarie), al fine di abbassare la probabilità di essere scoperti, per essere poi successivamente reinvestiti anche in attività regolari. Sull'analisi economica e per una valutazione dell'incidenza dell'attività di riciclaggio in Italia si veda MASCIANDARO D. (2007) e ARDIZZI G., PETRAGLIA C., PIACENZA M., SCHNEIDER F., TURATI G. (2013).

non sono rilevate²¹. Per i dettagli sulle varie fonti utilizzate per predisporre la base dati si rinvia alla Tavola 1 in Appendice.

La struttura *panel* consente di stimare le diverse specificazioni dell'equazione di domanda di transazioni in contanti tenendo conto dell'esistenza di eterogeneità residua non osservata fra le unità del campione, attraverso tecniche econometriche che includano effetti individuali specifici per ogni provincia. A tal fine, è stato utilizzato un modello Tobit *random effects* (Wooldridge, 2002) che ha il pregio – rispetto ad una regressione *standard* con effetti individuali casuali – di considerare la particolare distribuzione della variabile dipendente, caratterizzata da una massa di probabilità concentrata su valori positivi e molto prossimi a zero²². In particolare, la struttura dell'errore delle equazioni (1) e (2) è $\varepsilon_{it} = u_i + e_{it}$, dove u ed e sono rispettivamente gli effetti individuali provinciali e il disturbo statistico *standard*.

Una volta ottenute le stime dei parametri del modello, è stata applicata la procedura originariamente proposta da Tanzi (1983) di “azzeramento” delle variabili esplicative relative all'economia non osservata, al fine di quantificarne la dimensione economica calcolando la differenza tra il valore dei pagamenti in contanti in presenza di episodi di evasione e di crimine e la domanda di transazioni in contanti che si genererebbe nel caso in cui tali fenomeni fossero assenti ($EVAS1 = EVAS2 = CRIME = 0$), ovvero considerando unicamente la parte strutturale del-

²¹ Le province escluse dallo studio per via dell'indisponibilità dei dati su tale attività illegale per almeno uno degli anni analizzati sono: Aosta, Belluno, Caltanissetta, Crotone, Enna, Isernia, Matera, Nuoro, Oristano, Potenza e Vibo Valentia.

²² La Tavola 2 riporta un valore medio per *CASH* pari a 0,11 (mediana = 0,10), con un minimo di 0,01 ed un massimo di 0,24; inoltre il 75% delle osservazioni presenta valori inferiori a 0,14. Si noti anche che, prima di considerare la natura *censored* di *CASH* e adottare la metodologia di stima Tobit, sono state effettuate stime con tecniche *panel* tradizionali *fixed effects* (LSDV) e *random effects* (GLS) e il test di Hausman ha portato a non rifiutare il modello GLS, (i risultati ottenuti con questo modello sono coerenti con le stime Tobit e vengono discussi in ARDIZZI G., PETRAGLIA C., PIACENZA M., TURATI G., 2013). D'altro canto, CAMERON A.C. e TRIVEDI P.K. (2005) mettono in evidenza che un elemento di debolezza che caratterizza il modello LSDV consiste nell'elevato grado di imprecisione con cui vengono stimati i coefficienti nel caso in cui la variabilità nel tempo (*within*) del *panel* sia dominata dalla variabilità cross-sezionale (*between*), che è la situazione che si verifica per tutte le variabili del nostro modello (fatta eccezione per *INT*, si vedano i valori della deviazione *standard* riportati nella Tav. 2). Alla luce di queste considerazioni, come modello finale per la stima è stata adottata una specificazione Tobit *random effects* (l'implementazione di un modello Tobit *fixed effects*, peraltro, richiederebbe l'utilizzo di uno stimatore semiparametrico (cfr. HONORÉ B.E., 1992) al fine di ottenere stime dei parametri non distorte).

l'equazione (2). La decomposizione del valore totale dell'economia non osservata nelle sue due componenti di produzione sommersa e produzione illegale si basa su un calcolo analogo, in cui però l'eccesso di domanda rispetto alla parte strutturale viene stimato considerando solo un fenomeno per volta, ovvero imponendo $CRIME = 0$ per isolare l'impatto associato alla sola attività di evasione e $EVAS1 = EVAS2 = 0$ per stimare invece l'effetto specifico del crimine. Si noti che, per come è stata definita la variabile $CASH$, tutti i valori ottenuti sono espressi in rapporto alla domanda di pagamenti non in contanti; quindi, al fine di avere stime comparabili con gli studi precedenti sull'Italia (cfr. Graf. 1), si rende necessario moltiplicare tali risultati per il denominatore – in modo da ottenere un'espressione in termini assoluti del valore delle transazioni in contanti imputabili all'economia non osservata – e dividerli poi per il dato sul Pil rilevato a livello provinciale.

4. - Risultati

4.1 *Stima della domanda di pagamenti in contanti*

La Tavola 3 mostra i risultati delle stime. Le prime due colonne – MODELLI 1 e 2 – si riferiscono alla stima della sola equazione strutturale (1), rispettivamente senza e con il tasso di interesse incluso fra i regressori (INT), mentre nella terza e quarta colonna vengono presentate le stime del MODELLO completo 2 di domanda di pagamenti in contanti, in cui si tiene conto della presenza di transazioni nell'ambito dell'economia non osservata, solo da evasione (MODELLO 3) o includendo anche il ruolo del crimine (MODELLO 4).

TAV. 3

STIME TOBIT *RANDOM EFFECTS* DELL'EQUAZIONE DI DOMANDA DI PAGAMENTI IN CONTANTI^a

Regressore ^b	MODELLO 1	MODELLO 2	MODELLO 3	MODELLO 4
<i>YPC</i>	-0,040*** (0,004)	-0,035*** (0,004)	-0,030*** (0,003)	-0,026*** (0,004)
<i>BANK</i>	-0,060*** (0,010)	-0,036*** (0,010)	-0,037*** (0,011)	-0,061*** (0,013)
<i>ELECTRO</i>	-0,006*** (0,001)	-0,006*** (0,001)	-0,005*** (0,001)	-0,005*** (0,001)
<i>INT</i>	-	-0,009*** (0,002)	-0,011*** (0,002)	-0,010*** (0,002)
<i>EVAS1</i>	-	-	0,006*** (0,002)	0,006*** (0,002)
<i>EVAS2</i>	-	-	0,027*** (0,005)	0,010* (0,006)
<i>CRIME</i>	-	-	-	0,286*** (0,063)
costante	0,262*** (0,006)	0,240*** (0,006)	0,220*** (0,006)	0,222*** (0,006)
Osservazioni	364	364	364	364
Log-likelihood	934,60	950,59	959,08	963,96
Wald test (χ^2)	1537,98***	1038,21***	1969,51***	2563,29***
σ_u	0,025*** (0,001)	0,022*** (0,001)	0,022*** (0,001)	0,023*** (0,001)
σ_c	0,013*** (0,001)	0,012*** (0,001)	0,012*** (0,000)	0,012*** (0,000)
ρ	0,792 (0,019)	0,770 (0,020)	0,772 (0,019)	0,784 (0,017)

^a Variabile dipendente: *CASH*; MODELLO 1: equazione strutturale (1) senza il tasso d'interesse ($\alpha_4 = 0$); MODELLO 2: equazione strutturale (1) con il tasso d'interesse; MODELLO 3: equazione completa (2) senza l'indicatore di criminalità ($\alpha_7 = 0$); MODELLO 4: equazione completa (2) con l'indicatore di criminalità.

^b Gli errori *standard* sono riportati in parentesi; significatività delle stime: * = 10%, ** = 5%, *** = 1%.

Tutti i coefficienti hanno il segno atteso e sono statisticamente significativi (almeno al 10%). Confrontando le diverse specificazioni, si può osservare innanzitutto come i parametri ($\alpha_1 - \alpha_4$) che definiscono la componente strutturale della domanda risultino abbastanza stabili a perturbazioni del modello derivanti dall'aggiunta di nuovi fattori esplicativi legati allo svolgimento di attività sommerse e criminali: ciò può essere interpretato, da un lato, come un'indicazione della robustezza della specificazione (1) del modello che spiega il ricorso ai pagamenti in

contanti solo per transazioni ufficiali e, dall'altro lato, come evidenza della presenza di variabilità residua in *CASH* che l'equazione (2) aiuta a spiegare, integrando la (1) con gli effetti associati all'economia non osservata²³.

Il test LR conferma che tutte le estensioni della specificazione base (1) contribuiscono a migliorare il *fit* statistico del modello. In particolare, emerge un ruolo significativo del tasso di interesse come determinante della componente strutturale (MODELLO 1 *vs.* MODELLO 2, LR = 31,98^{***}) e, soprattutto, è confermata l'importanza delle attività illegali (traffico di stupefacenti e prostituzione), oltre a quelle di evasione fiscale, ai fini di una corretta valutazione dell'estensione complessiva dell'economia non osservata e della sua componente di produzione sommersa sulla base del CDA: dal confronto fra le stime dei MODELLI 3 e 4 (LR = 9,76^{***}), emerge come differenza più significativa la riduzione del coefficiente α_6 (da 0,027 a 0,010) una volta che si controllano gli effetti del crimine; ciò non sorprende data l'elevata correlazione rilevata sul campione fra *EVAS2* e *CRIME* (0,60) ed evidenza come la non inclusione della seconda variabile possa condurre ad una sovra-stima dell'economia sommersa, a causa dell'attribuzione a tale componente di un eccesso di domanda di pagamenti in contanti in realtà derivante dall'esercizio di attività illegali²⁴. Infine, per tutte le specificazioni il coefficiente ρ , che misura la quota di varianza del residuo totale del modello spiegata dagli effetti individuali (μ) rispetto alla quota spiegata dal disturbo statistico (e), è prossimo a 0,80, evidenziando l'importanza di utilizzare tecniche di stima *panel* che consentano di controllare la presenza di eterogeneità non osservata, imputabile – nel caso di un modello *random effects* – a *shock* casuali idiosincratichi ad ogni provincia del campione.

²³ È interessante notare, in particolare, che la stima del coefficiente α_1 relativo all'impatto di *YPC* – che include anche una stima della produzione sommersa (cfr. nota 14) – non aumenta in valore assoluto (al contrario, si osserva una lieve diminuzione passando dal MODELLO 2 ai MODELLI 3 e 4), come invece ci si aspetterebbe una volta che gli effetti del sommerso vengono valutati separatamente con l'inclusione nel modello degli indicatori *EVAS1* e *EVAS2* (entrambi di segno positivo). Da ciò è possibile dedurre che l'impatto esercitato dal livello di sviluppo economico sulla domanda di pagamenti in contanti è guidato principalmente dalla componente di produzione legata alle transazioni regolari.

²⁴ Si pensi, ad esempio, ai casi in cui certe attività produttive legali (di ristorazione o intrattenimento) vengono svolte congiuntamente ad attività illegali, talvolta come vera e propria copertura di traffici illeciti legati al "mercato" degli stupefacenti e della prostituzione per il riciclaggio di "denaro sporco".

4.2 Stima dell'economia sommersa

La stima del peso dell'economia non osservata per ciascuna provincia e ciascun anno è stata effettuata utilizzando la specificazione più completa dell'equazione (2), al fine di ottenere una valutazione separata della produzione sommersa e della componente attribuibile all'uso del contante nelle transazioni associate al traffico di stupefacenti e alla prostituzione (produzione illegale). Il riferimento è quindi alle stime dei parametri riportate nell'ultima colonna della Tavola 3 (MODELLO 4). Prima di procedere ad una valutazione del fenomeno in termini di incidenza media nel tempo e per macro-area geografica (Centro-Nord e Sud), si è ritenuto opportuno verificare la presenza di possibili valori estremi nelle stime ottenute, applicando il metodo – suggerito da Hadi (1992, 1994) – per l'identificazione degli *outlier* in campioni multivariati, che consente di eliminare le osservazioni “anomale” rispetto a più fattori congiuntamente considerati. Nello specifico, la verifica è stata eseguita in relazione ai valori dell'economia non osservata totale, della produzione sommersa e della produzione illegale e ha portato all'eliminazione di 26 osservazioni (livello di significatività al 5%), tutte caratterizzate da stime di una o più componenti eccessivamente elevate. I valori medi riportati nella Tavola 4 sono stati quindi ottenuti utilizzando le 338 osservazioni residue.

TAV. 4

VALORI MEDI DELL'ECONOMIA NON OSSERVATA IN % DEL PIL PER ANNO
E AREA GEOGRAFICA^a

	Produzione sommersa (evasione)	Produzione illegale (crimine)	Economia non osservata totale
ITALIA			
2005	14,5%	10,2%	24,6%
2006	15,0%	9,6%	24,5%
2007	18,0%	11,3%	29,3%
2008	18,5%	12,6%	31,2%
<i>Tutti gli anni</i>	<i>16,5%</i>	<i>10,9%</i>	<i>27,4%</i>
CENTRO-NORD			
2005	16,6%	11,5%	28,1%
2006	16,6%	11,0%	27,6%
2007	19,9%	13,0%	32,9%
2008	20,8%	14,6%	35,4%
<i>Tutti gli anni</i>	<i>18,5%</i>	<i>12,5%</i>	<i>31,0%</i>
SUD			
2005	9,7%	7,2%	16,9%
2006	11,3%	6,3%	17,6%
2007	13,6%	7,4%	21,0%
2008	13,6%	8,2%	21,8%
<i>Tutti gli anni</i>	<i>12,0%</i>	<i>7,3%</i>	<i>19,3%</i>

^a Medie calcolate sul campione con eliminazione di 26 *outlier* in base al metodo di HADI A.S. (1992, 1994).

Guardando innanzitutto alle stime calcolate su scala nazionale, emergono una serie di aspetti interessanti. In primo luogo, l'ordine di grandezza del sommerso – la media sui 4 anni è pari al 16,5% del Pil – è molto vicino ai valori ottenuti nei precedenti studi sull'Italia (in particolare le recenti stime dell'Istat, cfr. Graf. 1), ad eccezione dei risultati di Schneider e Enste (2000, 2002), che riportano valori assai più elevati (superiori al 25% da metà anni '90 fino al 2000). Tale discrepanza, come già suggerito da Zizza (2002), sembra essere imputabile alla mancata considerazione delle attività criminali nell'applicazione del CDA. Infatti, quantificando la produzione sommersa senza tenere conto degli effetti delle attività illegali (MODELLO 3 nella Tav. 3), la grandezza media che emerge è superiore e pari al 21,4% del Pil, un valore non distante dalle stime aggiornate per l'Italia presentate in Schneider (2010)²⁵ ma inferiore alla somma delle due componenti separate (evasione e crimine) che si ottiene stimando l'equazione (2) con l'inclusione di *CRIME* (MODELLO 4), prossima al 27,5%. L'incidenza del crimine è sostanziale: nel 2007 assorbe risorse pari in media all'11,3% del Pil, un valore in linea con le cifre calcolate per lo stesso anno da Eurispes (175 miliardi di euro, ovvero l'11,4% del Pil). Da ciò si deduce che, trascurando la componente criminale, si rischia non solo di imputare erroneamente a comportamenti di evasione una parte di economia non osservata derivante invece da attività illecite – che per loro natura non è possibile portare alla luce del sole al fine di recuperare il gettito perduto – ma anche di sotto-stimare il valore complessivo dell'economia non osservata. È bene inoltre rimarcare la maggiore affidabilità di queste stime rispetto agli altri lavori basati sul CDA, in virtù del fatto che – per la prima volta – viene utilizzata una misura *diretta* del valore delle transazioni effettuate in contanti, che non richiede quindi l'utilizzo dell'equazione di Fisher e le ipotesi restrittive sottostanti.

Con riferimento alla dinamica temporale, si può osservare un *trend* in crescita dal 2005 al 2008 per entrambe le componenti, sebbene l'incremento sia più marcato per la produzione sommersa (+4 punti percentuali) rispetto all'economia cri-

²⁵ L'incidenza media del sommerso negli anni dal 2005 al 2007 è pari a 23,3%. Merita però evidenziare che il confronto in questo caso risulta meno agevole rispetto ai valori calcolati fino al 2000 e presentati in SCHNEIDER F. e ENSTE D.H. (2000, 2002), dal momento che per gli anni più recenti le stime derivano da una combinazione della metodologia MIMIC (cfr. nota 7) con il CDA. Per tale ragione tali stime non sono state riportate nel Grafico 1.

minale (+2,4 punti percentuali), con un forte salto nel passaggio dal 2006 al 2007 (+3 punti percentuali e +1,7 punti percentuali rispettivamente)²⁶. Tale evidenza, sebbene in controtendenza rispetto alle stime sul sommerso dell'Istat (2010) e sull'evasione dell'Iva fornite dall'Agenzia delle Entrate (Corte dei Conti, 2011), può essere, almeno in parte, giustificata da due ordini di fattori. Innanzitutto, nel corso del 2007 l'economia italiana ha cominciato a risentire, al pari degli altri paesi dell'area euro, del ripiegamento ciclico mondiale causato dalla grave crisi finanziaria, con un forte rallentamento dei consumi e degli investimenti e un deciso peggioramento degli indicatori di fiducia delle imprese²⁷. Le aspettative negative degli operatori possono averli indotti ad una maggiore sottrazione di reddito imponibile al Fisco e al ricorso al mercato del lavoro nero, oppure a rivolgersi ad alcuni settori dell'economia illegale. Un secondo fattore che, a partire dal 2007, potrebbe aver stimolato la domanda di pagamenti in contanti da parte di chi opera nell'economia non osservata è l'entrata in vigore di una serie di norme volte a favorire una maggiore tracciabilità delle transazioni bancarie a fini di trasparenza e antiriciclaggio²⁸. Tali provvedimenti potrebbero aver provocato una contrazione dell'utilizzo (divenuto più rischioso) di strumenti bancari per finalità illecite (evasione o crimine), a favore del circolante. Questo implica che una parte del sommerso stimato dal 2007 in poi sarebbe imputabile ad attività irregolari preesistenti ma non catturate dal CDA (ad es. transazioni regolate in nero tramite assegni tratti su se stessi e girati al portatore, bonifici tra conti correnti intestati a nominativi fittizi, ecc.), dal momento che tale metodologia assume che tutti i pagamenti nell'ambito del sommerso avvengano in contanti.

²⁶ Variazioni nel tempo di entità analoga nelle stime del sommerso in Italia sono riscontrabili anche in studi precedenti. Ad esempio, nell'analisi di ZIZZA R. (2002) il sommerso aumenta di circa +2 punti percentuali tra il 1995 e il 1996, mentre BOVI M. e CASTELLUCCI L. (2001) calcolano un incremento di +2,5 punti percentuali tra il 1992 e il 1993.

²⁷ Cfr. BANCA D'ITALIA (2007, p. 8).

²⁸ In particolare, si ricorda il Regolamento CE n. 1781 del 2006 sugli obblighi di accompagnare i trasferimenti di fondi con i dati identificativi dell'ordinante (nonché quello di registrazione e verifica di tali dati) e il D.Lgs. n. 231 del 2007 in materia di antiriciclaggio che, tra le altre cose, limita il regime di circolazione degli assegni bancari e postali: i libretti di assegni di norma devono essere emessi con clausola non trasferibile, con il divieto di girata "al portatore" oltre certi limiti di importo. Le disposizioni sugli assegni in parte erano state anticipate da un Decreto di urgenza del Governo del gennaio 2007, in base al quale la banca rilascia al correntista libretti di assegni con clausola di non trasferibilità, salvo esplicita richiesta del cliente e sostenendo un maggior costo.

Per quanto concerne infine la disaggregazione delle stime a livello territoriale, la Tavola 4 mette in luce un divario significativo tra Centro-Nord e Sud, in termini sia di incidenza dell'evasione (18,5% *vs.* 12%) che del crimine (12,5% *vs.* 7,3%), che va però in controtendenza rispetto all'opinione largamente diffusa secondo cui è dal Mezzogiorno che deriverebbe il principale contributo alla formazione della nostra produzione sommersa e dell'economia illegale. In realtà, per quanto attiene alla componente di sommerso, le nostre stime sono supportate dall'evidenza emersa anche in altri studi che hanno cercato di indagare – con approcci diversi dal CDA – sull'incidenza del fenomeno nelle varie aree del paese. Bovi, Hermann, Pappalardo e Sica (2002, pp. 87-89), analizzando una serie storica dai primi anni Ottanta fino alla fine degli anni Novanta, osservano diversi periodi in cui l'intensità del sommerso è assai più elevata nelle zone del Nord (est o ovest a seconda degli anni) rispetto a quelle del Sud. In modo analogo, un recente rapporto UIL (2011) sulla diffusione del lavoro irregolare in Italia conferma anche per il periodo attuale una maggiore incidenza del fenomeno nel Centro-Nord: in base ai dati ricavati dall'attività ispettiva condotta dal Ministero del Lavoro da gennaio a ottobre 2010, tra le regioni con il più alto tasso di imprese irregolari quattro su cinque sono localizzate nel Centro-Nord (Lombardia, Liguria, Marche, Umbria); lo stesso vale per le percentuali più elevate di lavoratori in nero rispetto all'occupazione regolare, registrate in prevalenza nelle regioni del Centro-Nord. Anche negli studi di Pisani e Polito (2006) sull'evasione Irap e, soprattutto, di Marino e Zizza (2008) sull'evasione Irpef si mette in evidenza in molti casi una più elevata evasione fiscale nelle regioni del Centro-Nord rispetto al resto del paese²⁹. Da ultimo, i risultati delle indagini sull'evasione fiscale condotte nel 2011 sia dal Gruppo di lavoro "Economia non osservata e flussi finanziari" istituito dal Ministero dell'Economia (e presieduto dal presidente dell'Istat Giovannini), sia dalla Svimez su dati Istat, Agenzia delle Entrate e Ministero dell'Economia (Moro e Pica, 2011), vanno nella stessa direzione del presente lavoro.

²⁹ Dal lavoro di PISANI S. e POLITO C. (2006) per l'Agenzia delle Entrate sembra emergere, in particolare, una correlazione positiva tra i livelli di evasione IRAP stimati dagli autori a livello provinciale per gli anni 1998 e 2002 e la variabile dipendente utilizzata in questo studio per misurare l'utilizzo del contante nelle transazioni. Tuttavia, la differenza tra i periodi considerati non permette una verifica rigorosa di coerenza dei risultati. Se si guarda inoltre alle statistiche basate sugli Studi di Settore, emerge come la distribuzione territoriale del numero di contribuenti soggetti a tali indagini sia largamente favorevole al Nord (GRUPPO DI LAVORO "ECONOMIA NON OSSERVATA E FLUSSI FINANZIARI", 2011). Ciò potrebbe in parte spiegare la maggiore evasione della base imponibile Iva – quale *proxy* del sommerso – rilevata nelle regioni settentrionali.

Tale evidenza peraltro non dovrebbe stupire avendo in mente la distribuzione sul territorio delle attività produttive e le molteplici motivazioni discusse in precedenza sottostanti alla decisione di operare nel settore sommerso, evitando così una serie di adempimenti burocratici e fiscali. Una indagine campionaria condotta nel 2002 dall'Isae presso un campione di circa 4.000 imprese della manifattura e dei servizi indica la difficoltà di assunzione di nuovo personale come principale ostacolo percepito dalle imprese regolari, probabilmente per via della maggiore flessibilità delle imprese "in nero" nella gestione della manodopera in risposta a oscillazioni cicliche e all'andamento dei mercati. Seguono le motivazioni relative all'elevata pressione fiscale e contributiva, che incide in modo significativo sul grado di competitività. In particolare, disaggregando le risposte su base regionale, le imprese che segnalano una più elevata incidenza di questo tipo di ostacoli sono localizzate nel Centro-Nord (per maggiori dettagli, si veda Bovi, Hermann, Pappalardo e Sica, 2002). L'analisi di Cannari e D'Alessio (2007) sulle opinioni degli italiani riguardo l'evasione fiscale conferma l'effetto positivo della pressione fiscale locale sulla propensione ad evadere, mentre l'innalzamento della qualità dei servizi forniti dalla Pubblica Amministrazione tende a ridurla. Nel loro complesso queste evidenze aiutano a comprendere i risultati ottenuti nel nostro studio. Un'ulteriore possibile spiegazione del differenziale Nord-Sud stimato potrebbe essere legata ai differenziali nel livello dei prezzi, più bassi al Sud rispetto al Nord. A parità di unità di lavoro impiegate nella produzione sommersa, ciò condurrebbe infatti ad una remunerazione complessiva più bassa nel Mezzogiorno rispetto alle regioni settentrionali, quindi ad un minor impiego di contante per transazioni irregolari.

Con riferimento alla componente criminale, il risultato emerso di una maggiore incidenza riscontrata per il Centro-Nord probabilmente si giustifica con il fatto che l'utilizzo di contante per transazioni illegali riguarda specificamente attività criminali – traffico di stupefacenti e prostituzione – che, pur avendo "centri decisionali" localizzati in prevalenza al Sud, per effetto della mobilità delle risorse della criminalità organizzata e della concentrazione del "mercato al dettaglio" per questi beni e servizi nelle aree più ricche del paese, trovano una diffusione più intensa nelle province del Centro-Nord. Il recente lavoro di Varese (2011) sembra fornire supporto a questa tesi, identificando le condizioni che garantiscono alla criminalità organizzata di "attecchire" in aree diverse da quella di origine.

5. - Conclusioni

Questo studio si inserisce nel dibattito sulla valutazione dell'economia sommersa – una grandezza che in termini di incidenza sul Pil colloca l'Italia fra le prime non invidiabili posizioni in ambito UE – con l'obiettivo di proporre alcune innovazioni metodologiche utili a migliorare la misurazione del fenomeno. A tal fine, si suggerisce una rivisitazione del *Currency Demand Approach* (CDA), un metodo di stima indiretto in origine proposto da Tanzi (1980, 1983) e largamente utilizzato in letteratura, che si basa sull'ipotesi realistica che le transazioni “in nero” siano regolate prevalentemente tramite contanti al fine di evitarne la tracciabilità.

L'innovazione più rilevante riguarda l'impiego del *flusso* di contante prelevato dai conti correnti bancari come variabile dipendente nell'equazione di domanda di circolante (al posto dello *stock* di liquidità), che elimina la necessità di ricorrere alla teoria quantitativa della moneta e alle ipotesi restrittive sottostanti per stimare il valore delle transazioni effettuate nel sommerso. In secondo luogo, sono stati utilizzati degli indicatori diretti di presenza di comportamenti di *evasione fiscale* (tributaria e contributiva), permettendo così di superare un potenziale problema di specificazione inappropriata del modello in relazione alle diverse possibili determinanti del fenomeno, che oltre al livello della pressione fiscale dovrebbero includere anche altri fattori quali la regolamentazione del mercato del lavoro, la *tax morale* dei cittadini e l'efficienza della Pubblica Amministrazione. Un terzo aspetto saliente dell'analisi consiste nell'aver tenuto conto della presenza di transazioni in contanti di origine *criminale* (traffico di droga e prostituzione), che – in base a classificazioni largamente accettate in letteratura – identificano un'importante componente (illegale) della cosiddetta economia *non osservata*, un aggregato più ampio di cui la produzione sommersa rappresenta la parte di più diretto interesse in questo studio. Oltre che dal punto di vista metodologico – ovvero della precisione delle stime ottenute – la distinzione della produzione sommersa dall'economia criminale rileva anche sotto il profilo delle implicazioni, molto diverse, relative alle politiche di contrasto. Infatti, mentre al settore sommerso è possibile associare un gettito potenziale sottratto alle finanze pubbliche, il cui recupero richiederà un'attenta analisi delle motivazioni sottostanti alle decisioni di evasione al fine di individuare strumenti di intervento efficaci per incrementare la *compliance* (ricordando che l'approccio CDA tradizionale ha già individuato nel livello della pressione fiscale una possibile “leva”), nel caso del crimine l'azione di contrasto deve mirare a reprimere l'esercizio dell'attività illegale e la repressione non porterà con sé alcun gettito aggiuntivo. La stima separata di

queste due componenti acquista inoltre una valenza particolare nel caso italiano, considerata la forte eterogeneità delle basi imponibili e il diverso radicamento delle attività illegali sul territorio, che rendono interessante valutare il contributo del Centro-Nord e del Sud alla formazione dell'economia non osservata e verificare se esistono differenze di rilievo nell'incidenza relativa del fenomeno a seconda che si consideri la componente di evasione fiscale o quella legata al crimine.

Dai risultati, ottenuti su un campione di 91 province italiane sul periodo dal 2005 al 2008, emerge un valore medio della produzione sommersa in Italia pari al 16,5%, in linea con recenti stime da fonti statistiche ufficiali basate su metodi microeconomici di misurazione (approccio dell'*input di lavoro*, Istat, 2010). La dimensione del sommerso risulta tuttavia inferiore ai valori ottenuti per l'Italia in altri noti studi internazionali fondati sul CDA (es. Schneider e Enste, 2000, 2002; Schneider, 2010). Tale discrepanza, come suggerito già da Zizza (2002), sembra essere dovuta alla mancata considerazione del ruolo del crimine. Non a caso, estendendo il modello in modo da tenere conto delle attività illegali, è confermata l'importanza di questa componente, che assorbe risorse pari in media all'11% del Pil, generando così un valore complessivo dell'economia non osservata superiore al 27% del Pil. La conclusione a cui si perviene è che, trascurando la componente criminale, si rischia non solo di imputare erroneamente a evasione fiscale una parte di economia non osservata derivante invece dall'esercizio di attività illegali, ma anche di sotto-stimare la dimensione totale dell'economia non osservata. Al di là di ogni riscontro con le stime emerse in precedenti lavori basati sul CDA, merita soprattutto rimarcare la maggiore affidabilità delle stime qui presentate, dal momento che – per la prima volta – è stata adottata una misura *diretta* del valore delle transazioni in contanti, evitando così l'applicazione dell'equazione di Fisher e delle ipotesi poco realistiche riguardo la velocità di circolazione della moneta.

Per quanto riguarda infine l'evidenza disaggregata per aree territoriali, è emerso che le province del Centro-Nord, in media, esibiscono un'incidenza maggiore sul Pil sia dell'evasione sia del crimine rispetto alle province del Sud, un risultato che pare contraddire l'opinione assai diffusa secondo cui il Mezzogiorno sarebbe il principale responsabile della formazione della nostra produzione sommersa come pure di quella illegale. Se il risultato sulle differenze territoriali del sommerso stimola ancora di più ad indagare in futuro sulle possibili motivazioni sottostanti alla maggior tendenza ad evadere stimata per il Centro-Nord (dal livello della pressione fiscale locale al modo in cui l'efficienza della Pubblica Amministrazione influenza il grado di *compliance* dei contribuenti), quello sull'economia illegale

invita a riflettere sulla capacità delle organizzazioni criminali di “esportare” traffici illeciti (droga e prostituzione) nelle aree più ricche del paese dove è presumibilmente concentrata la domanda pagante.

Appendice

Lo studio è stato realizzato costruendo una base dati relativa a 91 province italiane osservate dal 2005 al 2008 che integra le informazioni provenienti da quattro diverse fonti: Banca d'Italia (BdI), Guardia di Finanza (GdF), Istat ed Eurostat. Le variabili monetarie sono fonte BdI e si riferiscono alle segnalazioni periodiche effettuate dal sistema bancario a livello provinciale. I dati sul valore del Prodotto interno lordo (Pil) provinciale sono fonte Eurostat. Le informazioni sugli episodi di evasione fiscale sono estratte da un database fornito dalla GdF riguardante le attività di controllo e accertamento di violazioni per ogni provincia sul periodo 2005-2008. L'indicatore di criminalità è costruito sulla base delle informazioni relative al numero di delitti a livello provinciale denunciati ogni anno dalle Forze di Polizia all'Autorità Giudiziaria e raccolte dall'Istat (disponibili al sito <http://giustiziaincifre.istat.it>).

TAV. 1

DESCRIZIONE DELLE VARIABILI UTILIZZATE NELLA STIMA DELLE EQUAZIONI (1)-(2)

Variabile	Descrizione	Fonte
<i>CASH</i>	Importo dei prelievi di contante (da conti correnti bancari) sul totale dei pagamenti non in contanti	BdI
<i>Fattori strutturali</i>		
<i>YPC</i>	Pil provinciale <i>pro-capite</i>	Eurostat
<i>BANK</i>	Numero di conti correnti bancari <i>pro-capite</i>	BdI
<i>ELECTRO</i>	Valore dei pagamenti elettronici sul Pil provinciale	BdI e Eurostat
<i>INT</i>	Tasso di interesse passivo sulla raccolta bancaria	BdI
<i>Fattori associati al sommerso fiscale</i>		
<i>EVAS1</i>	Numero di verifiche specifiche effettuate dalla Guardia di Finanza (pesate per un indice di concentrazione provinciale del Pil)	GdF e Eurostat
<i>EVAS2</i>	Numero di controlli con rilievi effettuati dalla Guardia di Finanza su ricevute fiscali e registratori di cassa (pesati per un indice di concentrazione provinciale del Pil)	GdF e Eurostat
<i>Fattori associati al sommerso criminale</i>		
<i>CRIME</i>	Quota di delitti connessi a violazioni della normativa sugli stupefacenti e sullo sfruttamento e il favoreggiamento della prostituzione rispetto al totale dei delitti denunciati (pesata per un indice di concentrazione provinciale del Pil)	Istat e Eurostat

TAV. 2

STATISTICHE DESCRITTIVE DELLE VARIABILI UTILIZZATE NELLA STIMA DELLE EQUAZIONI (1)-(2)

Variabile	Media	Deviazione Standard			Minimo	Massimo
		Totale	Between	Within		
ITALIA ^a						
CASH	0,108	0,048	0,046	0,013	0,010	0,236
YPC (x10 ⁴ €)	2,491	0,596	0,590	0,099	1,235	3,908
BANK	0,584	0,193	0,189	0,042	0,236	1,177
ELECTRO	2,100	1,728	1,598	0,672	0,538	16,638
INT	1,247	0,488	0,265	0,410	0,472	2,909
EVAS1	1,151	0,594	0,575	0,159	0,222	3,839
EVAS2	0,204	0,215	0,207	0,063	0,001	1,233
CRIME	0,023	0,020	0,019	0,004	0,001	0,116
CENTRO-NORD ^b						
CASH	0,090	0,041	0,039	0,011	0,010	0,204
YPC (x10 ⁴ €)	2,823	0,335	0,318	0,110	2,061	3,908
BANK	0,684	0,129	0,124	0,036	0,304	1,177
ELECTRO	2,399	1,962	1,802	0,800	0,538	16,638
INT	1,299	0,504	0,261	0,432	0,472	2,910
EVAS1	1,067	0,522	0,507	0,136	0,221	2,746
EVAS2	0,149	0,186	0,177	0,059	0,001	1,232
CRIME	0,022	0,021	0,020	0,003	0,001	0,115
SUD ^c						
CASH	0,148	0,038	0,036	0,016	0,063	0,236
YPC (x10 ⁴ €)	1,703	0,216	0,210	0,062	1,234	2,218
BANK	0,347	0,077	0,057	0,053	0,236	0,581
ELECTRO	1,390	0,479	0,479	0,077	0,806	2,729
INT	1,122	0,424	0,235	0,355	0,474	2,480
EVAS1	1,350	0,699	0,678	0,205	0,387	3,839
EVAS2	0,335	0,224	0,215	0,072	0,037	0,983
CRIME	0,025	0,016	0,015	0,006	0,004	0,095

^a I valori si riferiscono ad un *panel* bilanciato composto da 91 Province osservate dal 2005 al 2008, per un totale di 364 osservazioni;

^b I valori si riferiscono ad un *panel* bilanciato composto da 64 Province osservate dal 2005 al 2008, per un totale di 256 osservazioni.

^c I valori si riferiscono ad un *panel* bilanciato composto da 27 Province osservate dal 2005 al 2008, per un totale di 108 osservazioni.

BIBLIOGRAFIA

- ALVAREZ F. - LIPPI F., «Financial Innovation and the Transactions Demand for Cash», *Econometrica*, no. 77(2), 2009, pages 363-402.
- ARDIZZI G. - PETRAGLIA C. - PIACENZA M. - SCHNEIDER F. - TURATI G., «Money Laundering as a Financial Sector Crime. A New Approach to Measurement, with an Application to Italy», Center for Economic Studies & IFO Institute, *CESifo Working Paper Series*, no. 4127, febbraio, 2013.
- ARDIZZI G. - PETRAGLIA C. - PIACENZA M. - TURATI G., «Servizi di pagamento, prelievo di contante e stima del sommerso: i risultati di una indagine sul caso italiano», *Rivista Bancaria - Minerva Bancaria*, n. 4, 2011, pp. 5-37.
- , ---, ---, ---, ---, «Measuring the Underground Economy with the Currency Demand Approach: A Reinterpretation of the Methodology, with an Application to Italy», *Review of Income and Wealth*, 2013, in corso di pubblicazione.
- ARDIZZI G. - TRESOLDI C., «Spunti di riflessione sull'uso del contante nei pagamenti», *Banca Impresa Società*, n. 2, 2003, pp. 153-188.
- BANCA CENTRALE EUROPEA, *Bollettino economico*, edizione speciale, maggio, 2008.
- BANCA D'ITALIA, «I bilanci delle famiglie italiane nell'anno 2008», *Supplementi al Bollettino statistico - Indagini campionarie*, Roma, 2008.
- , *Bollettino economico trimestrale*, Roma, anni vari.
- , *Relazione annuale*, Roma, 2007.
- , *Relazione annuale*, Roma, 2009.
- BOVI M. - CASTELLUCCI L., «Cosa sappiamo dell'economia sommersa in Italia al di là dei luoghi comuni? Alcune proposizioni empiricamente fondate», *Economia Pubblica*, n. 6, 2001, pp. 77-119.
- BOVI M. - HERMANN A. - PAPPALARDO C. - SICA F., «Il sommerso: cause, intensità territoriali, politiche di regolarizzazione», in ISAE, *Rapporto Trimestrale - Priorità nazionali: trasparenza, flessibilità, opportunità*, n. 9, 2002, pp. 55-100.
- CAGAN P., «The Demand for Currency Relative to Total Money Supply», *Journal of Political Economy*, no. 66, 1958, pages 303-328.
- CAMERON A.C. - TRIVEDI P.K., *Microeconometrics: Methods and Applications*, Cambridge University Press, New York, 2005.
- CANNARI L. - D'ALESSIO G., «Le opinioni degli italiani sull'evasione fiscale», Banca d'Italia, *Temi di Discussione*, n. 618, Roma, 2007.
- CHIARINI B. - MARZANO E., «Dimensione e dinamica dell'economia sommersa: un approfondimento del currency demand approach», *Politica Economica*, n. 3, 2005, pp. 303-334.
- CORTE DEI CONTI, *Rapporto 2011 sul Coordinamento della finanza pubblica*, 2011.

- DE PLOEY W. - DENECKER O. - KEHRMANN T., *European Payment Profit Pool Analysis: Casting Light in Murky Waters*, Second Compendium on Payments Industry, McKinsey & Company, June, 2005.
- DREHER A. - KOTSOGIANNIS C. - MCCORRISTON S., «How Do Institutions Affect Corruption and the Shadow Economy?», *International Tax and Public Finance*, no. 16(4), 2009, pages 773-796.
- DREHER A. - SCHNEIDER F., «Corruption and the Shadow Economy: An Empirical Analysis», *Public Choice*, no. 144(2), 2010, pages 215-238.
- DREHMANN M. - GOODHART C.A.E., «Is Cash Becoming Technologically Outmoded? Or Does it Remain Necessary to Facilitate Bad Behaviour? An Empirical Investigation into the Determinants of Cash Holdings», Financial Markets Group Research Centre, *Discussion Paper*, no. 358, LSE, Roma, 2000.
- EURISPES, *Rapporto Italia 2008*, Roma, Istituto di Studi Politici Economici e Sociali, 2008.
- FEIGE E.L., «How Big is the Irregular Economy?», *Challenge*, no. 12, 1979, pages 5-13.
- FELD L. - FREY B.S., «Tax Compliance as the Result of a Psychological Tax Contract: The Role of Incentives and Responsive Regulation», *Law and Policy*, no. 29(1), 2007, pages 102-120.
- FERWERDA J. - DELEANU I. - UNGER B., «Revaluating the Tanzi-Model to Estimate the Underground Economy», Tjalling C. Koopmans Research Institute, *Discussion Paper*, no. 10-04, Utrecht School of Economics, febbraio, 2010.
- FRIEDMAN E. - JOHNSON S. - KAUFMANN D. - ZOIDO-LOBATÓN P., «Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries», *Journal of Public Economics*, no. 76(3), 2000, pages 459-493.
- GILES D.E.A., «Measuring the Hidden Economy: Implications for Econometric Modeling», *Economic Journal*, no. 109, 1999, pages 370-380.
- GOODHART C. - KRUEGER M., «The Impact of Technology on Cash Usage», Financial Markets Group Research Centre, *Discussion Paper*, no. 374, LSE, 2001.
- GRUPPO DI LAVORO "ECONOMIA NON OSSERVATA E FLUSSI FINANZIARI", *Rapporto finale sull'attività*, Ministero dell'Economia, 2011.
- HADI A.S., «A Modification of a Method for the Detection of Outliers in Multivariate Samples», *Journal of the Royal Statistical Society, Series B*, no. 56, 1994, pages 393-396.
- . «Identifying Multiple Outliers in Multivariate Data», *Journal of the Royal Statistical Society, Series B*, no. 54, 1992, pages 761-771.
- HONORÉ B.E., «Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects», *Econometrica*, no. 60(3), 1992, pages 533-565.
- ISTAT, «La misura dell'economia sommersa secondo le statistiche ufficiali. Anni 1992-1999», *Conti Nazionali - Statistiche in Breve*, Roma, Istituto Nazionale di Statistica, 2003.

- ISTAT, «La misura dell'economia sommersa secondo le statistiche ufficiali. Anni 2000-2008», *Conti Nazionali - Statistiche in Breve*, Roma, Istituto Nazionale di Statistica, 2010.
- LIPPERT O. - WALKER M., *The Underground Economy: Global Evidences of its Size and Impact*, Vancouver, The Frazer Institute, 1997.
- LIPPI F. - SECCHI A., «Technological Change and the Demand for Currency: An Analysis with Household Data», Banca d'Italia, *Temi di Discussione*, no. 697, Roma, 2008.
- MARINO M.R. - ZIZZA R., *L'evasione dell'Irpef: una stima per tipologia di contribuente*, XX Conferenza della Società Italiana di Economia Pubblica, Pavia, 25-26 settembre, 2008.
- MASCIANDARO D., «Il riciclaggio bancario e finanziario: conviene combatterlo?», in BANCA DELLA MAREMMA, *Evoluzione degli strumenti di pagamento. Efficienza e tutela della legalità*, Grosseto, 2007.
- MOGENSEN G.V. - KVIST H.K. - KORMENDI E. - PEDERSEN S., «The Shadow Economy in Denmark 1994: Measurement and Results», Rockwool Foundation Research Unit, *Study*, no. 3, Copenhagen, 1995.
- MORO F. - PICA F., «Italia unita nell'evasione fiscale. Basta accuse al Mezzogiorno», SVIMEZ, *nota di ricerca*, maggio 2011, www.svimez.it.
- OECD, *Measuring the Non-Observed Economy - A Handbook*, Parigi, 2002.
- PALMIERI I., «Metodologie utilizzate per quantificare l'evasione fiscale con particolare riferimento ai metodi fondati sugli accertamenti», Agenzia delle Entrate, *Documenti di lavoro dell'Ufficio Studi*, n. 2, 2004.
- PISANI S. - POLITO C., «Analisi dell'evasione fondata sui dati IRAP - Anni 1998-2002», Agenzia dell'Entrate, *Documenti di lavoro dell'Ufficio Studi*, n. 1, 2006.
- ROGOFF K., «Blessing or Curse? Foreign and Underground Demand for Euro Notes», *Economic Policy - A European Forum*, 1998, pages 261-303.
- SCHIFF M. - OZDEN C. (eds.), *International Migration, Remittances, and the Brain Drain*, World Bank, Washington, DC, 2005.
- SCHNEIDER F., «The Influence of Public Institutions on the Shadow Economy: An Empirical Investigation for OECD Countries», *Review of Law and Economics*, no. 6(3), 2010, pages 441-468.
- .-, «The Shadow Economy in Europe. Using Payment Systems to Combat the Shadow Economy», *A.T. Kearney Research Report*, settembre, 2009.
- SCHNEIDER F. - ENSTE D.H., «Shadow Economies: Size, Causes and Consequences», *Journal of Economic Literature*, n. 38(1), 2000, pages 77-114.
- .-, -.-, *The Shadow Economy: Theoretical Approaches, Empirical Studies, and Political Implications*, Cambridge University Press, UK, 2002.
- SMITH P., «Assessing the Size of the Underground Economy: The Canadian Statistical Perspectives», *Canadian Economic Observer*, no. 7(5), May, 1994.

- TANZI V., «The Underground Economy in the United States: Annual Estimates 1930-1980», *IMF Staff Papers*, no. 30(2), 1983, pages 283-305.
- . , «The Underground Economy in the United States: Estimates and Implications», Banca Nazionale del Lavoro, *Quarterly Review*, no. 135(4), 1980, pages 427-453.
- THOMAS J.J., *Informal Economic Activity*, LSE Handbooks in Economics, London, Harvester Wheatsheaf, 1992.
- TORGLER B. - SCHNEIDER F., «The Impact of Tax Morale and Institutional Quality on the Shadow Economy», *Journal of Economic Psychology*, no. 30(2), 2009, pages 228-245.
- UIL, *2° Rapporto UIL sul lavoro sommerso*, Roma, Servizio Politiche del Lavoro e della Formazione, 2011.
- VARESE F., *Mafie in movimento*, Einaudi, Torino, 2011.
- WOOLDRIDGE J.M., *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Massachusetts, MIT Press, 2002.
- ZIZZA R., «Metodologie di stima dell'economia sommersa: un'applicazione al caso italiano», Banca d'Italia, *Temi di Discussione*, n. 463, dicembre, 2002.

Previsioni delle spese del bilancio dello Stato attraverso i flussi di contabilità finanziaria

Giuseppe Bianchi*

Ministero dell'Economia e delle Finanze, Roma

Tatiana Cesaroni**

Ministero dell'Economia e delle Finanze, Roma

Ottavio Ricchi***

Ministero dell'Economia e delle Finanze, Roma

In this paper, we model and forecast monthly budget balance expenditures. The annual dimension of budget figures is integrated with an approach based on higher frequency data. To this end we develop an econometric model (disaggregated at expenditures category level) based on behavioral equations and budget balances identities linking data at different frequencies. The aggregate expenditure forecasts are obtained as a linear combination of the macroaggregates forecasts. This approach determines an improvement of forecast performance with respect to pure benchmark autoregressive models. The econometric model also represents a useful instrument to support the government forecasts based on more traditional tools. [JEL Classification: C32; C50; H50].

Keywords: state budget expenditures; infra annual financial indicators; econometric models; forecast performance.

* <giuseppe.bianchi@tesoro.it>, Ragioneria Generale dello Stato, MEF.

** <tatiana.cesaroni@tesoro.it>, Ragioneria Generale dello Stato - Servizio Studi Dipartimentale.

*** <ottavio.ricchi@tesoro.it>, Ragioneria Generale dello Stato - Servizio Studi Dipartimentale.

Gli autori desiderano ringraziare Mario Marzoli, Alessia Alessandrini, Giuseppe Taliano e Teresa Valentini per l'aiuto fornito sull'uso e l'interpretazione dei dati. Un ringraziamento va inoltre a Francesca di Brisco, Giuseppe Venanzoni ed ai *referee* anonimi per i loro suggerimenti.

1. - Introduzione

La programmazione della spesa del bilancio dello Stato e il successivo monitoraggio dei flussi finanziari rappresentano aspetti fondamentali ai fini del conseguimento degli obiettivi programmatici di finanza pubblica.

Un'esplicita formalizzazione di tale assunto viene dal recente dettato normativo concretizzatosi nell'approvazione della legge n. 196 del 2009 che sancisce il rafforzamento della manovra di bilancio su base triennale. Esiste però anche una valenza infrannuale dell'esigenza programmatica legata al monitoraggio e al controllo della spesa. Per rispettare obiettivi di spesa annuale, ed eventualmente intervenire in maniera tempestiva correggendo tendenze sfavorevoli, occorre, infatti, disporre di un profilo atteso di spesa all'interno dell'anno al quale raffrontare la sua evoluzione effettiva. L'esigenza si raccorda, risultando parimenti disponibili proiezioni delle entrate, a quella di definire un profilo sempre più attendibile del fabbisogno per potere ottimizzare la politica delle emissioni del debito pubblico.

Quest'articolo illustra un primo tentativo di modellare un profilo mensile dei pagamenti del bilancio dello Stato per l'Italia a partire dalle sue stesse grandezze e al tempo stesso di riconciliarne la dimensione tipicamente annuale con un approccio basato su serie storiche a maggiore frequenza (mensile). In primo luogo sono stimate alcune equazioni di *benchmark* di tipo autoregressivo che tengono conto dell'elevata stagionalità dei pagamenti. Successivamente si valuta la capacità predittiva rispetto a tali *benchmark* di equazioni che utilizzano quali regressori, oltre ai *lag* dei pagamenti, alcune grandezze che nell'ottica del ciclo di bilancio, si trovano "a monte" rispetto ai pagamenti. Di conseguenza, si stimano equazioni che includono gli stanziamenti, gli impegni e i residui in corso di formazione. La riconciliazione contabile delle grandezze di bilancio avviene utilizzando in sequenza delle opportune identità. Si rileva inoltre che alcune variabili hanno una valenza tipicamente annuale e sono osservate solo con tale frequenza; queste sono ad esempio gli stanziamenti iniziali – in fase di approvazione della legge di bilancio, e i residui – a livello di rendiconto.¹

Il lavoro si articola nel modo seguente. Il secondo paragrafo illustra il contesto normativo all'interno del quale si trova ad operare il tentativo di modellizzazione.

¹ Il lavoro svolto ha anche riguardato lo sviluppo di regole di proiezione per le economie sui residui, in grado di proiettare a partire dall'anno di provenienza dei residui la loro perenzione e la messa a punto di raccordi tra la modellizzazione a livello annuale delle grandezze di bilancio e la modellizzazione delle stesse a livello mensile. Questi aspetti non costituiscono ad ogni modo oggetto di trattazione del presente lavoro.

Si mettono in particolare evidenza l'accentuata attenzione del legislatore verso il controllo, monitoraggio e programmazione dei conti pubblici (in prospettiva formulati in termini di cassa). Il paragrafo esplicita anche che il modello realizzato può rappresentare, in forma prototipale, una piattaforma (a livello di macro aggregati) utile per operare previsioni a breve termine dei flussi mensili del bilancio dello Stato e una definizione dei valori attesi dei principali aggregati del bilancio all'interno dell'esercizio finanziario. Il terzo paragrafo esamina i principali tentativi effettuati in Italia dalla letteratura di modellazione a fini previsivi di serie storiche di finanza pubblica e dà conto delle tecniche econometriche adottate in alcuni di essi. Il quarto paragrafo illustra le caratteristiche del bilancio dello Stato e dei relativi dati provenienti dai sistemi informativi della Ragioneria Generale dello Stato (RGS) che sono oggetto di analisi; nella parte finale s'illustrano inoltre le tecniche utilizzate per trasformare tali dati in serie storiche. Il quinto paragrafo riporta l'analisi descrittiva effettuata sui dati di flusso (l'ammontare mensile dei pagamenti) e l'approccio utilizzato per il trattamento della stagionalità. Il paragrafo contiene inoltre una prima analisi delle interrelazioni tra l'insieme delle variabili considerate e i pagamenti al fine di avere prime indicazioni sulla capacità predittiva delle prime. Il sesto paragrafo descrive il modello costruito che è basato sulla stima di equazioni in cui si utilizzano variabili che nell'ottica del ciclo di bilancio si trovano "a monte" rispetto ai pagamenti, sull'inserimento delle equazioni in una struttura caratterizzata da identità contabili e sulla possibilità del passaggio alla frequenza annuale con il calcolo dei residui previsti. Il settimo paragrafo contiene delle prime indicazioni ottenute dall'utilizzo del modello: in primo luogo se ne raffronta la capacità previsiva rispetto a un modello basato soltanto sulle equazioni univariate in precedenza stimate; in secondo luogo si presenta un confronto tra le previsioni annuali di spesa per il 2008 e il 2009 contenute nella Relazione Unificata sull'Economia e la Finanza Pubblica e quelle prodotte dal modello.

L'articolo si conclude con una valutazione di tutte le direzioni di lavoro aperte, tra le quali la valutazione del livello ottimale di disaggregazione dell'analisi (pur restando all'interno di una dimensione "macro" del bilancio), la possibilità di introdurre grandezze esterne al bilancio nella previsione di alcuni aggregati di spesa (in particolare per le componenti di spesa obbligatorie), la proiezione delle entrate di bilancio (e la conseguente previsione dei saldi) e, in prospettiva, il raccordo con i dati di Tesoreria.

2. - Controllo, monitoraggio e programmazione dei flussi di cassa alla luce del recente contesto normativo e il ruolo del bilancio dello Stato

Come anticipato nell'introduzione, riteniamo utile riassumere gli aspetti principali del quadro normativo all'interno del quale opera il tentativo di modellizzazione presentato. Questo per meglio evidenziare il contributo fornito dal modello proposto.

La legge di riforma della contabilità e della finanza pubblica n. 196 del 2009, in un'ottica di contenimento della spesa pubblica e di attenta pianificazione delle emissioni di debito pubblico, pone particolare rilievo alle prescrizioni relative al monitoraggio e controllo dei conti pubblici e della programmazione dei flussi di cassa². A tal riguardo sono di particolare rilevanza i compiti affidati alla RGS dall'articolo 14. Tra questi si ricordano, perché più rilevanti nel contesto del presente articolo, la valutazione dell'andamento delle principali grandezze e della loro coerenza rispetto agli obiettivi fissati dal Documento di Economia e Finanza attraverso il controllo in corso di gestione e la verifica a consuntivo del raggiungimento degli obiettivi stessi (art. 14, comma 1, lett. *b*) e il monitoraggio degli effetti finanziari delle misure previste dalla manovra di bilancio e dei principali provvedimenti adottati in corso d'anno (art. 14, comma 1, lett. *c*). Inoltre l'articolo 17, nel disciplinare la copertura finanziaria delle leggi di spesa, definisce una clausola di salvaguardia che implica il monitoraggio delle spese autorizzate e o previste da nuovi provvedimenti legislativi.

Il legislatore nella l. n. 196 del 2009 ha dedicato particolare attenzione agli andamenti del bilancio in termini di cassa e a seguire la loro evoluzione nel corso della gestione in "tempo reale". Ne sono testimonianza le ulteriori specifiche in aggiunta alle "tradizionali" pubblicazioni relative al monitoraggio dei flussi di cassa (art. 14, co. 4)³.

Va inoltre evidenziato che le indicazioni presenti nell'articolo 14 si raccordano con il disposto degli artt. 40 e 42 che riguardano la revisione della struttura del

² Per una più approfondita descrizione si veda *La riforma della contabilità e della finanza pubblica - Novità, riflessioni e prospettive*, Ragioneria Generale dello Stato, 2010.

³ Le relazioni trimestrali di cassa (di maggio, settembre e novembre) devono contenere i valori riscontrati nel corso dell'anno, coerentemente con le grandezze utilizzate per la programmazione, dei saldi di cassa dell'amministrazione centrale e delle amministrazioni pubbliche. Inoltre, sempre con riferimento al monitoraggio dei flussi di cassa, viene istituzionalizzata la pubblicazione mensile, entro il mese successivo a quello di riferimento, della relazione sul conto consolidato di cassa dell'amministrazione centrale (di cui il "vecchio aggregato" settore statale è comunque la principale componente).

bilancio dello Stato, la disciplina della sua gestione e il potenziamento della funzione del bilancio di cassa. In tali articoli si afferma un approccio di programmazione degli andamenti del bilancio dello Stato che si spinge fino al livello della singola unità responsabile della spesa. Si prevede infatti (art. 42, primo co., lettera c)) che il dirigente responsabile predisponga un piano finanziario che tenga conto della fase temporale di assunzione delle obbligazioni sulla base delle quali ordina e paga la spesa.

Questa prescrizione è testimonianza di un intento di costruire una rete d'informazioni relativa alla tempistica attesa dei pagamenti che è raccordabile, in un'ottica di controllo di gestione finanziario, ad una attività di monitoraggio e controllo della spesa a carattere infra annuale. Per poter rispettare obiettivi di spesa annuale, ed eventualmente intervenire in maniera tempestiva per correggere tendenze sfavorevoli, occorre disporre per ciascuno degli aggregati di competenza di un profilo atteso di spesa all'interno dell'anno al quale raffrontare la sua evoluzione effettiva. La possibilità di effettuare tali confronti si basa sulla disponibilità di informazione in tempo reale per ciascuna delle voci di spesa più rilevanti.

Sia pure meno rappresentativo degli andamenti di finanza pubblica rispetto al passato, il bilancio dello Stato costituisce comunque il punto di riferimento più importante per la finanza pubblica italiana e il principale luogo, dove si concentra e si rappresenta la decisione politica sulle diverse componenti di entrata e di spesa e sui saldi di bilancio. In termini di spesa erogata al netto degli interessi la percentuale di competenza delle amministrazioni centrali (di cui il bilancio dello Stato è la principale componente) è pari soltanto al 25% del totale delle spese del conto consolidato della Pubblica Amministrazione. Tuttavia prima che si esegua il consolidamento, la percentuale di spesa in uscita dal bilancio dello Stato, è pari a più del 60%. Come è noto, una parte rilevante delle somme in uscita dal bilancio è costituita da trasferimenti verso gli altri enti pubblici a cui compete la decisione di spesa. Il bilancio dello Stato continua a essere molto rilevante in termini di determinazione dell'ammontare della spesa perché ad esempio tagli dei trasferimenti impattano, anche per via di meccanismi istituzionali, sul livello della spesa degli altri sottosettori. Ne sono testimonianza le recenti vicende di tagli ai trasferimenti operati agli enti territoriali ⁴.

⁴ Il nuovo processo delineato dalla legge n. 196 prevede una maggiore condivisione con le autonomie locali riguardo alle decisioni sugli obiettivi complessivi di finanza pubblica e dunque in prospettiva una minore rilevanza del bilancio qualora con il pieno realizzarsi del Federalismo Fiscale si assistesse ad un "dimagrimento" del bilancio dello Stato tramite la riduzione dei trasferimenti e una maggiore capacità impositiva a livello locale.

La conoscenza e la previsione dei suoi flussi a una frequenza mensile risulta propedeutica al conseguimento degli obiettivi di controllo, monitoraggio e programmazione appena illustrati dal momento che il bilancio dello Stato rappresenta il primo “mattoncino” a partire dal quale si costruiscono tutti gli altri aggregati di finanza pubblica. Con riferimento in particolare alla spesa pubblica, sulla quale questo contributo si concentra, si ritiene prioritaria in sede di effettuazione di previsioni l’esigenza di partire dall’aggregato “iniziale”, e cioè dai pagamenti del bilancio dello Stato, per poi operare gli opportuni raccordi⁵. Sicuramente, per effettuare questi ultimi nella misura migliore occorre risolvere sia problemi di sfasamenti temporali (consolidamento con i dati di tesoreria per il fabbisogno del settore statale) che di integrazione con aggregati di riferimento più ampi (per le amministrazioni centrali e, naturalmente, per il passaggio successivo alle amministrazioni pubbliche).

La legge n. 196 del 2009 opera in entrambe le direzioni. Tra i diversi interventi previsti nell’ambito dell’art. 40 (delega al Governo per il completamento della revisione del bilancio dello Stato) rientra la revisione del conto riassuntivo del tesoro allo scopo di garantire maggiore chiarezza e significatività alle informazioni in esso contenute tramite l’integrazione tra i dati contabili del bilancio dello Stato e quelli di tesoreria, co. 2, lettera *o*). Inoltre, nel co. 1, lettera *b*), del menzionato art. 42, sempre in funzione del potenziamento del bilancio di cassa, si prevede il raccordo tra le autorizzazioni di cassa del bilancio dello Stato e la gestione di tesoreria.

La tendenza quindi del disposto legislativo è quella di restituire significatività al bilancio dello Stato nel senso di dare maggiore contezza di quali somme stanziare e successivamente impegnate, si tramutano in spesa effettiva e quando ciò avviene⁶.

Per quanto riguarda il raccordo del bilancio dello Stato con gli aggregati più ampi del conto delle amministrazioni pubbliche, si ricorda che uno degli aspetti pregnanti della legge n. 196 consiste nella previsione di un piano dei conti che consenta di migliorare il consolidamento.

⁵ Per una attenta descrizione dei principali problemi di raccordo con le altre grandezze vigenti prima della approvazione della legge n. 196 si veda *I principali saldi di finanza pubblica: definizioni, utilizzo, raccordi* (2008). Si ricorda inoltre che tale legge ha posto particolare enfasi nei principi di coordinamento e armonizzazione dei conti, per cui, in prospettiva, i problemi di raccordo dovrebbero ridursi.

⁶ All’intento di fornire maggiore chiarezza è finalizzata anche la lettera *p*), del co. 2, dell’art. 40. Il cui disposto prevede, con alcune eccezioni, la progressiva eliminazione delle gestioni contabili operanti a valere su contabilità speciali o conti correnti di tesoreria i cui fondi siano stati costituiti mediante somme originariamente iscritte in stanziamenti di spesa del bilancio dello Stato.

L'andamento mensile del bilancio dello Stato è già oggetto di attento monitoraggio da parte della RGS e il risultato di tali controlli confluisce in una pubblicazione attualmente trasmessa alle Camere⁷. In un'ottica previsiva di valori mensili si ricorda infine il «Rapporto mensile sull'andamento delle entrate tributarie e contributive», redatto dal Dipartimento delle finanze e dal Dipartimento della Ragioneria generale dello Stato, secondo le disposizioni della legge di contabilità e finanza pubblica (art. 14, comma 5, l. n. 196 del 2009).

L'idea portata avanti nel presente articolo e motivata dall'illustrazione del contesto normativo di riferimento è quella di aggiungere alle procedure attualmente esistenti un ulteriore strumento di carattere statistico a frequenza mensile. Tale strumento, andando a integrare l'informazione fornita dai modelli con frequenza annuale utilizzati presso il MEF, potrebbe rappresentare un supporto all'effettiva previsione e al monitoraggio infra annuale delle grandezze di bilancio, a cominciare dalle spese.

L'intento degli autori è stato quello di costruire un modello basato su stime econometriche, che possa fornire indicazioni di varia natura relativamente al bilancio dello Stato, quali:

- a) un profilo atteso infra annuale dei pagamenti (supporto alla programmazione);
- b) aggiornamenti delle previsioni mensili a breve termine sulla base dei risultati di monitoraggio all'interno dell'esercizio finanziario (ottica di monitoraggio/controllo di gestione finanziario);
- c) l'implicazione di determinati interventi al livello di bilancio – quali variazioni degli stanziamenti – sul profilo dei pagamenti (ottica di simulazione);
- d) previsioni sullo *stock* di residui alla fine dell'esercizio finanziario il tutto mantenendo coerenza contabile tra le variazioni delle diverse grandezze di bilancio.

La capacità per un tale strumento di dare indicazioni soddisfacenti è funzione, oltre che della qualità e del contenuto informativo dei dati, di un insieme di scelte riguardanti il livello di aggregazione (passaggio da titoli a categorie o a raggruppamenti di queste) e la modellazione delle serie storiche. Il modello realizzato dagli autori e presentato in questo lavoro non ha immediate valenze operative ma opportuni affinamenti successivi, nelle direzioni indicate all'interno del paragrafo finale, potranno assicurargli tale caratteristica.

A prescindere dai futuri possibili utilizzi del modello, il valore aggiunto che si ritiene di dare in questo contributo consiste nell'introduzione di un approccio statistico/econometrico su dati del bilancio ad alta frequenza e nell'inizializzazione

⁷ *Monitoraggio degli incassi e dei pagamenti del bilancio dello Stato.*

di uno strumento di analisi e previsione che, alla luce di successivi approfondimenti, potrebbe fornire anche utili indicazioni in termini di valutazione dell'impatto finanziario delle politiche di bilancio.

3. - Le precedenti esperienze in letteratura nell'utilizzo di tecniche di modellizzazione

Prima di passare all'illustrazione dei dati utilizzati e dei risultati conseguiti giova ricollegarsi ai precedenti contributi in materia. La selezione dei lavori presi in rassegna ha seguito diversi criteri. In primo luogo era d'interesse dare conto degli sforzi di modellizzazione di aggregati di finanza pubblica vicini il più possibile a quelli da noi analizzati. Il numero di contributi dedicati alle analisi del bilancio dello Stato è risultato abbastanza ridotto, ma si è anche tenuto conto di diversi lavori dedicati alla previsione del fabbisogno del settore statale nelle sue diverse componenti (di spesa e di entrata). Questi aspetti erano di particolare interesse in quanto caratterizzati da problematiche in termini di stima delle equazioni molto vicine a quelle da noi affrontate. È stato quest'ultimo il secondo criterio di scelta. Infine si è scelto di fare riferimento anche a lavori che, pur non connotandosi come analisi empiriche tipo statistico, potessero dare indicazioni sulla scelta degli aggregati di riferimento da utilizzare. In particolare questi contributi forniscono indicazione sulle scelte di disaggregazione da effettuare in prospettiva.

In Italia l'esperienza di modellizzazione di serie storiche di finanza pubblica con frequenza mensile ha conosciuto una fase di particolare sviluppo tra la fine degli anni '80 e i primi anni '90, con una propaggine alla fine degli anni '90. Dopo circa un decennio di mancanza di ulteriori apporti, solo negli ultimi anni si segnala una "ripresa delle attività".

Una delle prime esperienze è stata portata avanti da Artoni, Bernardi e Lippi (1985) con un tentativo di analisi delle serie storiche del bilancio dello Stato (alcune delle quali a frequenza mensile) improntato a un approccio prettamente operativo. Pur riconoscendo che nel contesto particolarmente disaggregato della struttura del bilancio si rendeva "inevitabile l'uso preminente di modelli deterministici e istituzionali", lo studio segnalava l'auspicabilità di migliorare le tecniche a quel tempo in uso di previsione delle grandezze inserite nel bilancio annuale. L'analisi esplorativa sui dati di bilancio conteneva il primo tentativo di modellare alcune serie storiche di bilancio (analizzate a frequenza mensile, trimestrale ed annuale) utilizzando modelli autoregressivi.

Questo segnale fu raccolto presso l'allora Ministero del Tesoro istituendo un

apposito gruppo di lavoro⁸ i cui sforzi si concentrarono in prevalenza sulla valutazione e previsione con tecniche econometriche del fabbisogno del settore statale⁹ nonché sulla analisi ex-post degli errori di previsione, anche in prospettiva di utilizzare questi ultimi per migliorare le stime. Come riferito in Giovannini (1991), utilizzando dati del Conto Riassuntivo del Tesoro, il lavoro era caratterizzato da un approccio fortemente disaggregato che determinava la stima di 26 equazioni univariate.

Le capacità previsive delle equazioni risultavano migliorate introducendo al loro interno “variabili d’intervento” (*dummies* puntuali o a carattere stagionale). Tuttavia la scelta di quest’ultime – per ammissione dell’autore – rappresentava un fattore di criticità ai fini previsivi¹⁰ poiché il loro utilizzo implicava una continua attività di manutenzione delle equazioni e di valutazione riguardo al loro eventuale inserimento in funzione di conoscenze istituzionali. Utilizzando per la valutazione della *performance* previsiva delle equazioni le usuali statistiche (ME e MAE) sugli errori di previsione commessi nel campione su un arco temporale di quattro mesi (corrispondente alle esigenze operative della previsione), lo studio forniva risultati considerati soddisfacenti. Si riconosceva, tuttavia, che i valori previsivi di fabbisogno “ufficiali” prodotti dal Ministero del Tesoro mostravano capacità previsive migliori. Questi ultimi erano ottenuti partendo dalle stime econometriche e però facendo ricorso anche a informazioni “esterne” di natura istituzionale¹¹ pervenute successivamente all’utilizzo delle equazioni. L’autore concludeva al riguardo argomentando che, nonostante le previsioni ufficiali risultassero “migliorative” rispetto a quelle fornite dalle equazioni, queste ultime

⁸ Coordinato da Enrico Giovannini e composto da alcuni funzionari del Tesoro.

⁹ L’impulso a stimare tale aggregato proveniva anche dall’esigenza e dalla forte volontà di quel periodo di migliorare e razionalizzare la politica di emissioni del debito pubblico ai fini della quale, come si è chiarito, è molto rilevante disporre di previsioni affidabili del fabbisogno primario.

¹⁰ L’introduzione di variabili di intervento in previsione è particolarmente critica per quanto riguarda le variabili di gettito fiscale. Infatti, una variazione nella struttura dell’imposizione risulta difficilmente modellabile in una equazione di tipo univariato basato solo sui valori passati delle variabili. In tal caso si predilige utilizzare modelli istituzionali con una modellazione delle basi imponibili; naturalmente come si vedrà l’utilizzo di tali variabili apre una diversa serie di problematiche.

¹¹ Queste riguardavano prevalentemente futuri movimenti di tesoreria suggeriti da “intenzioni di prelevamento o versamento da parte di diversi enti o di istituzioni del settore statale”. I valori previsivi ufficiali erano utilizzati nei tavoli operativi come base di discussione con le autorità monetarie riguardo alle forme di copertura delle necessità finanziarie.

dovevano avere rappresentato un nuovo e valido punto di partenza dal momento che l'istituzione della prassi del loro utilizzo all'interno del processo previsivo aveva portato ad una rilevante riduzione degli errori di previsione rispetto al passato.¹²

In un contributo successivo l'articolo di Giovannini, Zanardi (1991) pone l'accento sulla rilevanza assunta dall'aggregato fabbisogno del settore statale quale indicatore della consistenza effettiva dei flussi finanziari attivati dallo stato. Ciò era dovuto alla recente decisione di accentrare presso la Tesoreria le disponibilità liquide degli enti destinatari di trasferimenti da parte del bilancio dello Stato e in precedenza operate su conti correnti detenuti da tali enti presso il settore bancario. La movimentazione (tiraggio) da parte degli enti delle somme disponibili presso la Tesoreria secondo tempistiche irregolari e/o comunque diverse rispetto a quelle dell'accreditamento ricevuto presso il bilancio dello Stato rendeva i flussi del bilancio dello Stato meno rappresentativi (quantomeno dal punto di vista della tempistica) dell'andamento dei conti pubblici. Le decisioni di bilancio prese dal parlamento venivano così a "essere filtrate dai comportamenti di spesa degli enti pubblici". Il fabbisogno del settore statale, quale combinato di due distinte gestioni – quella del bilancio dello Stato e quella della tesoreria – era dunque divenuto l'aggregato rilevante. Una analisi effettuata sugli errori di previsioni compiuti nel passato dagli uffici preposti¹³ portava l'autore a fornire alcuni suggerimenti per migliorare la loro capacità previsiva. Questi erano perseguibili muovendosi lungo due direzioni: prendendo in considerazione la stagionalità/regolarità di alcune voci dei pagamenti e andando a migliorare il flusso informativo dagli enti che effettuavano il tiraggio. Il percorso suggerito era quello di costruire, a partire da una previsione (obiettivo) annuale sufficientemente affidabile, una iniziale ripartizione teorica mensile utilizzando coefficienti di correzione stagionale per le

¹² Si osserva tuttavia al riguardo che è possibile che una parte del miglioramento fosse piuttosto dovuto ad un semplice affinamento delle tecniche previsive non modellistiche o della capacità di ottenere informazioni relazionandosi con le altre amministrazioni pubbliche.

¹³ Come punto di partenza del suo contributo, Zanardi valuta la capacità previsiva per le singole voci disaggregate delle previsioni qui di sopra denominate "ufficiali" del fabbisogno del settore statale. L'analisi rivelava che per alcune di esse si verificavano errori sistematici, che tendevano a non compensarsi. Le voci per cui gli errori risultavano generalmente maggiori avevano una doppia matrice: quelle corrispondenti a pagamenti degli enti locali (che godevano di una ampia discrezionalità rispetto ai tempi ed alla intensità del prelievo dalla tesoreria) e quelle per le spese di pagamento del personale in quiescenza e in servizio (sia pure in misura minore), difficili da prevedere in quanto soggette a modifiche e innovazioni legislative.

diverse variabili e di formulare la previsione finale tenendo conto delle conoscenze operative sviluppate dai funzionari del Tesoro.

L'articolo di Zanardi segnalava inoltre che il passaggio alla Tesoreria unica, avvenuto nel 1986, aveva sì portato a una razionalizzazione e minor costo della gestione complessiva della liquidità del sistema e a una maggiore controllabilità dell'aggregato complessivo del fabbisogno del settore statale, tuttavia aveva aumentato – quantomeno a livello mensile – le difficoltà di prevedere le uscite che in precedenza erano “semplicemente” cadenzate dalle uscite dal bilancio dello Stato con accreditamenti sugli istituti bancari tesorieri degli enti locali.

Qualche anno dopo Pisauro (1994) tornava sull'argomento della Tesoreria. Al termine della sua analisi concludeva che il ruolo assunto dalla Tesoreria dello Stato rappresentava un estremo tentativo di correzione di un problema che risiedeva nel bilancio dello Stato e che consisteva nella sottovalutazione di alcuni programmi di spesa incompressibili senza modifiche legislative¹⁴. L'osservazione era dunque da stimolo al miglioramento della produzione di previsioni – innanzitutto annuali – della spesa e in genere alla valutazione dell'impatto dei provvedimenti di spesa.

La necessità del monitoraggio del fabbisogno del settore statale nelle sue componenti rilevanti ha trovato una risposta nel gruppo di lavoro coordinato da Giarda (1997) presso il Ministero del Tesoro. Scopo del gruppo di lavoro era ripristinare una tempestiva capacità di valutare l'effetto finanziario (netto) sul fabbisogno delle diverse tipologie di movimenti dei conti di Tesoreria¹⁵. Anche se il gruppo di lavoro non si concentrò sulla previsione di valori di fabbisogno ma soltanto sulla tempestività e correttezza della stima dei valori, resta rilevante ai fini di quest'articolo l'approccio scelto per ripartire i settori di spesa da studiare. A tale scelta si accennerà nella parte conclusiva del presente lavoro.

Non risultano agli scriventi, per circa un decennio, ulteriori contributi successivi rilevanti per le finalità del presente lavoro. Questo non significa che non si siano realizzati affinamenti della capacità previsiva degli aggregati in questione

¹⁴ I debiti che risultavano in Tesoreria non venivano inclusi nel fabbisogno del settore statale ma ripianati con emissioni di titoli del debito pubblico causando una incoerenza tra i flussi e la crescita dello *stock*.

¹⁵ Contestualmente alla formazione del gruppo di lavoro la legge finanziaria per il 1997 si era fatta carico, tra le altre misure volte a contrastare le principali “patologie” del sistema, di ridurre le giacenze sui conti di Tesoreria subordinando i pagamenti di bilancio verso enti alla riduzione al 20% delle giacenze rispetto ai valori di fine 1996.

all'interno delle strutture ministeriali sia in termini di modellistica¹⁶, sia di buona pratica presso gli uffici deputati alle previsioni. Tuttavia è probabilmente venuto a ridursi l'interesse da parte della comunità accademica su tale tema forse perché alcune delle questioni e problematiche sono apparse meno pressanti¹⁷. Non si tornerà qui sulle motivazioni per le quali si giustifica a nostro parere il lavoro in questione, ci si limita a ricordare nuovamente che il bilancio dello Stato e a seguire il fabbisogno del settore statale restano, comunque, il primo tassello e il punto di partenza dal quale osservare i fenomeni di finanza pubblica.

Per un approccio recente volto a utilizzare tecniche econometriche per l'analisi dei dati a frequenza mensile al fine di prevedere i valori del fabbisogno del settore statale per l'Italia, bisogna attendere il contributo di Pozzuoli ed Ercoli (2008). Gli autori utilizzano le serie storiche mensili del fabbisogno elaborate presso la RGS e relative al periodo 2002-2007 al fine di prevederne gli andamenti nel breve termine. Il lavoro, dopo un'analisi iniziale dell'accuratezza *ex post* delle previsioni contenute nei principali documenti di finanza pubblica, procede ad un attento esame delle proprietà statistiche dei dati mensili di fabbisogno al fine di costruire modelli statistici di previsione a breve termine. Il primo tentativo di modellizzazione dei dati di fabbisogno riguarda l'eliminazione della componente stagionale e di trend dalla serie storica attraverso l'uso di tradizionali modelli stocastici. Il risultato conduce alla stima di un'equazione univariata di tipo Sarima, selezionata fra tre modelli Sarima alternativi, che su un orizzonte temporale da uno a sei passi sembra fornire risultati soddisfacenti. Il secondo tentativo di stima consiste in un'analisi disaggregata del fabbisogno, attraverso la stima di modelli di previsione Sarima separati per le singole componenti di entrata e di spesa. Tuttavia, l'analisi disaggregata non mostra risultati incoraggianti tanto che la *performance* di previsione del fabbisogno ottenuta come differenza tra entrate e uscite risulta peggiore della previsione diretta del saldo. Un ultimo tentativo di modellizzazione del fabbisogno del settore statale viene effettuato attraverso la stima di un modello uni-

¹⁶ Si pensi ad esempio alla continua manutenzione e miglioramenti dei modelli previsivi di finanza pubblica all'interno della Ragioneria Generale dello Stato relazionati nel volume *I documenti programmatici: ruolo, strutture, processi e strumenti del MEF* (2006).

¹⁷ Da una parte con il passaggio all'indebitamento netto delle AAPP come misura determinante delle amministrazioni pubbliche il fabbisogno del settore statale è diventato meno (direttamente) rilevante, dall'altra la ormai comprovata elevata capacità di gestire il finanziamento del fabbisogno da parte delle strutture ministeriali competente potrebbe aver reso meno evidente alla comunità scientifica l'importanza di prevederne in maniera sempre più precisa l'importo mensile.

variato a componenti non osservate. Tuttavia, i risultati della *performance* di previsione non si discostano complessivamente da quelli ottenuti attraverso l'uso dei modelli Sarima come evidenziato peraltro anche dagli stessi autori. Non si procede nell'articolo all'utilizzo di stime multivariate (modelli multiequazionali e/o VAR) poiché si afferma che i dati sul fabbisogno si rendono disponibili abbastanza rapidamente (i dati relativi a ciascun mese si rendono ormai disponibili all'inizio del mese successivo) mentre gli indicatori dell'economia reale che potrebbero esserne utili alla stima (in quanto correlati con le serie oggetto di esame) sono aggiornati con una minore tempestività¹⁸.

Tra i recenti approcci di analisi sui saldi di finanza pubblica relativi ad altri paesi figurano il contributo di Moulin, Salto, Silvestrini e Veredas (2004) in cui gli autori presentano un approccio per la previsione del saldo del settore statale francese basato sull'utilizzo di flussi di entrate e spese mensili modellati attraverso modelli autoregressivi.¹⁹ Secondo gli autori l'esercizio previsivo svolto mostrerebbe che le previsioni un passo avanti ottenute utilizzando i modelli aggregati a livello annuale per le singole categorie migliorerebbe le previsioni ufficiali effettuate dal governo e basate solo sul giudizio degli analisti. Ulteriori contributi concernenti l'uso di dati infrannuali del settore statale sono infine quelli di Pedragal e Perez (2008) e Leal, Pedregal e Perez (2009) che seguono un approccio basato su modelli a frequenza mista per effettuare monitoraggio e previsioni dei deficit della pubblica amministrazione rispettivamente dell'Area Euro e della Spagna ed Onorante ed altri (2010) che usano sempre modelli a frequenza mista per prevedere variabili fiscali dell'Area Euro.

Sebbene i lavori di Moulin, Salto, Silvestrini e Veredas (2004), Pozzuoli e Ercoli (2008) Pedragal e Perez (2008) e Leal, Pedregal e Perez (2009) rappresentino

¹⁸ I dati di finanza pubblica, quanto meno quelli relative alle amministrazioni centrali, sono molto più tempestivi di quelli relative all'andamento dell'economia che – ad esempio – potrebbero essere utili alla previsione delle entrate. Nel caso delle spese Giovannini aveva trovato che l'introduzione di variabili ulteriori all'interno di alcune equazioni di spesa o di pagamento migliorava tali equazioni ma la loro *performance* previsiva peggiorava dovendo proiettare i valori di tali variabili che erano disponibili con ritardo rispetto a quelle di finanza pubblica.

¹⁹ Dopo aver stimato modelli Arima mensili per le entrate e le spese disaggregate a livello di categorie inferiscono i parametri di un modello di previsione annuale a partire dai parametri stimati dei modelli mensili tramite un'aggregazione temporale. Dal punto di vista operativo quest'ultima si basa sull'individuazione degli ordini $p d q$ del processo come pure dei parametri e della varianza del modello Arima che descrive il modello annuale tramite un sistema di equazioni che, a partire dagli ordini $p d q$ noti a livello mensile, determina i corrispondenti valori a livello annuale.

validi tentativi di impiego di modelli econometrici di serie storiche al fine di prevedere nel breve termine il fabbisogno del settore statale e/o di aggregati di finanza pubblica più ampi, presentano l'inconveniente insito nell'uso di tutti modelli statistici univariati, ovvero quello di utilizzare soltanto l'informazione passata per prevedere gli andamenti futuri del fenomeno. Tali modellizzazioni non consentono, infatti, di cogliere tempestivamente variazioni nelle serie dovuti a *shock* e cambiamenti di natura esterna alla serie che si sta modellando. In questo lavoro, come sarà mostrato in modo più approfondito nei paragrafi successivi, si cerca di superare l'inconveniente legato a tali approcci utilizzando equazioni disaggregate per le singole componenti di spesa (a livello di titoli) che utilizzino come regressori non soltanto valori passati della serie che si desidera prevedere ma anche l'informazione contenuta negli indicatori mensili provenienti dal bilancio dello Stato collegati "contabilmente" con le variabili che si desidera prevedere e che potremmo definire *leading* per i pagamenti in quanto da un punto di vista contabile (in termini del ciclo di bilancio) più a monte rispetto alle serie che si desidera prevedere. Tale approccio rappresenta pertanto una apprezzabile innovazione rispetto alle usuali tecniche di previsione autoregressive per la possibilità di utilizzare le informazioni provenienti da variabili interne allo stesso bilancio ai fini previsivi.

4. - Il ciclo del bilancio dello Stato e le fonti utilizzate; il trattamento dei dati e le serie storiche mensili

Per una più agevole lettura dei restanti paragrafi, dedicati alla spiegazione della costruzione di serie storiche mensili del bilancio dello Stato e alle scelte di modellizzazione effettuate si è ritenuto opportuno premettere una breve spiegazione relativa al ciclo del bilancio dello Stato. Quest'aspetto di contabilità pubblica è infatti una competenza estremamente specialistica e di non immediata comprensione per i non addetti ai lavori.

Il ciclo del bilancio ha inizio con l'approvazione della legge di bilancio (prima fase) in cui il parlamento fissa i limiti e i contenuti della gestione finanziaria dello Stato e ne è autorizzata l'esecuzione²⁰. Tale atto, in sostanza, dà il via all'esercizio finanziario e quindi alla gestione del bilancio (seconda fase). Le autorizzazioni di

²⁰ La legge di bilancio è formata apportando al disegno di legge di bilancio le variazioni determinate dalla legge finanziaria. A partire dal 2010, con la riforma della legge di contabilità e finanza pubblica (legge n. 196 del 2009), le variazioni vengono determinate dalla legge di stabilità, *fonte*: Glossario della Ragioneria Generale dello Stato.

spesa sono inserite nei sistemi contabili e identificate con il termine “stanziamenti”. Esse costituiscono limiti giuridici oltre i quali l’amministrazione non è autorizzata a spendere. Questi limiti possono essere variati nel corso dell’anno sia in occasione della cosiddetta “legge di assestamento di bilancio”, sia tramite provvedimenti amministrativi. Facendo riferimento ai relativi stanziamenti, l’amministrazione dello stato, durante la gestione, effettuerà impegni di spesa (cioè si vincolerà tramite degli atti che abbiano valore giuridico ad effettuare delle spese) che successivamente daranno luogo a dei pagamenti. Entrambe le misure (impegni e pagamenti) sono naturalmente registrate nei sistemi informatici.

Per chiarire le finalità per le quali è stato effettuato il presente lavoro, va rilevato che essendo l’ammontare dei pagamenti vincolato dalla autorizzazione del bilancio, in teoria, non si necessita di una previsione vera e propria²¹. Piuttosto appare utile tracciare un profilo di pagamenti “atteso” che consenta un monitoraggio sullo stato di attuazione della gestione del bilancio in corso d’anno. Ciò per costatare se il profilo dei pagamenti (e anche degli impegni) in corso di gestione è coerente (compatibile) con il livello finale autorizzato per la fine dell’anno. Tradizionalmente, il vincolo effettivamente considerato stringente in fase di gestione del bilancio è quello sugli impegni, mentre i pagamenti tendono a seguirne il profilo. Tuttavia accade regolarmente che, a fine anno, al momento della chiusura del bilancio (3° fase), esista ancora un ammontare rilevante di impegni che non abbia dato luogo a pagamenti. In tal caso si verifica la formazione dei cosiddetti residui, che costituiscono un’ulteriore grandezza registrata nei sistemi contabili²². Di conseguenza, durante ciascun esercizio finanziario, una parte dei pagamenti avverrà non in funzione degli impegni presi nel corso della gestione ma degli impegni presi durante gli anni precedenti e non tramutatisi in pagamenti (i cosiddetti pagamenti in conto residui). Data la rilevanza e l’entità dei residui normalmente accumulati, è importante seguire con altrettanta attenzione anche il comportamento di tali variabili.

Quando tutta l’informazione relativa all’esercizio finanziario è disponibile, viene registrata come fonte ufficiale per il valore annuale definitivo di tutte le variabili di bilancio. Tutte le grandezze rilevanti sono quindi prospettate all’interno

²¹ Questa affermazione verrà meglio articolata nel paragrafo conclusivo.

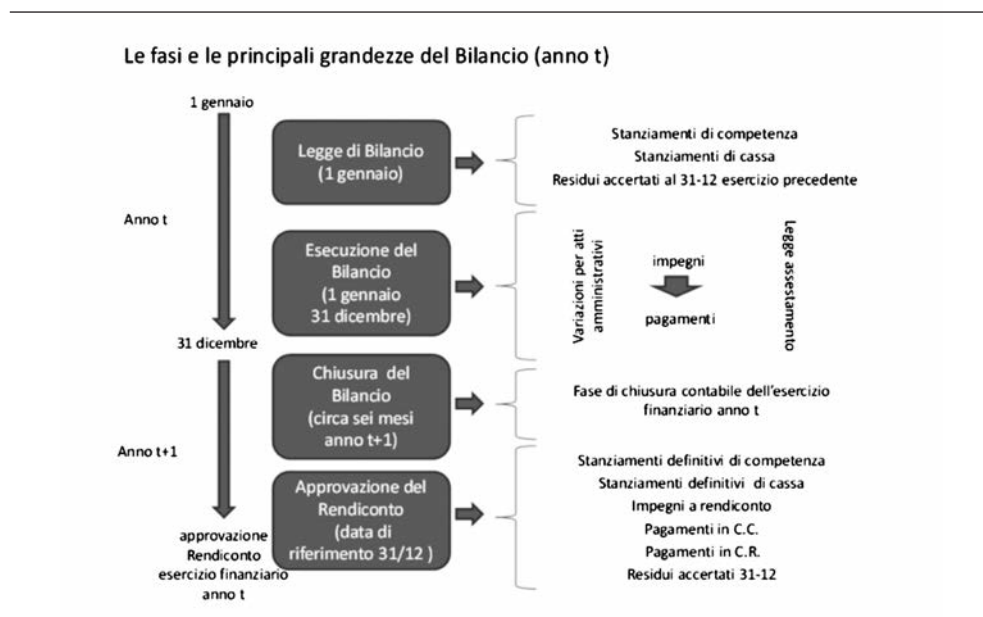
²² I residui che si formano come risultato di impegni che non hanno ancora dato luogo a pagamenti sono definiti residui propri. Esiste una seconda categoria di residui che vengono generati da stanziamenti che non hanno ancora dato luogo ad impegni; in questo caso però non esiste un automatismo, occorre uno specifico provvedimento amministrativo di “conservazione” dei residui. Questa seconda categoria di residui è definita con il termine di residui impropri.

del Rendiconto generale dello Stato (4° fase). In esso sono riassunti e mostrati i risultati della complessiva gestione svolta nell'anno finanziario, con distinto e simultaneo riferimento alle gestioni di competenza, di cassa e dei residui (conto del bilancio), e delle variazioni intervenute nel patrimonio dello Stato per effetto della gestione del bilancio o anche per altre cause (conto generale del patrimonio). Tuttavia, trattandosi di variabili i cui valori si osservano *ex-post*, diversi mesi dopo la chiusura dell'esercizio, il Rendiconto non rappresenta una fonte d'informazioni rilevante per la modellazione che si propone di effettuare.

La Figura 1 mostra le quattro fasi del bilancio descritte. Nel lavoro sono state considerate le grandezze dal momento dell'approvazione della Legge di bilancio, quindi gli stanziamenti di competenza e cassa e i residui previsti. Durante il periodo di esecuzione del bilancio²³, oltre a considerare le variazioni agli stanziamenti per atti amministrativi e per Legge di assestamento sono stati utilizzati gli impegni, i pagamenti di competenza e quelli di cassa. Nella fase di rendicontazione sono state considerate le variabili riferite agli stanziamenti definitivi di competenza e di cassa, gli impegni a rendiconto, i pagamenti di competenza, di cassa e i residui al 31 dicembre dell'esercizio finanziario.

FIG. 1

LE FASI DEL BILANCIO E LE PRINCIPALI GRANDEZZE



²³ Per il dettaglio delle variabili utilizzate si veda la Tavola 1 del paragrafo 4.

I dati utilizzati nel presente lavoro provengono dalle Basi Dati della RGS. Prima di analizzare le serie storiche, riteniamo utile chiarire, sommariamente, la struttura degli archivi e le loro funzionalità.

La RGS, per la sua missione istituzionale, si è dotata nel tempo di una serie di sistemi gestionali. Per la gestione del bilancio lato spese, oggetto del presente lavoro, essa ha a disposizione un insieme di dati provenienti da più sistemi gestionali che, a loro volta alimentano il Data Warehouse²⁴ RGS. Quest'ultimo ha il compito, tra l'altro, di integrare e rendere coerenti i dati provenienti da diverse fonti. Per quanto riguarda i dati dei pagamenti, questi, oltre che provenire dai sistemi gestionali, affluiscono anche da una fonte complementare e cioè dalla Tesoreria. Mentre i dati relativi ai pagamenti di Tesoreria possiedono un ciclo annuale, e quindi si riferiscono all'anno solare in quanto sono registrati nel momento in cui vengono eseguiti, i sistemi gestionali della Ragioneria registrano a volte i dati, ad esempio dei pagamenti, successivamente al momento in cui viene effettivamente effettuata la transazione. Accade quindi che, per alcuni aggregati, la registrazione delle transazioni effettuate avvenga nei mesi successivi al termine dell'esercizio finanziario. Nel momento della chiusura del Rendiconto tutte le transazioni sono state registrate.

Le serie storiche mensili analizzate nel paragrafo successivo ai fini della costruzione del modello di previsione e simulazione dei pagamenti sono: gli stanziamenti iniziali nel seguito denominati CP_LB e le relative variazioni infra annuali (CP), gli impegni in conto competenza (CP_IMP), la disponibilità in conto competenza (CP_DIS)²⁵ ovvero la quantità dello stanziamento ancora utilizzabile a sua volta scomponibile in disponibilità di fondo (CP_DIS_F) e disponibilità non di fondo (CP_DIS_NF)²⁶, i pagamenti in conto competenza (CP_PAG) e in conto residui (RS_PAG), i residui propri (RS_AL) e impropri (RS_STAN) Per maggiore chiarezza la tavola seguente riporta le sigle con la descrizione delle serie storiche utilizzate nel lavoro.

²⁴ Il DW RGS adotta un modello architetturale "Corporate", che prevede la realizzazione di un Enterprise Data Warehouse (EDW) centralizzato, con dati univoci, ottimizzati, razionalizzati e di elevato dettaglio, rispondente al modello dati aziendale e da più Data Mart, ambiti di maggiore aggregazione, alimentati dall'EDW e messi a disposizione secondo schemi dati multidimensionali e specializzati per aree tematiche.

²⁵ Essa è definita come Stanziamento Definitivo di competenza (+Assegnato in C/C nelle interrogazioni che prevedono la Ragioneria) – Totale Accantonato C/C – Impegnato C/C – Conservazione di Fondi.

²⁶ La differenza deriva dall'esistenza all'interno del bilancio dello Stato di capitoli su cui possono essere emessi mandati di pagamento (CP_DIS_NF) e altri capitoli per i quali non si può esercitare tale opzione (CP_DIS_F).

TAV. 1

DESCRIZIONE DELLE VARIABILI DI BILANCIO UTILIZZATE NEL MODELLO

Variabili a frequenza mensile	
CP_LB	Stanziamenti di competenza assegnati nella legge di bilancio (stanziamenti iniziali)
CP	Variazioni agli Stanziamenti di competenza. Rappresenta il volume delle variazioni all'istante t
S_TOT_CP	Stanziamenti iniziali di bilancio + variazioni in corso d'anno
CP_IMP	Impegni in conto competenza. Rappresentano il volume totale degli impegni all'istante t
CP_PAG	Pagamenti in conto competenza
RS_PAG	Pagamenti in conto residui
CS_PAG	Pagamenti di cassa (Pagamenti in c. competenza + Pagamenti in c. residui)
RS_ACC_AL	Residui propri accertati
RS_AL	Residui propri. Sono dati dalla somma dei residui correnti derivanti: dal trasporto degli Ordini di Pagare e degli Ordini di Accreditamento, dai Ruoli di Spesa Fissa insoluti e per i quali risulta emesso un impegno, dalle spese impegnate e non ordinate dall'Amministrazione centrale ovvero dai Funzionari Delegati e spese ordinate dai Funzionari Delegati e non pagate e dalle spese ordinate dall'Amministrazione centrale con impegno contemporaneo, non ordinate dai Funzionari Delegati ovvero ordinate ma non pagate.
RS_STAN	Residui impropri o di stanziamento. Riguardano stanziamenti di spese, in genere di conto capitale, che, non impegnate alla chiusura dell'esercizio, vengono tuttavia fatte transitare nel conto dei residui. Non costituiscono debiti per lo Stato
ST_RS_ACC_AL	Stock di residui accertati all'istante t . Sono dati dalla differenza tra residui propri accertati e residui pagati
CP_PAG_BKI	Pagamenti in conto competenza provenienti dalla Tesoreria
RS_PAG_BKI	Pagamenti in conto residui provenienti dalla Tesoreria
CS_PAG_BKI	Pagamenti di cassa provenienti dalla Tesoreria
Ulteriori variabili analizzate e non presenti nella versione attuale del modello (frequenza mensile)	
CP_DIS	Disponibilità di competenza (differenza tra stanziamento definitivo di competenza totale accantonato in conto competenza e impegni in conto competenza)
CP_DIS_F	Fa riferimento a capitoli su cui non possono essere emessi mandati di pagamento
CP_DIS_NF	Fa riferimento a capitoli su cui possono essere emessi mandati di pagamento

Una caratteristica dell'informazione che si estrae dalla suddetta base dati è che la rappresentazione del dato contabile si estende su di un orizzonte di circa 18 mesi: successivamente alla chiusura dell'esercizio finanziario di riferimento, la

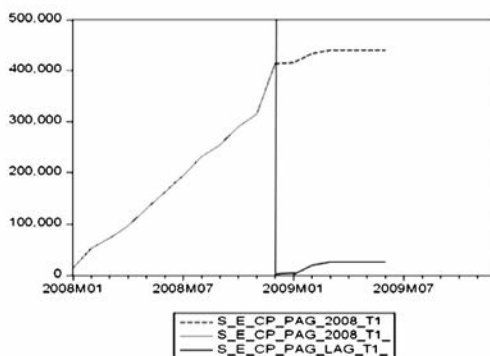
contabilità relativa all'anno resta aperta perché, come spiegato, si continuano a recepire/registrare informazioni sulle transazioni avvenute.²⁷

In termini di costruzione di serie storiche la scelta è stata quella di scomporre le 18 osservazioni relative a ciascun anno in due blocchi. Si veda, come esempio, il Grafico 1. Il primo blocco – a sinistra della linea verticale – comprende tutti i valori osservati nei primi 12 mesi (osservati cioè nel corso dell'esercizio finanziario), il secondo blocco comprende i valori osservati successivamente alla chiusura dell'anno per i successivi sei mesi (linea tratteggiata). Il livello della variabile nel mese di gennaio successivo è pari al livello a fine anno incrementato del valore registrato nel mese di gennaio. Le due variabili “derivate” da quella iniziale corrispondono alle due linee continue e la parte di linea tratteggiata corrisponde ai valori presenti nei sistemi negli ultimi sei mesi osservati.

Le serie storiche a carattere mensile sono ottenute “saldando” le variabili con dodici osservazioni (del primo blocco) tra i vari anni ovvero, considerando soltanto i valori osservati in corso d'anno. Ad esempio la serie storica dei pagamenti in conto competenza (CP_PAG) verrà ottenuta saldando le osservazioni della variabile CP_PAG per gli anni 2003, 2004, 2005 e così via. Le serie relative agli ultimi sei mesi (secondo blocco) sono utilizzate nell'ambito della modellizzazione annuale delle grandezze di bilancio e sono caratterizzate, in termini di codifica, dal suffisso _LAG. Il Grafico 1 mostra i valori delle variabili in questione relativi all'anno finanziario 2008 per il titolo 1 (da cui il suffisso finale).

GRAF. 1

VALORE DEFINITIVO DELLA VARIABILE PAGAMENTI IN CONTO COMPETENZA DEL TITOLO I RELATIVAMENTE ALL'ESERCIZIO FINANZIARIO 2008



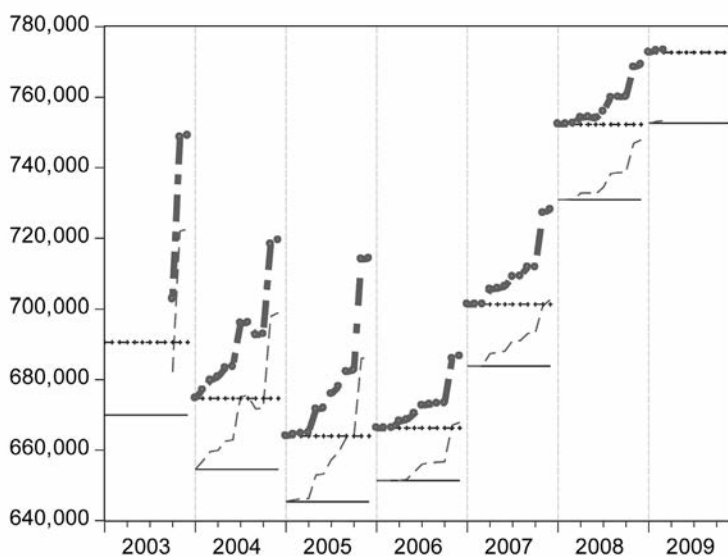
²⁷ I valori tendono a stabilizzarsi intorno al mese di marzo ma il dato definitivo che verrà ufficializzato si ottiene soltanto a chiusura del consuntivo.

Nei grafici seguenti si mostrano brevemente le serie storiche in livelli ottenute per gli stanziamenti e per gli impegni. Gli stanziamenti che si prendono in considerazione sono i valori totali, che risentono in quanto tali di diversi eventi nel corso dell'anno. In primo luogo si considerano gli stanziamenti iniziali che provengono dalla legge di bilancio, vi sono poi variazioni dovute a incrementi o decrementi legati alla assegnazione di fondi al taglio di stanziamenti nel corso dell'anno ed al provvedimento di assestamento.²⁸

Il Grafico 2 mostra gli stanziamenti iniziali (in conto competenza e di cassa) e i valori totali stanziati nel corso dell'anno ottenuti come somma dei valori iniziali (linee orizzontali) e delle successive variazioni. Il Grafico 3 mostra lo *stock* del totale degli impegni in conto competenza. I valori sono espressi in milioni di euro.

GRAF. 2

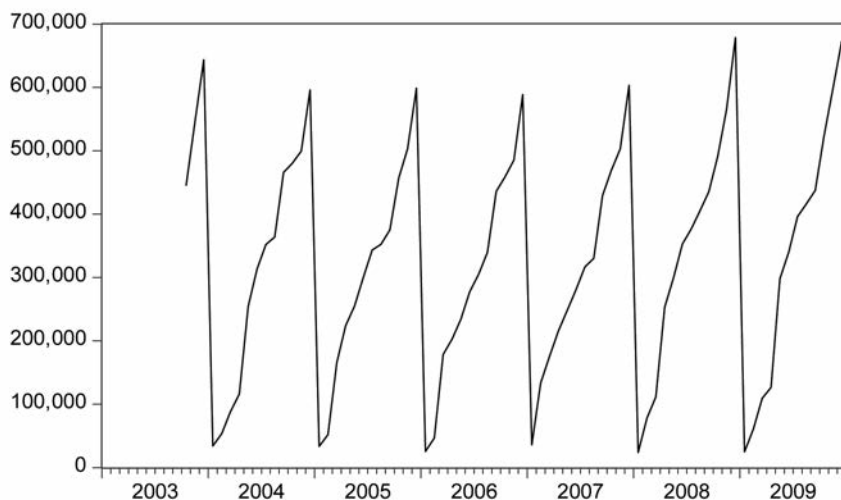
STANZIAMENTI DI COMPETENZA DI CASSA



²⁸ Si ricorda che l'analisi si concentra sul bilancio dello Stato nel suo insieme e che, dunque, sia per questa grandezza che per le successive, le riallocazioni di stanziamenti tra diversi ministeri sono "neutrali".

GRAF. 3

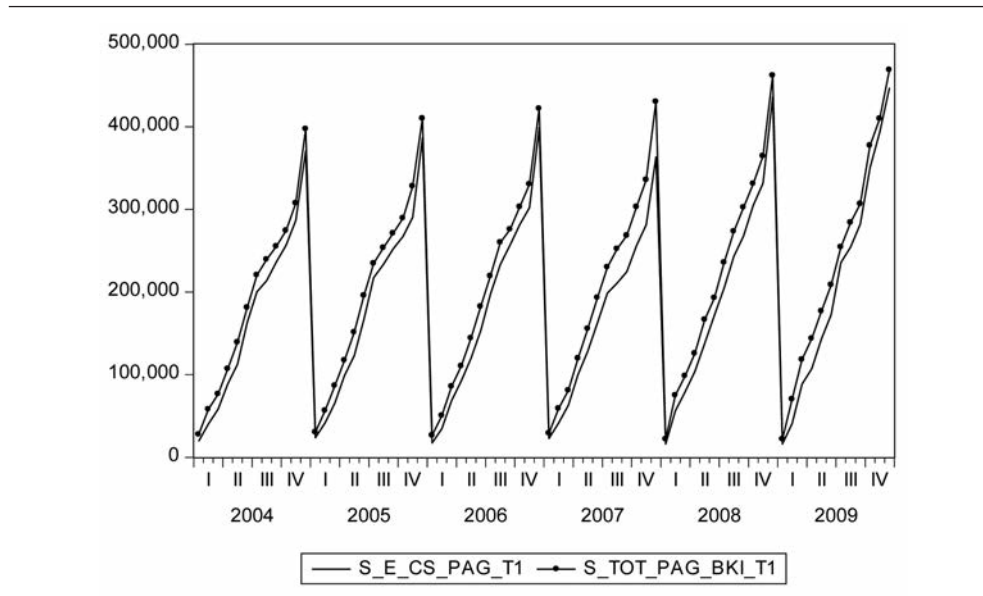
IMPEGNI IN CONTO COMPETENZA



La variabile di cui si vuole prevedere l'andamento sono i pagamenti dal bilancio dello Stato. Di essa esistono nei sistemi informatici della RGS due versioni: una è di fonte interna e proviene dalle rilevazioni dei sistemi gestionali, l'altra è fornita dalla Tesoreria (entrambe le variabili sono rappresentate nei Grafici 4 e 5). Più nel dettaglio, il Grafico 4 confronta i pagamenti mensili di cassa di fonte RGS provenienti dalla gestione ed il flusso dei pagamenti di tesoreria relativamente al titolo I. Il Grafico 5 confronta le due serie relativamente al 2008 considerando però per i pagamenti RGS anche le registrazioni di operazioni riferite all'esercizio finanziario e contabilizzate nei 6 mesi successivi alla fine dell'anno finanziario.

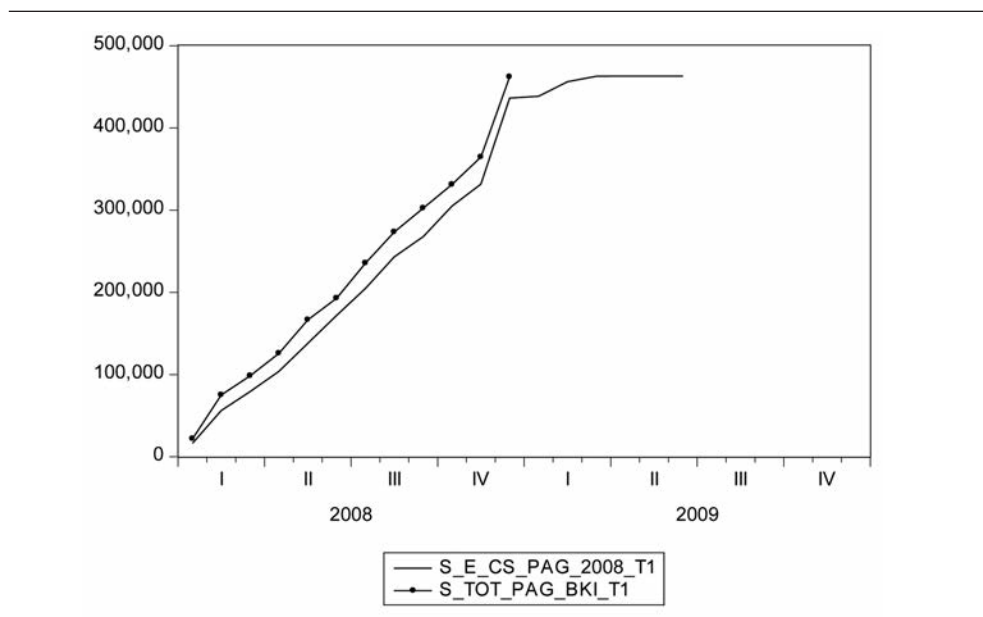
GRAF. 4

% PAGAMENTI DI CASSA DAL BILANCIO DELLO STATO E DALLA TESORERIA
(spese correnti)



GRAF. 5

PAGAMENTI DI GESTIONE E DI TESORERIA NEL 2008
(spese correnti)

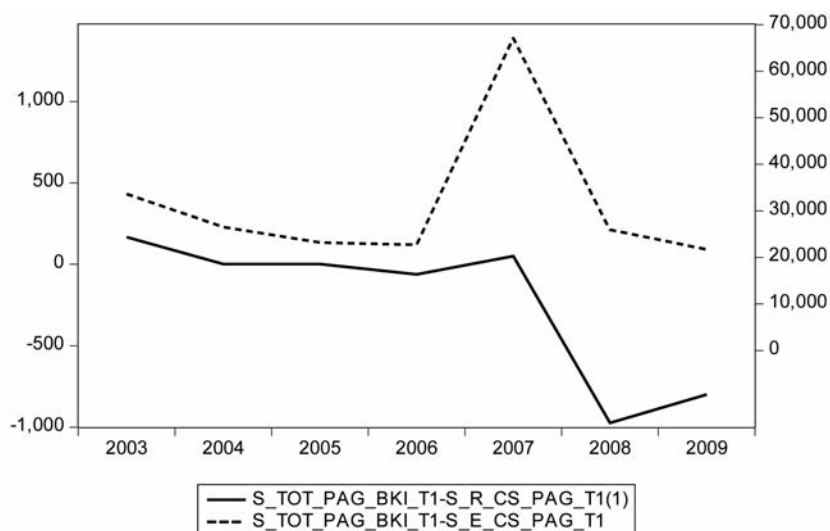


Osservando il Grafico 4 si evince che i pagamenti RGS risultano costantemente al di sotto di quelli di tesoreria a causa di uno sfasamento temporale nella registrazione tra i due sistemi contabili. A causa di tale sfasamento temporale, il valore dei pagamenti RGS raggiunge quello di fonte Tesoreria soltanto nei mesi successivi alla chiusura dell'esercizio. Tale caratteristica si evince in modo più chiaro nel Grafico 5 in cui i dati di fonte RGS, raggiungono (a meno di una differenza, in termini percentuali, trascurabile) quelli Tesoreria, nei sei mesi successivi alla chiusura dell'anno finanziario.

Per osservare più attentamente l'entità degli scostamenti evidenziati nei Grafici 4 e 5 il Grafico 6 riporta le differenze tra la serie i pagamenti di tesoreria 31 dicembre (s_{tot_bki}) e la serie dei pagamenti di gestione RGS al 31 dicembre ($s_{e_cs_pag}$) e la differenza tra serie storica dei pagamenti di tesoreria al 31 dicembre e la serie storica annuale dei dati di rendiconto ($s_{r_cs_pag}$) in corrispondenza al valore riferito al 18° mese.

GRAF. 6

DIFFERENZE TRA LIVELLO DI FONTE TESORERIA AL 31 DICEMBRE E VALORI DEI PAGAMENTI DESUNTI DAL RENDICONTO A CONSUNTIVO
(valori in milioni di Euro)



Come si nota, la differenza tra il valore dei pagamenti di tesoreria e il valore dei pagamenti di bilancio (linea tratteggiata - scala a destra), varia notevolmente nel corso degli anni e non esiste una tendenza di cui si possa eventualmente tenere conto in sede di analisi previsiva con uno scostamento anomalo nel 2007. Tale scostamento potrebbe essere in parte dovuto a un incremento dei pagamenti effettuati con strumenti diversi da quelli registrati nel circuito RGS rispetto a quelli registrati in tesoreria e per tale ragione da esso non intercettati. Lievi scostamenti permangono anche confrontando le differenze dei pagamenti di tesoreria con i dati di rendiconto (linea continua - scala a sinistra). La scala di queste ultime differenze è però notevolmente inferiore.

Nonostante la maggiore tempestività insita nel meccanismo di registrazione dei pagamenti tesoreria, come verrà mostrato successivamente, nelle equazioni del modello si stimano e prevedono i pagamenti di fonte RGS in quanto presentano raccordi contabilmente coerenti con tutte le altre voci presenti nei sistemi informativi della Ragioneria. A partire dalle previsioni dei pagamenti RGS il profilo dei pagamenti di tesoreria viene poi tracciato tramite l'utilizzo di equazioni di raccordo che sostanzialmente stimano e prevedono il dato di tesoreria a partire dai pagamenti RGS.

5. - Le serie storiche dei flussi: stagionalità dei dati, prime analisi sulle interazioni tra variabili di bilancio e ulteriori caratteristiche

Dopo la breve illustrazione delle principali variabili di bilancio in termini di livelli (*stock*), nel paragrafo si presenta un approfondimento relativo alle caratteristiche dei dati di flusso, che sono poi quelli utilizzati nelle stime econometriche dei modelli, ottenuti decumulando gli *stock* delle serie storiche mensili del bilancio.²⁹

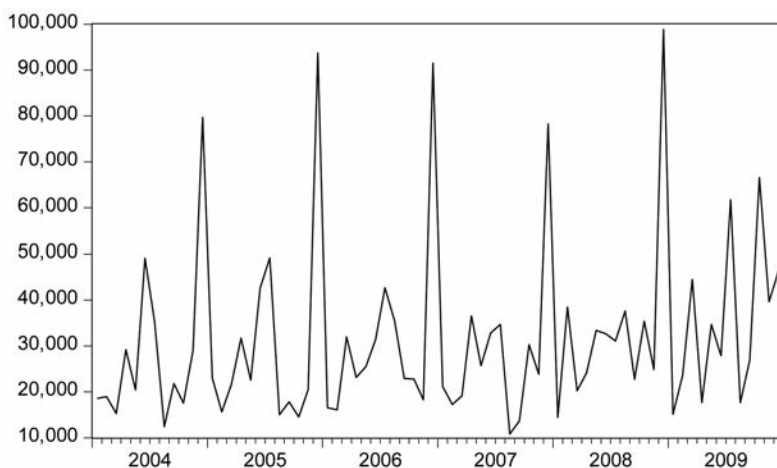
Una prima caratteristica di cui tener conto in fase di modellizzazione delle serie di bilancio è la presenza di picchi di stagionalità in corrispondenza di alcuni particolari mesi dell'anno. La stagionalità rilevata è connaturata all'annualità stessa del bilancio dello Stato; al suo determinarsi concorrono alcune prassi e regolamentazioni o indicazioni legislative e la presenza di scadenze di tipo amministrativo che portano al concentrarsi di stanziamenti, impegni o pagamenti in

²⁹ Ovviamente, data la natura contabile del dato, i valori della serie di flusso e di *stock* coincideranno per il mese di gennaio. L'analisi dei dati mensili dal punto di vista econometrico può essere effettuata però soltanto ricostruendo il dato di flusso relativo a ciascun mese a partire dalle serie degli *stock*.

determinati mesi dell'anno. Al fine di evidenziare la presenza di stagionalità in tali serie, i Grafici 7a, 7b, 7c e 7d riportano in via esemplificativa, i flussi mensili, dei pagamenti in conto competenza (cp_pag) mentre i Grafici 8a, 8b, 8c e 8d dei pagamenti in conto residui (rs_pag); in entrambi i casi si espongono i dati rispettivamente del titolo I, del titolo II, del titolo III e totali.

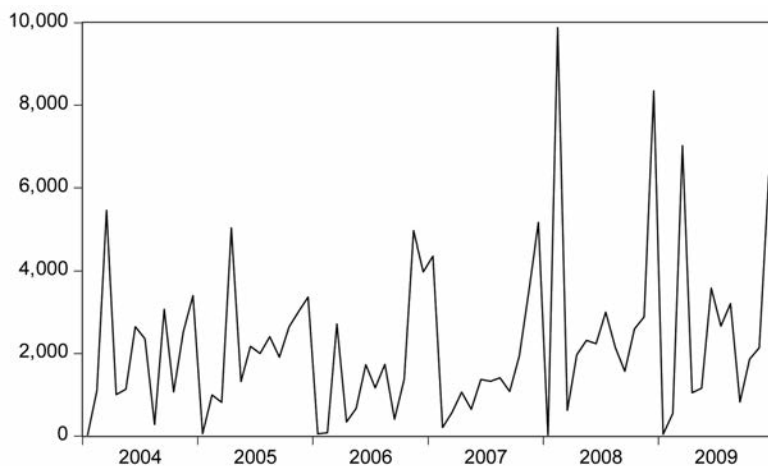
GRAF. 7a

FLUSSI PAGAMENTI IN C.C. TITOLO I



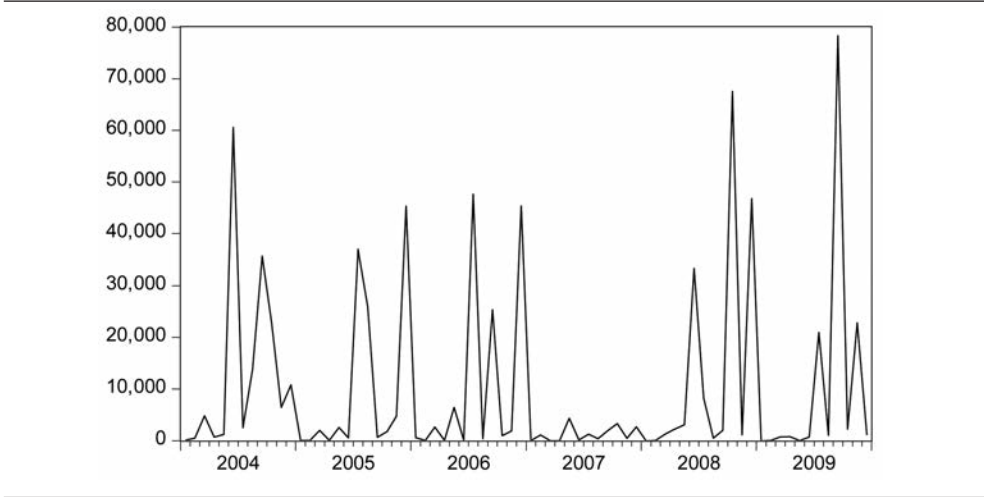
GRAF. 7b

FLUSSI PAGAMENTI IN C.C. TITOLO II



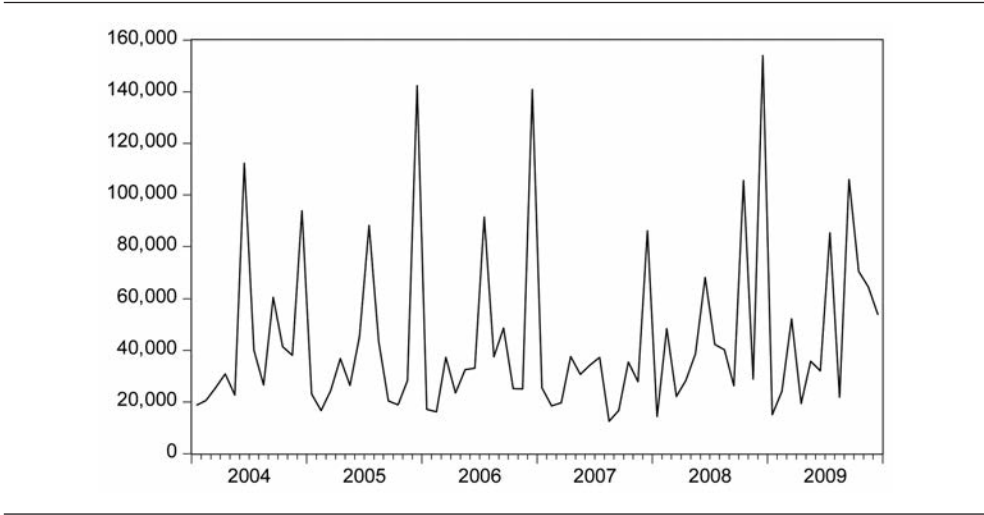
GRAF. 7c

FLUSSI PAGAMENTI IN C.C. TITOLO III



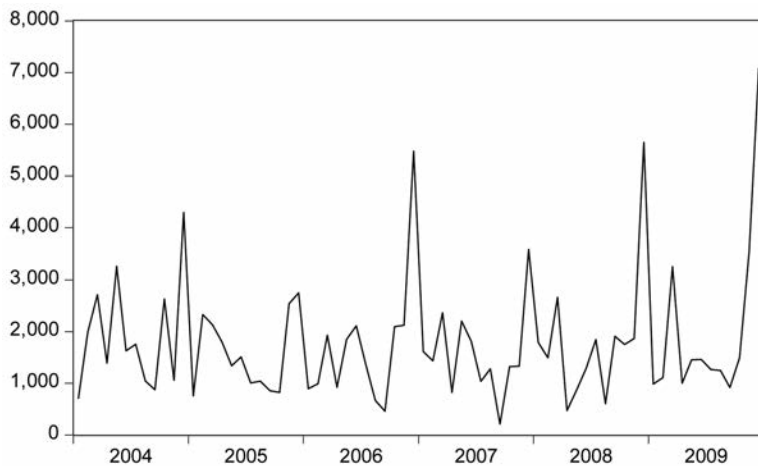
GRAF. 7d

FLUSSI PAGAMENTI IN C.C. TOTALI



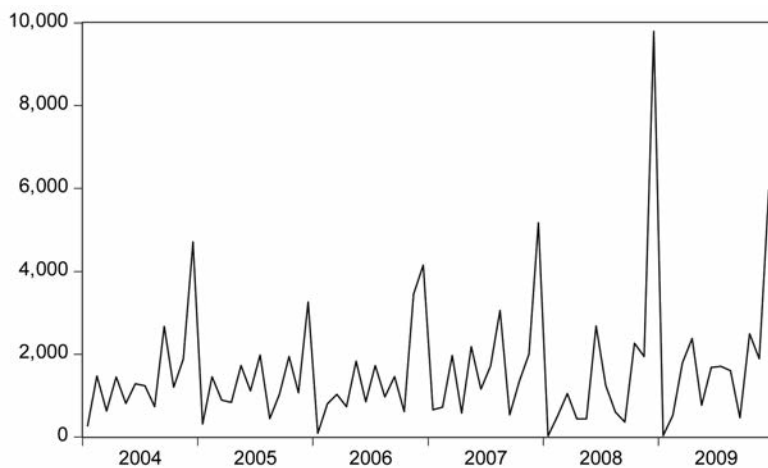
GRAF. 8a

FLUSSI PAGAMENTI IN C.R. TITOLO I



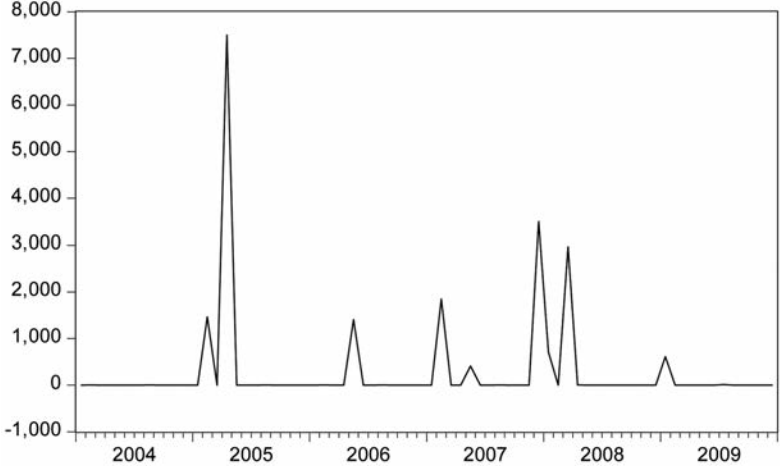
GRAF. 8b

FLUSSI PAGAMENTI IN C.R. TITOLO II



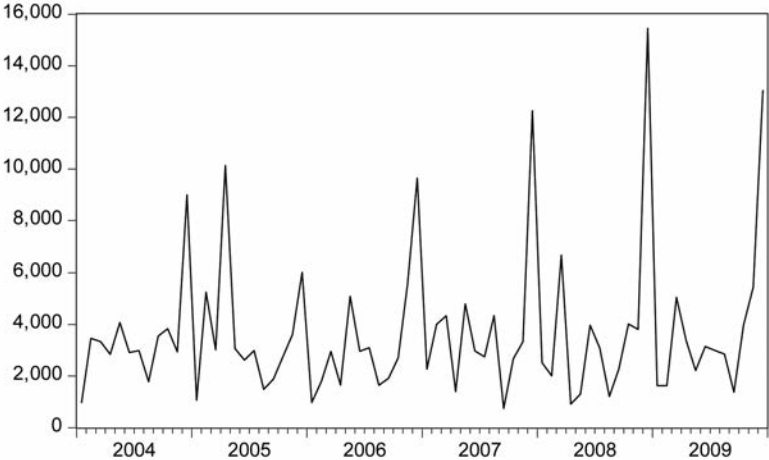
GRAF. 8c

FLUSSI PAGAMENTI IN C.R. TITOLO III



GRAF. 8d

FLUSSI PAGAMENTI IN C.R. TOTALI



L'analisi grafica mostra, per tutte le serie dei pagamenti in conto competenza, una presenza di picchi di stagionalità. Disaggregando ulteriormente le serie storiche a livello di titoli in categorie di spesa,³⁰ è possibile individuare le voci di bilancio che concorrono in modo più significativo a determinare la stagionalità. Per il titolo I (spese correnti), ad esempio, emerge che la stagionalità è particolarmente determinata da picchi stagionali nei consumi intermedi, nei trasferimenti correnti ad amministrazioni pubbliche e negli interessi passivi. Anche la dinamica dei pagamenti in conto residui evidenzia forti componenti stagionali con un aumento dei picchi in corrispondenza della fine dell'anno. Tale evidenza per il titolo I può essere spiegata dal fatto che i residui dei pagamenti comprendono voci contrattuali (es. Fondo Unico di Amministrazione) il cui pagamento avviene di solito con ritardi "regolari" rispetto all'esercizio finanziario di competenza. Per altri tipi di variabili di bilancio (disponibilità, impegni, ecc.) per le quali per ragioni di brevità non si riportano i grafici, la stagionalità appare spesso evidente anche negli altri due titoli.³¹

Al fine di verificare in modo più formale la presenza di stagionalità nelle serie storiche dei pagamenti di bilancio, sono stati inoltre condotti *test* di stagionalità sui dati. L'applicazione di tali *test* alle serie dei pagamenti ha identificato la presenza di stagionalità confermando quanto già evidenziato nell'analisi grafica. Data la presenza di picchi stagionali nei dati, nel modello econometrico di previsione e stima delle spese di bilancio che verrà presentato nel paragrafo 6, sono state stimate equazioni basate sulle differenze dodicesime (variazioni tendenziali). L'uso di serie differenziate con un ordine corrispondente al tipo di stagionalità consente infatti di rimuovere la presenza di radici stagionali nelle variabili oggetto di previsione, permettendo di riottenere agevolmente la previsione sui livelli (valori assoluti) a partire da quella sulle differenze.

Si è visto nel paragrafo 3 che, la maggioranza dei contributi empirici, ha cercato di modellare l'evoluzione dei flussi di bilancio sulla base di modelli autoregressivi. I tentativi effettuati nel presente lavoro si sono proposti di superare l'approccio tradizionale e arrivare a stime dove le equazioni includessero oltre ai ritardi della variabile dipendente, ulteriori regressori – in altri lavori definiti "variabili d'intervento" – interni al bilancio dello Stato. Nell'analizzare le serie storiche, l'approccio seguito è consistito nel valutare (anche empiricamente) se, all'interno del ciclo del bilancio dello Stato, esistessero variabili che potessero essere utili per pre-

³⁰ I grafici relativi alle categorie non vengono riportati per brevità.

³¹ È il caso degli impegni per i quali, i picchi di stagionalità, si riscontrano anche nel titolo III.

vedere i flussi dei pagamenti. Ci si è interrogati ad esempio sull'opportunità di utilizzare l'informazione relativa al mese t sugli impegni per prevedere i pagamenti nei mesi successivi; oppure sulla possibilità di prevedere a loro volta gli impegni, utilizzando l'informazione sugli stanziamenti. Si è cercato anche di valutare il ruolo giocato da altre grandezze del bilancio (disponibilità, residui, ecc.) nell'aumentare la capacità previsiva dei pagamenti.

Al fine di individuare i potenziali regressori da utilizzare nel modello di previsione delle spese, è stata effettuata un'analisi esplorativa preliminare delle cross correlazioni tra le principali serie storiche ed alcune loro trasformazioni elementari di bilancio suddivise per titoli³². I risultati relativi alle cross correlazioni dei dati di bilancio con i pagamenti forniscono nel complesso due indicazioni principali. In primo luogo per i primi due titoli esiste una rilevante correlazione contemporanea tra pagamenti ed impegni; in secondo luogo, la struttura dei ritardi di altre variabili del bilancio non sembra potenzialmente in grado di aumentare in maniera rilevante il potere predittivo di equazioni dei pagamenti. Riguardo alla relazione tra pagamenti e impegni con il primo aspetto si pone innanzitutto la questione di effettuare ulteriori approfondimenti riguardante il nesso di causalità tra le due serie³³. Esiste inoltre un ulteriore aspetto dal punto di vista della modellizzazione della serie:

³² I risultati delle cross correlazioni sono riportati in *APPENDICE*.

³³ *Test* preliminari di Causalità nel senso di Granger tra serie di bilancio relative alle diverse fasi (impegni, pagamenti), sia relativamente alle variabili in conto competenza che in conto residui, hanno mostrato assenza di causalità degli impegni verso i pagamenti per tutti e tre i titoli di spesa. L'assenza di un nesso di causalità nel senso di Granger degli impegni verso i pagamenti, non implica ad ogni modo un'assenza di un nesso di causa effetto tra le due variabili ma riflette semplicemente la mancanza di una causalità di tipo "sequenziale" che rappresenta il presupposto del *test* stesso. A tal proposito SPRITES P. *et AL.* (1993) e PEARL J. (2000) hanno proposto altri tipi di *test* di analisi di causalità "in contemporanea" che prescindono dalla presenza di una sequenza temporale. Tale concetto di causalità si applica in particolare per quei dati che vengono registrati in un momento predefinito ma che astrae dal preciso momento storico in cui si è verificato l'evento. Un tale *framework* è coerente anche con quanto avviene per i dati del bilancio la cui registrazione contabile spesso prescinde dal momento in cui si produce il corrispondente effetto sulle altre variabili. L'utilizzo di *test* di causalità di questo tipo potrebbe dunque risultare più appropriato per l'analisi in questione. Il risultato del *test* di Granger, apparentemente contro-intuitivo, può essere inoltre spiegato dal fatto che i *test* sono stati condotti a livello di titoli di spesa e che all'interno dei titoli esistono differenti sfasamenti tra impegni di spesa e relativi pagamenti a seconda delle categorie di spesa. Una parte rilevante dei pagamenti del titolo I, circa il 40%, è composta dai ruoli di spesa fissa quali ad esempio gli stipendi il cui pagamento avviene contestualmente al relativo impegno. In aggregato (a livello di titolo) la sequenzialità del pagamento rispetto all'impegno può non apparire così evidente e produrre un risultato di non significatività statistica del *test* di causalità nel senso di Granger.

l'elevata correlazione contemporanea tra pagamenti e impegni potrebbe essere proficuamente utilizzata qualora fosse possibile prevedere con precisione gli impegni (tale aspetto verrà approfondito nel prossimo paragrafo).

In secondo luogo l'analisi delle correlazioni riportata in appendice fornisce esclusivamente un'indicazione dei possibili legami "elementari" esistenti tra variabili di bilancio poiché riporta soltanto le correlazioni delle differenze dodicesime dei pagamenti con alcune grandezze di bilancio trasformate in differenze dodicesime oppure ottenute a partire da aggregazioni tra indicatori di bilancio. Come si vedrà, nelle equazioni del modello, sono state incluse anche altre trasformazioni delle variabili di bilancio, ottenute a partire dalle variabili base che nelle analisi esplorative non vengono riportate. In sostanza i regressori scelti (più approfonditamente descritti nel paragrafo 6) vengono selezionati sia in base alla significatività econometrica, sia in base conoscenze a priori sui loro legami di tipo contabile (legate al ciclo di bilancio) con i pagamenti.

6. - Il modello econometrico

In questa sezione viene presentato il modello econometrico multiequazionale di previsione e modellizzazione dei flussi delle spese di bilancio. Il primo paragrafo introduce le equazioni stimate da noi definite "pseudo" comportamentali in quanto riproducono non un comportamento di tipo economico ma di tipo amministrativo legato al ciclo del bilancio e stimate per le singole voci di bilancio, mentre il secondo paragrafo riporta le principali identità contabili del modello.

6.1 Equazioni "pseudo" comportamentali

Nel lavoro di costruzione del modello spese si è proceduto alla stima di equazioni dove le variabili dipendenti sono le variazioni dei flussi degli stanziamenti ($f_{s_e_cp}$), degli impegni ($f_{s_e_cp_imp}$) e dei pagamenti ($f_{s_e_cp_pag}$) in conto competenza ed ($f_{s_e_rs_pag}$) in conto residui per ciascuno dei tre titoli. Le variabili previste dal modello sono, come già accennato nel paragrafo 2, le differenze dodicesime dei flussi (suffisso Δ^{12}), mentre i regressori utilizzati nelle equazioni sono espressi come differenze dodicesime, flussi oppure *stock* (indicati con l'apice *S* sopra la variabile).

Le specificazioni a cui si è giunti, coerentemente con l'intento sul quale si è impostato il lavoro di ricerca, hanno portato ad un superamento dell'utilizzo delle sole variabili dipendenti ritardate ed alla introduzione all'interno di diverse equa-

zioni di almeno una variabile di “interazione” e di *feedback* il cui valore è aggiornato all’interno della struttura del modello. Soltanto per quanto riguarda le variazioni degli stanziamenti di bilancio (stanziamenti infra-annuali) per il momento non si è individuata un’ulteriore variabile esplicativa e quindi la struttura statistica è rimasta inalterata rispetto ad equazioni autoregressive considerate nel lavoro come *benchmark*. D’altra parte, come illustrato nel terzo paragrafo, l’importo delle variazioni degli stanziamenti risulta molto contenuto rispetto ai valori stabiliti dalla legge di bilancio.

Nel caso degli impegni la variabile d’interazione che si è aggiunta alla struttura autoregressiva è la variazione rispetto all’anno precedente del rapporto tra impegni effettuati fino al mese precedente e il totale degli stanziamenti. Questa variabile agisce in maniera simile a un termine di correzione dell’errore: qualora il valore della variabile risulti inferiore rispetto al valore registrato nello stesso periodo dell’anno precedente il valore predetto dall’equazione degli impegni per il mese successivo aumenta. L’idea di fondo è quindi che il valore nel corso dell’anno del rapporto assuma un profilo abbastanza stabile, valori anomali di tale rapporto tenderanno ad essere compensati nei mesi successivi. Il valore tenderà a 1 nell’ultimo mese, senza però mai arrivarci (una parte degli stanziamenti non viene impegnata).

Per quanto riguarda le equazioni sulle variabili di pagamenti in conto competenza si sono introdotti due nuovi regressori: le variazioni contemporanee degli impegni e uno o più termini ritardati tra i valori dei livelli dello scostamento d’impegni e pagamenti in corso d’anno (*cp_imp-cp_pag*).

Sull’introduzione del termine contemporaneo degli impegni va rilevato che questo è essenziale per migliorare la *performance* previsiva delle equazioni di pagamenti. La scelta si giustifica con il fatto che l’impegno, per alcune categorie³⁴, temporalmente simultaneo al pagamento (da qui l’esigenza di introdurre il termine contemporaneo), ne costituisce dal punto di vista contabile e giuridico il presupposto. Si è consapevoli che l’introduzione di tale regressore può risultare premiante ai fini della previsione dei pagamenti soltanto se gli errori commessi nel prevedere gli impegni non portino ad un peggioramento complessivo delle previsioni³⁵.

Per quanto riguarda i pagamenti in conto residui in questa versione del modello sono state inserite equazioni *benchmark* autoregressive indicate con BM.

³⁴ Si precisa che il peso ‘finanziario’ di tali categorie è elevato (ad esempio le spese per il personale).

³⁵ Questa scelta è oggetto di valutazione successivamente nelle simulazioni che confrontano le capacità previsive del modello stimato contro quelle del modello *benchmark* statistico.

L'utilizzo dei termini di residui ritardati (sia in flussi sia in *stock*) non ha per il momento consentito di conseguire dei miglioramenti. Su quest'aspetto sono in corso ulteriori approfondimenti.

Per la maggior parte delle equazioni sono state inoltre introdotte delle variabili *dummy* nelle stime per tenere conto degli effetti di particolari episodi che hanno portato a variazioni inattese della crescita tendenziale dei pagamenti.³⁶ La Tavola 2 riporta le equazioni stimate con i corrispondenti valori dei coefficienti e degli *standard errors* suddivise per titoli.

Data l'esistenza di scostamenti tra le due variabili di cui si è dato conto nel paragrafo 4. I valori dei pagamenti di bilancio sono inoltre raccordati all'interno del modello con quelli di tesoreria tramite ulteriori equazioni. Queste sono contraddistinte da un elevato potere predittivo (gli R^2 sono prossimi a 1) dal momento che i due gruppi di variabili sono altamente correlati.

Tale scostamento risulta, in parte, dovuto alla diversa natura del dato utilizzato per effettuare le previsioni mensili dei pagamenti di Tesoreria rispetto al valore di rendiconto. Il modello spese utilizza, infatti, come già detto, i dati dei flussi mensili del "sistema informativo bilancio" per prevedere i pagamenti di cassa. Il contenuto informativo di tale previsione viene poi usato per prevedere i pagamenti di tesoreria tramite un'equazione di raccordo basata su una regressione dei pagamenti di bilancio su quelli di tesoreria. Il dato di tesoreria così previsto, viene successivamente confrontato con il valore di rendiconto che risulta disponibile solo a partire dal mese di giugno/luglio dell'anno successivo a quello di stima e che quindi incorpora i valori delle variabili usate nel modello rivisti fino a giugno.

La Tavola 2 riporta le stime dei coefficienti e degli *standard errors* delle equazioni dei modelli degli stanziamenti, impegni e pagamenti per i tre titoli state ottenute utilizzando lo stimatore OLS. Le usuali diagnostiche di stima mostrano inoltre normalità dei residui, assenza di autocorrelazione.³⁷

³⁶ Un esempio d'introduzione di variabile *dummy* riguarda l'equazione per i pagamenti in conto competenza del titolo I. Nella seconda metà del 2007 per motivi amministrativi si registrò un ritardo nella registrazione all'interno dei sistemi contabili dei pagamenti, da qui il segno negativo assunto dalla variabile *dummy*. Peraltro tale ritardo fu recuperato nel mese di gennaio del 2008.

³⁷ L'apice *S* posto sopra alla variabile sta ad indicare che si tratta di uno *stock*, ottenuto sommando i flussi mese per mese. L'apice *ST* sta invece ad indicare che si tratta di uno *stock* totale. Quest'ultimo caso riguarda esclusivamente la variabile stanziamenti per la quale al fine di ottenere lo *stock* occorre sommare allo stanziamento infra annuale anche lo stanziamento iniziale.

TAV. 2

EQUAZIONI STIMATE PER STANZIAMENTI, IMPEGNI E PAGAMENTI
(campione di stima 2005:1 2007:12)

Stanziamenti in conto competenza

$$(x_i = f_s_e_cp_ti \ i = 1, 2, 3)$$

- Titolo I $\Delta^{12}x_t = c(1) + c(2) * \Delta^{12}x_{t-12} + c(3) * x_{t-12}$
(BM) $c(1) = 88,15 (299,39) \ c(2) = -0,484 (0,100) \ c(3) = -0,358 (0,089)$
Estimated S.E. = 1365,558
- Titolo II $\Delta^{12}x_t = c(1) + c(2) * \Delta^{12}x_{t-12} + c(3) * x_{t-12}$
(BM) $c(1) = 326,035 (132,723) \ c(2) = -0,17 (0,124) \ c(3) = -0,75 (0,142)$
Estimated S.E. = 720,86
- Titolo III $\Delta^{12}x_t = c(1) + c(2) * \Delta^{12}x_{t-12} + c(3) * x_{t-12}$
(BM) $c(1) = -294,3291 (557,112) \ c(2) = -0,32 (0,092) \ c(3) = -0,60 (0,128)$
Estimated S.E. = 2604,04

Impegni in conto competenza

$$(y_i = f_s_e_cp_imp_ti \ i = 1, 2, 3)$$

- Titolo I $\Delta^{12}y_t = c(1) + c(2) * \Delta^{12}y_{t-12} + c(3) * y_{t-12} + c(4) (y_{t-1}^S/x_{t-1}^{ST} - y_{t-13}^S/x_{t-13}^{ST}) + c(5) * [(x_{t-1} - x_{t-2}) - (x_{t-2} - x_{t-3})] + c(6) * [(y_{t-3} - z_{t-3})/y_{t-3}^S]$
 $c(1) = 13913,5 (5142,001) \ c(2) = -0,436 (0,13) \ c(3) = -0,19 (0,09)$
 $c(4) = -145979,5 (65927,10) \ c(5) = 0,44 (0,2171) \ c(6) = -39631,92 (23143,62)$
Estimated S.E. = 8451,839
- Titolo II $\Delta^{12}y_t = c(1) + c(2) * \Delta^{12}y_{t-12} + c(3) * (\Delta^{12}y_{t-1} - \Delta^{12}y_{t-2}) + c(4) * (y_{t-3}^S/x_{t-1}^{ST} - y_{t-13}^S/x_{t-13}^{ST}) + c(5) * (y_{t-1}/(y_{t-13} - z_{t-13}))$
 $c(1) = -17436,428 (1094,876) \ c(2) = -0,41 (0,15) \ c(3) = 0,04 (0,12)$
 $c(4) = -17536,34 (6072,402) \ c(5) = 437,522 (460,274)$
Estimated S.E. = 1544,652
- Titolo III $\Delta^{12}y_t = c(1) + c(2) * \Delta^{12}y_{t-12} + c(3) * (y_{t-1}^S/x_{t-1}^{ST} - y_{t-13}^S/x_{t-13}^{ST})$
 $c(1) = 431,57 (3787,516) \ c(2) = -0,07 (0,11) \ c(3) = -172541,8 (39277,35)$
Estimated S.E. = 19181,32

Pagamenti in conto competenza

$$(z_i = f_s_e_cp_pag_ti \ i = 1, 2, 3) \ w_i = f_s_e_cp_pag_tif_s_e_cp_pag_bki_ti$$

- Titolo I $\Delta^{12}z_t = c(1) + c(2) * \Delta^{12}z_t + c(3) * \Delta^{12}z_{t-12} + c(4) * (z_{t-12}) + c(5) * y_{t-12} + c(6) * [(y_{t-1} - z_{t-1}) - (y_{t-7} - z_{t-7})] + c(7) * D_{2007,07-2007,12} + c(8) * (w_{t-1}^S - w_{t-13}^S)$
 $c(1) = 2375,645 (3275,059) \ c(2) = 0,4308 (0,122) \ c(3) = -0,47 (0,17) \ c(4) = -0,0497 (0,177)$
 $c(5) = -0,01 (0,2079) \ c(6) = 0,268 (0,1632) \ c(7) = -14715,14 (4384,159) \ c(8) = -50580,55 (46517,46)$
Estimated S.E. = 6404,742
- Titolo II $\Delta^{12}z_t = c(1) + c(2) * \Delta^{12}z_t + c(3) * \Delta^{12}z_{t-1} + c(4) * z_{t-12} + c(5) * y_{t-12} + c(6) * (y_{t-12} - z_{t-12})$
 $c(1) = -1235,426 (575,75) \ c(2) = 0,475 (0,07) \ c(3) = -0,126 (0,083)$
 $c(4) = -0,731 (0,147) \ c(5) = 0,22 (0,1) \ c(6) = 0,295 (0,09)$
Estimated S.E. = 815,2151
- Titolo III $\Delta^{12}z_t = c(1) + c(2) * \Delta^{12}z_{t-1} + c(3) * z_{t-12} + c(4) * y_{t-12} + c(5) * (y_{t-12} - z_{t-12})$
 $c(1) = 911,71 (4330,86) \ c(2) = -0,1544 (0,0959) \ c(3) = -0,848 (0,133)$
 $c(4) = -0,166 (0,09) \ c(5) = 0,17 (0,063)$
Estimated S.E. = 17064,69

./.

<i>Pagamenti in conto residui</i>	
$(z^R = f_{s_e_rs_pag_ti} \ i = 1, 2, 3)$	
Titolo I (BM)	$\Delta^{12}z^R_t = c(1) + c(2) * z^R_{t-12} + c(3) * \log(\text{trend}) + c(4) * D_{2006,12}$ $c(1) = -368,22 \ (1795,79) \ c(2) = -0,64 \ (0,05) \ c(3) = 317,6 \ (448,77)$ $c(4) = 3565,762 \ (591,61)$ <i>Estimated S.E.</i> = 577,64
Titolo II (BM)	$\Delta^{12}z^R_t = c(1) + c(2) * \Delta^{12}z^R_{t-1} + c(3) * \Delta^{12}z^R_{t-4} + c(4) * \Delta^{12}z^R_{t-12} + c(5) * D_{2006,11} + c(6) * D_{2007,08}$ $c(1) = -54,79 \ (105,5) \ c(2) = -0,24 \ (0,11) \ c(3) = 0,32 \ (0,11) \ c(4) = -0,57 \ (0,123)$ $c(5) = 1739,064 \ (558,81) \ c(6) = 2495,42 \ (518,32)$ <i>Estimated S.E.</i> = 503,2529
Titolo III (BM)	$\Delta^{12}z^R_t = c(1) + c(2) * \Delta^{12}z^R_{t-1} + c(3) * \Delta^{12}z^R_{t-2} + c(4) * \Delta^{12}z^R_{t-12} + c(5) * D_{2005,04}$ $c(1) = 264,958 \ (127,71) \ c(2) = -0,092 \ (0,067) \ c(3) = -0,001 \ (0,067)$ $c(4) = -1,038 \ (0,09) \ c(5) = 7254,374 \ (751,435)$ <i>Estimated S.E.</i> = 734,67

* I valori tra parentesi indicano gli *std errors* delle stime.

6.2 Il modello e le identità contabili

Le equazioni di comportamento stimate sono inserite all'interno della struttura del modello. Quest'ultimo contiene alcune identità che permettono di effettuare un raccordo tra le variabili in esso contenute rendendo le loro proiezioni contabilmente coerenti.

Un primo tipo d'identità riguarda la trasformazione delle variabili da variazioni tendenziali dei flussi a flussi (ovvero i valori mensili degli stanziamenti, impegni, pagamenti) e, infine, da flussi a livelli (i valori totali registrati alla fine di ogni mese). Si vanno inoltre a calcolare variabili quali la differenza tra impegni e pagamenti che a fine anno concorre a determinare i residui di nuova formazione, oltre ad altre variabili (quali la massa spendibile) che al momento non concorrono a dare dei *feedback* all'interno del modello ma ne rappresentano solo un *output*.

Tutte le variabili sono ricondotte anche a una dimensione annuale in modo da poter confluire in un modello nel quale sono proiettate anche le grandezze non disponibili a frequenza mensile nell'arco dell'esercizio finanziario (quali i residui accertati a rendiconto, le economie, l'ammontare della perenzione). Per raggiungere coerenza contabile con i dati annuali "ufficiali" di rendiconto si è provveduto a individuare regole di proiezione per le variabili di stanziamento, pagamento e impegno che sono registrate nei sistemi contabili durante i sei mesi successivi alla chiusura dell'esercizio finanziario (e che nei loro codici-serie ricevono un suffisso *_LAG*). Il valore finale di ciascuna variabile è pari al valore di dicembre del suo livello proveniente dal modello mensile sommato al valore proiettato della variabile *_LAG*. Per quanto riguarda i residui esistono raccordi con-

tabili che tengono conto della perenzione (in funzione dell'anno di provenienza), dei pagamenti sui residui e delle ulteriori variabili che ne determinano gli importi di fine anno stabiliti all'interno del rendiconto. La componente annuale del modello, contenendo prevalentemente raccordi contabili, non è dettagliatamente illustrata in questa sede.

7. - Uso del modello econometrico ai fini previsivi

In questa sezione viene valutata la capacità predittiva del modello econometrico di previsione e monitoraggio dei flussi di spesa mensili.

Si è visto nel paragrafo precedente che per prevedere il totale delle spese di bilancio si utilizzano modelli distinti sia per i tre titoli della spesa che per le due tipologie di pagamento; i pagamenti in conto competenza e quelli in conto residui. La previsione aggregata del totale dei pagamenti di cassa è poi ricostruita sommando i valori delle previsioni disaggregate per conto competenza e conto residui dei tre titoli.

Al fine di valutare la *performance* di previsione del modello econometrico, se ne comparano le previsioni con quelle di un modello *benchmark*. Quest'ultimo consiste in equazioni univariate autoregressive³⁸ stimate, anche in questo caso, per ciascuno dei tre titoli di spesa del bilancio e distinte tra pagamenti in conto competenza e conto residui. Le equazioni *benchmark* e quelle del modello spese per i pagamenti in conto competenza sono riportate nella tavola seguente.³⁹

³⁸ I *benchmark* autoregressivi sono stati selezionati sulle serie in differenze dodicesime (d12) tramite l'utilizzo di una procedura del lavoro *general to specific* disponibile del modello spese nel software PCGIVE.

³⁹ Si ricorda che in questa versione i pagamenti in conto residui coincidono con i pagamenti in conto residui del modello *benchmark*.

TAV. 3

CONFRONTO TRA EQUAZIONI DEL MODELLO BENCHMARK ED EQUAZIONI
DEL MODELLO SPESE PER I PAGAMENTI IN CONTO COMPETENZA

Equazioni dei pagamenti in conto competenza del modello *Benchmark*

$$\begin{aligned} \text{CP_Pag_t1} \quad & \Delta^{12}z_t = c(1) + c(2) * \Delta^{12}z_{t-1} + c(3) * \Delta^{12}z_{t-2} + c(4) * \Delta^{12}z_{t-12} + c(5) * D_{2007,06\%2008,1} \\ \text{CP_Pag_t2} \quad & \Delta^{12}z_t = c(1) + c(2) * \Delta^{12}z_{t-1} + c(3) * \Delta^{12}z_{t-12} + c(4) * z_{t-12} + c(5) * D_{2007,11\text{-end sample}} \\ & + c(6) * D_{2008,1-2008,2} \\ \text{CP_Pag_t3} \quad & \Delta^{12}z_t = c(1) + c(2) * \Delta^{12}z_{t-1} + c(3) * \Delta^{12}z_{t-12} \end{aligned}$$

Equazioni dei pagamenti in conto competenza del modello spese

$$\begin{aligned} \text{CP_Pag_t1} \quad & \Delta^{12}z_t = c(1) + c(2) * \Delta^{12}y_t + c(3) * \Delta^{12}z_{t-12} + c(4) * (z_{t-12}) + c(5) * y_{t-12} + c(6) * [(y_{t-1} - z_{t-1}) - \\ & (y_{t-7} - z_{t-7})] + c(7) * D_{2007,07-2007,12} + c(8) * (w_{t-1}^s - w_{t-13}^s) \\ \text{CP_Pag_t2} \quad & \Delta^{12}z_t = c(1) + c(2) * \Delta^{12}y_t + c(3) * \Delta^{12}z_{t-1} + c(4) * z_{t-12} + c(5) * y_{t-12} + c(6) * (y_{t-12} - z_{t-12}) \\ \text{CP_Pag_t3} \quad & \Delta^{12}z_t = c(1) + c(2) * \Delta^{12}z_{t-1} + c(3) * z_{t-12} + c(4) * y_{t-12} + c(5) * (y_{t-12} - z_{t-12}) \end{aligned}$$

Le strutture dei ritardi di entrambi i modelli di previsione (modello spese e modello *benchmark*) sono ottenute specificando le equazioni tra il 2004:1 e il 2007:12. A partire da tali specificazioni si effettuano previsioni fuori dal campione (*out of sample*) da uno a sei passi avanti utilizzando schemi di tipo *recursive*. Per ciascun passo di previsione si considera una finestra di previsione di 24 mesi. Le previsioni un passo avanti partiranno pertanto dal 2008:1 e termineranno il 2010:1, le previsioni 2 passi avanti inizieranno il 2008:2 e termineranno il 2010:2 e così via. Poiché la versione attuale del modello spese utilizza equazioni autoregressive per la previsione dei residui mensili dei tre titoli di spesa, nelle Tavole 9 e 10 vengono riportati solo i valori del *Root mean square forecast error* (RMSFE) e *Mean absolute error* (MAE) delle previsioni dinamiche *recursive* per $h = 1, 2, 3, 4, 5$ e 6 passi avanti delle differenze dodicesime dei pagamenti in conto competenza prodotte dai modelli relativi ai pagamenti suddivisi per titoli.

TAV. 4

RMS(F)E - RECURSIVE ESTIMATES

Modelli	RMSFE h=1	RMSFE h=2	RMSFE h=3	RMSFE h=4	RMSFE h=5	RMSFE h=6
CP_Pag_t1_BM	14.092	14.597	15.491	15.157	15.145	15.089
CP_Pag_t2_BM	2.690	2.364	2.021	1.905	2.051	2.128
CP_Pag_t3_BM	24.993	25.442	25.914	25.642	26.012	27.198
CP_Pag_t1	11.502	13.075	12.650	13.896	13.301	14.178
CP_Pag_t2	2.615	2.453	1.859	1.721	1.788	1.674
CP_Pag_t3	20.653	20.542	21.172	21.143	21.460	22.884

TAV. 5

MAE - RECURSIVE ESTIMATES

Modelli	MAE h=1	MAE h=2	MAE h=3	MAE h=4	MAE h=5	MAE h=6
CP_Pag_t1_BM	10.864	10.896	11.293	10.926	11.169	11.079
CP_Pag_t2_BM	1.836	1.724	1.590	1.486	1.565	1.679
CP_Pag_t3_BM	14.420	14.983	15.267	14.934	14.988	16.537
CP_Pag_t1	8.304	8.611	8.579	9.925	9.347	9.955
CP_Pag_t2	1.975	1.680	1.467	1.387	1.433	1.318
CP_Pag_t3	12.167	11.849	12.205	12.013	12.148	13.588

I risultati riportati in Tavola 4 indicano che il “modello spese” comprendente i dati di contabilità finanziaria produce generalmente una *performance* di previsione migliore rispetto al puro *benchmark* autoregressivo per tutti e tre i titoli tranne che per il secondo passo del titolo secondo. Anche guardando alla tavola che riporta i valori del MAE (Tav. 5) i risultati sono simili e mostrano che il modello *benchmark* fornisce risultati migliori soltanto per le previsioni al primo passo del titolo secondo.

Per disporre di una misura più accurata della valutazione della *performance*, nel seguito si riportano i risultati (Tav. 6) dell'indice di Theil definito dal rapporto tra il RMS(F)E del modello spese e quella del modello *benchmark*:

$$(1) \quad Theil = \frac{\sqrt{\sum_{t=T_1}^{T_2-h} \left(Y_{t+h}^b - \hat{Y}_{i,t+h}^b \right)^2}}{\sqrt{\sum_{t=T_1}^{T_2-h} \left(Y_{t+h}^b - \hat{Y}_{0,t+h}^b \right)^2}}$$

dove T_1 e T_2-h indicano rispettivamente la prima e l'ultima data relative alle stime *out of sample*, Y_{t+h}^b è il valore effettivo della spesa in $t+h$, $\hat{Y}_{i,t+h}^b$ rappresenta la previsione al passo h del modello spese e $\hat{Y}_{0,t+h}^b$ indica la previsione analoga ottenuta con il modello *benchmark*. Un valore di tale indicatore minore di 1 indica che il modello spese presenta una *performance* di previsione migliore rispetto al *benchmark*.

TAV. 6

RADICE DELL'ERRORE QUADRATICO MEDIO RELATIVO

Modelli	h=1	h=2	h=3	h=4	h=5	h=6
CP_Pag_t1	0,816208	0,895732	0,816603	0,916804	0,878244	0,939625
CP_Pag_t2	0,97248	1,037648	0,919842	0,903412	0,87177	0,786654
CP_Pag_t3	0,826351	0,807405	0,81701	0,824546	0,825004	0,841385

Dall'esame dei risultati della Tavola 6 si evince che il modello spese mostra una *performance* di previsione migliore del *benchmark* per tutti e tre titoli dei pagamenti ad eccezione del secondo passo di previsione del titolo secondo. Per il titolo I il modello spese è in grado di fornire risultati abbastanza buoni al primo e al terzo passo di previsione.

Al fine valutare se le differenze tra le previsioni del modello utilizzato e del *benchmark* risultano statisticamente significative nel seguito si riportano i risultati del *test* di Diebold Mariano. Tale *test* sottopone a verifica l'ipotesi H_0 che due modelli previsivi abbiano la stessa capacità previsiva contro l'ipotesi alternativa che invece essa sia diversa.⁴⁰

TAV. 7

TEST DI DIEBOLD - MARIANO (1995) PER LA CAPACITÀ PREVISIVA A LIVELLO DEL 5% (valori del *P value*)

Modelli	h=1	h=2	h=3	h=4	h=5	h=6
CP_Pag_t1	0,1458	0,4325	0,005	0,0019	0,0221	0,0163
CP_Pag_t2	0,7247	0,8236	0,2660	0,0006	0,3304	0,0181
CP_Pag_t3	0,0475	0,0206	0,0471	0,0394	0,0593	0,064

Il *test* evidenzia che per il titolo I la previsione del modello spese risulta significativamente diversa da quella del modello *benchmark* a partire dal terzo passo

⁴⁰ La statistica *test* è data da:

$$S = \frac{\bar{d}_t}{\frac{\text{cov}(d_t - d_{t-1})^{\frac{1}{2}}}{T}}$$

dove $d_t = (e_{t+h}^{MOD-S})^2 - (e_{t+h}^{BM})^2$ indica la differenza tra gli errori di previsione ottenuti usando il modello spese e quelli ottenuti stimando attraverso il modello benchmark e dove $\bar{d}_t = \frac{1}{T} \sum_{t_0}^T d_t$.

di previsione. Per il titolo II i risultati del *test* non consentono invece di rifiutare l'ipotesi nulla di assenza di differenza nell'abilità previsiva dei due modelli per la quasi totalità dei passi di previsione considerati ad eccezione del quarto e del sesto. Le previsioni del modello spese risultano invece significativamente diverse dal modello *benchmark* per quasi tutti i passi di previsione del titolo III.⁴¹

Si è detto che la previsione delle spese di bilancio può avere una valenza sia infrannuale che annuale. Nel caso di previsione annuale, quest'ultima viene aggiornata man mano che si aggiorna la previsione a livello mensile in seguito alla disponibilità di nuove informazioni. Se ad esempio si dispone di un dato mensile di spesa fino a giugno di un determinato anno, si effettuerà la previsione mensile per i restanti 6 mesi e la previsione annuale sarà ottenuta sommando le previsioni dei flussi di spesa di tali mesi a quelle del dato cumulato di giugno. Una volta disponibile il dato effettivo di luglio, si effettuerà una nuova previsione per i restanti 5 mesi dell'anno e la previsione annuale verrà aggiornata sommando al dato cumulato di luglio i valori dei flussi mensili stimati per restanti 5 mesi e così via. A partire dalle previsioni mensili è dunque possibile ottenere una previsione annuale delle spese di bilancio che diverrà più precisa man mano che nuovi dati infrannuali si renderanno disponibili.

Nella tavola seguente sono riportate le previsioni delle spese di bilancio per il 2008 e per il 2009 effettuate in corso d'anno basate su tecniche di tipo contabile e riportate nella Relazione Unificata sull'Economia e la Finanza Pubblica (RUEF), le previsioni ottenute attraverso l'uso del modello econometrico spese e il valore effettivo a rendiconto di tali spese. I valori riportati sono suddivisi inoltre in spese correnti in conto capitale e finali.

Per coerenza, poiché le stime contenute nella RUEF si basano su un set informativo di dati disponibili rispettivamente fino a febbraio per il 2008 e fino a marzo per il 2009, anche le previsioni del modello presentate nella tavola utilizzano i dati disponibili fino a quelle date. Al fine di valutare eventuali miglioramenti di accuratezza della previsione del modello spese in seguito all'aggiunta di nuovi dati, viene inoltre presentata una previsione del valore dei pagamenti di bilancio a fine anno effettuata nel mese di settembre (colonne IV e VIII).

⁴¹ È in corso di sviluppo una nuova stima delle equazioni del modello dei titoli II e III al fine di migliorarne ulteriormente la capacità previsiva.

TAV. 8

CONFRONTO TRA PREVISIONI DELLE SPESE DI BILANCIO PER IL 2008 E IL 2009
CONTENUTE NELLA RUEF (quadro di sintesi delle gestioni di cassa) E OTTENUTE
CON IL MODELLO SPESE (milioni di euro)

	Rendiconto 2008	Stime RUEF 2008 (marzo)	Stime mod. spese mar. 08	Stime mod. spese sett. 08	Rendiconto 2009	Stime RUEF 2009 (aprile)	Stime mod. spese apr. 09	Stime mod. spese sett. 09
Pagamenti correnti	463.064,5	461.469	437.044	434.115	469.577	471.912	469.858	471.517
Pagamenti c. capitale	60.956,6	652.66	431.66,30	48.464	53.670	64.142	59.904	55.882
Totale pagamenti (spese finali)	524.021,07	526.735	480.210	482.579	523.247	536.054	529.762	527.399

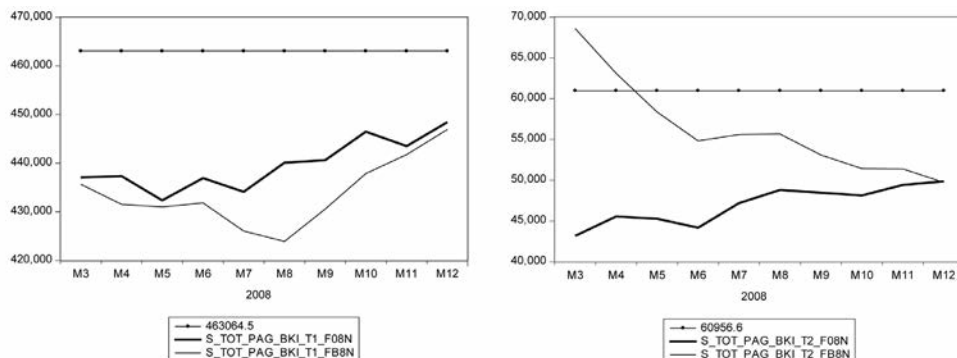
I risultati del confronto tra previsioni e valori effettivi a rendiconto mostrano che, relativamente al 2009, il modello spese ha fornito una previsione più vicina al valore dei pagamenti a rendiconto rispetto alle previsioni contenute nella RUEF sia per le spese correnti che in conto capitale. Per il 2008 al contrario, le previsioni RUEF, risultano più realistiche rispetto a quelle prodotte dal modello che sottovaluta eccessivamente sia pagamenti correnti che in conto capitale. Le stime ottenute nel mese di settembre riducono ulteriormente l'errore rispetto al dato a rendiconto sia per il 2008 sia per il 2009 relativamente al titolo II. Per il titolo I l'errore aumenta lievemente rispetto al 2008.

Il confronto con le previsioni ufficiali mostra pertanto che, sebbene il modello spese qui utilizzato possa essere oggetto di ulteriori perfezionamenti, l'ausilio di uno strumento di previsione basato su tecniche statistico-econometriche, in grado di fornire previsioni annuali di spesa in corso d'anno, appare particolarmente opportuno considerando che le previsioni "ufficiali" riferite all'aggregato bilancio dello Stato e contenute nei documenti programmatici del Ministro dell'Economia sono unicamente quelle contenute nella RUEF e formulate nel primo trimestre dell'anno. A tal riguardo il modello spese permette di colmare tale carenza consentendo di poter aggiornare mese per mese le previsioni fatte ad inizio anno sui valori a rendiconto.

Una valutazione sistematica di come l'errore di previsione varia man mano che ulteriori informazioni a livello mensile si rendono disponibili nel corso dell'anno, si evince dai Grafici 9 e 10 che riportano sia per il 2008 che per il 2009 i vari passi di previsione per spese correnti e in conto capitale ottenute con il modello spese (linea spessa) e con i modelli *benchmark* (linea sottile).

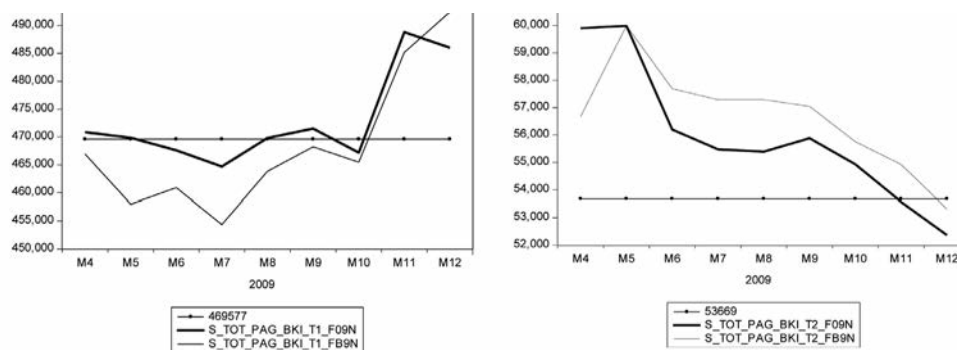
GRAF. 9

PREVISIONI DA 1 A 10 PASSI AVANTI DELLE SPESE CORRENTI E IN CONTO CAPITALE NEL 2008 OTTENUTE SIA CON IL MODELLO SPESE CHE CON IL MODELLO *BENCHMARK*



GRAF. 10

PREVISIONI DA 1 A 9 PASSI AVANTI DELLE SPESE CORRENTI E IN CONTO CAPITALE NEL 2009 OTTENUTE SIA CON IL MODELLO SPESE CHE CON IL MODELLO *BENCHMARK*



In tutti e quattro i grafici si nota che complessivamente le previsioni ai diversi passi del modello spese si avvicinano maggiormente al valore di fine anno desunto dal rendiconto rispetto alle previsioni dei modelli *benchmark*.

Osservando la dinamica delle previsioni rispetto al valore di rendiconto, un altro elemento che emerge riguarda il gap che permane anche a fine anno (sebbene nel corso d'anno si riduca almeno per il 2008) tra i valori previsti a novembre dal modello ed il valore di rendiconto.

8. - Ulteriori considerazioni e direzioni di sviluppo del lavoro

Il modello presentato in questo lavoro è da considerarsi una versione pilota, rispetto alla quale si prospettano numerose ulteriori analisi da effettuare in funzione di possibili evoluzioni.

Un primo ambito di approfondimento riguarda la decisione dell'opportuno livello di aggregazione da utilizzare a regime, considerando l'eventuale passaggio da titoli alle categorie. Una tale scelta non necessariamente comporterebbe la necessità di modellare tutte le categorie disponibili separatamente ma si potrebbero considerare "raggruppamenti" di queste tenendo conto di un insieme di fattori (o esigenze) eventualmente da bilanciare.

Innanzitutto si dovrebbe valutare l'eventuale miglioramento della capacità previsiva proveniente da una disaggregazione più fine delle componenti del bilancio dello Stato. La letteratura in materia riferisce in sostanza che⁴² si ottiene un vantaggio nel prevedere singolarmente le componenti di un aggregato se le serie storiche disaggregate presentano dei comportamenti statistici diversi tra di loro e se gli errori di previsioni nel modellarle separatamente sono tra loro correlati in maniera negativa (tendono a compensarsi una volta che si riaggrega la previsione). In ogni caso, va aggiunto che soprattutto nell'ambito di una istituzione di *policy*, il guadagno di capacità previsiva andrebbe soppesato rispetto al maggiore costo computazionale proveniente dall'utilizzo di un numero più ampio di serie storiche ed alla tempestività con la quale si vuole procedere. In un tale contesto vanno valutati anche fattori di natura istituzionale quali la disponibilità di informazione (che prescinde e può risultare per alcune voci di spesa migliore) delle previsioni econometriche e la necessità di inserire ed integrare lo strumento statistico all'interno di un processo in cui diversi attori contribuiscono ad elaborare le previsioni per singoli aggregati di spesa utilizzando diverse metodologie (non si conviene dunque che un modello produca delle voci aggregate che non consentano di dialogare in maniera opportuna con tali soggetti⁴³).

Nel caso specifico del modello oggetto di questo lavoro, la scelta più facile riguarda la categoria degli interessi a carico del bilancio dello Stato. Questa voce di spesa, al contrario delle altre, è già materia di attentissima modellizzazione a livello

⁴² NEWBOLD P., HARVEY D.I. (2002).

⁴³ All'interno della Ragioneria Generale dello Stato operano diversi Ispettorati responsabili rispettivamente, per ricordare solo alcuni aspetti, della spesa sociale, di quella per il personale e dei rapporti con gli enti locali (e di conseguenza dei trasferimenti verso questi a carico del bilancio dello Stato).

mensile.⁴⁴ Risulta dunque opportuno, anche con il fine ultimo di contribuire alla formulazione di previsioni funzionali alla politica di emissione di titoli del debito, giungere a delle previsioni di spesa primaria. Per quanto riguarda ulteriori disaggregazioni del titolo primo non si è ancora proceduto sistematicamente. Una prova in corso di attuazione riguarda l'ulteriore suddivisione delle spese correnti andando a modellare separatamente dal resto la spesa per acquisti di beni intermedi e quella in trasferimenti. La spesa primaria, dunque si suddividerebbe nelle due categorie appena nominate e in una componente residua.

Un approccio più sistematico richiede tuttavia di tenere conto di altri aspetti. Esiste in primo luogo l'indicazione di natura istituzionale, rilevata nella sezione di rassegna della letteratura, di tenere separate le voci di spesa per le quali i pagamenti contabilizzati nel bilancio dello Stato impattano direttamente sul fabbisogno da quelle che passando dal Conto di Tesoreria. Si ricorda, infatti, che le seconde sono meno correlate (almeno nella tempistica) rispetto ai pagamenti e che la loro previsione è in maggior misura "migliorabile" tenendo conto di informazioni legate ai contatti istituzionali.

Merita inoltre un'adeguata analisi, l'eventualità di procedere a una disaggregazione della spesa (anche) in funzione della diversa natura autorizzatoria delle sue componenti⁴⁵ che determina una loro diversa reattività di fronte a interventi, anche generalizzati di riduzione della spesa.

⁴⁴ Modelli di previsione della spesa degli interessi in funzione dello *stock* di titoli del debito e della curva dei tassi.

⁴⁵ A tal riguardo, i sistemi contabili della Ragioneria hanno fino ad ora distinto tra "fabbisogno", spesa per "fattori legislativi" (spesa di funzionamento) e "oneri inderogabili" spesa obbligatoria. Le spese di adeguamento al fabbisogno, ossia spese non predeterminate legislativamente e quantificate tenendo conto delle esigenze delle amministrazioni, risultano più facilmente "aggredibili" essendo possibile controllarne la dinamica in via amministrativa. Tagli agli stanziamenti per le relative voci di spesa produrranno un impatto immediato nel breve sugli impegni e sulla spesa. Per quanto riguarda i fattori legislativi, questi rappresentano le spese autorizzate da un provvedimento legislativo che ne stabiliva l'importo, considerato quale limite massimo di spesa, e il periodo di iscrizione in bilancio. La riduzione degli importi stanziati deve passare da una risoluzione parlamentare. Infine, la componente di spesa obbligatoria riguarda in via prevalente diritti soggettivi individuali (stipendi, assegni, pensioni), la spesa per interessi e altre categorie considerate tali per espressa disposizione normativa. Riguardo alle spese classificate in quest'ultima categoria, la variazione del loro importo non può che procedere anche in questo caso tramite provvedimenti legislativi (caratterizzati però da una maggiore difficoltà politica di attuazione rispetto a quelli legati ai fattori legislativi). Le voci di spesa obbligatoria avranno tuttavia una loro dinamica che in assenza d'interventi concreti da parte della PA potrebbe non essere impattata (da qui la possibilità che la spesa ecceda gli stanziamenti e che questi debbano essere rivisti verso l'alto).
./.

Un'ulteriore direzione di approfondimento riguarda la possibilità di inserire grandezze contabili "a carattere intermedio". Esempi di tali variabili sono la massa spendibile e la disponibilità. Quest'ultima variabile è teoricamente quella che dovrebbe condizionare in maniera più prossima gli impegni essendo pari alla differenza tra lo stanziato e quanto attualmente già impegnato diminuita dagli accantonamenti di bilancio. Gli accantonamenti sostanzialmente equivalgono a degli importi che, sia pure stanziati, non sono più impegnabili. Al momento sono in corso soltanto alcune prove per valutare l'apporto in termini previsivi di queste variabili. Tuttavia la loro introduzione porterebbe a una espansione "verticale" del modello con un "appesantimento" legato alla necessità di inserire ulteriori identità di raccordo e di ipotesi comportamentali per le nuove variabili (ad esempio, per introdurre la disponibilità occorrerebbe poter prevedere gli accantonamenti).

Le capacità previsive del modello spese e la loro migliore *performance* rispetto a modelli puramente statistici, sono in parte condizionati dai valori delle variabili di stanziamento annuali oggetto di previsione in sede di legge di bilancio (modificati nel corso dell'anno da variazioni degli stanziamenti). Questi valori sono dunque esogeni rispetto al modello. Le previsioni di impegno e di pagamento che scaturiscono dal modello, pur se non completamente vincolate ad essi, risentono dei loro valori in funzione di parametri stimati dalle equazioni. Una dinamica particolarmente elevata (rispetto agli stanziamenti) dei valori registratisi nei primi mesi dell'anno tenderebbe a essere ridimensionata nella simulazione previsiva per i mesi successivi. Esiste, in particolare per queste componenti di spesa, un problema di valutazione "esogena" da parte del previsore sull'ammontare degli stanziamenti. Considerazioni simili valgono anche per simulazioni non di carattere previsivo ma volte a valutare l'impatto di una riduzione degli stanziamenti. È da ritenersi che una variazione degli stanziamenti relativa a voci discrezionali possa impattare più rapidamente e con maggiore "certezza" sulla spesa.

APPENDICE 1

Cross correlazioni di diverse variabili di bilancio espresse in differenze dodicesime con le differenze dodicesime dei pagamenti di bilancio totali e suddivisi per titoli

TAV. 9

CROSS CORRELAZIONI DEI DATI DI BILANCIO CON LE DIFFERENZE DODICESIME DEI PAGAMENTI IN CONTO COMPETENZA DEL TITOLO I (spese correnti) (F_S_E_CP_PAG_T1D12)

Lag	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
f_s_e_CP_IMP_T1D12	-0,4	0,04	0,27	-0,1	0,06	0	-0,06	-0,14	0,2	0	-0,24	-0,04	0,65
f_s_e_cp_dis_t1D12	0,1	-0,09	-0,03	0,02	-0,12	-0,16	0,24	0,15	-0,19	0,09	0,01	0,10	-0,31
f_s_e_cp_dis_f_t1D12	-0,11	-0,18	0,15	0,09	-0,09	-0,38	0,24	0,17	0,03	0	-0,02	-0,09	0
f_s_e_cp_dis_nf_t1D12	0,13	-0,06	-0,06	0	-0,11	-0,09	0,21	0,13	-0,21	0,1	0,02	0,13	-0,34

TAV. 10

CROSS CORRELAZIONI DEI DATI DI BILANCIO CON LE DIFFERENZE DODICESIME DEI PAGAMENTI IN CONTO COMPETENZA DEL TITOLO II (spese in conto capitale) (F_S_E_CP_PAG_T2D12)

Lag	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
f_s_e_cp_imp_t2D12	-0,27	0,13	0,15	0,01	-0,04	0,01	-0,1	0,1	0	0,04	0,11	-0,3	0,8
f_s_e_cp_dis_t2 D12	-0,02	0,12	-0,07	0,07	0,2	0,08	0	0,03	0	0,06	-0,04	0,08	-0,15
f_s_e_cp_dis_f_t2 D12	-0,41	0,18	-0,08	-0,16	0,07	0,03	-0,2	0	-0,02	0,08	-0,14	-0,21	0,43
f_s_e_cp_dis_nf_t2 D12	0,03	0,12	-0,16	0,14	0,11	0	0,04	0,06	0,04	0,15	-0,21	0,52	-0,44

TAV. 11

CROSS CORRELAZIONI DEI DATI DI BILANCIO CON LE DIFFERENZE DODICESIME DEI PAGAMENTI IN CONTO COMPETENZA DEL TITOLO III (rimborso prestiti) (F_S_E_CP_PAG_T3D12)

Lag	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
f_s_e_cp_imp_t3	-0,08	0,02	-0,08	0,15	0,06	-0,22	0,2	-0,27	0,33	-0,17	0,20	-0,07	0,06
f_s_e_cp_imp_t3D12	-0,15	0,08	-0,11	0,17	0,01	-0,17	0,24	-0,27	0,29	-0,21	0,15	0,06	0,1
f_s_e_cp_dis_t3D12	0,15	-0,1	0,13	-0,18	0,02	0,17	-0,25	0,26	-0,28	0,21	-0,14	-0,05	-0,11
f_s_e_cp_dis_f_t3D12	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na
f_s_e_cp_dis_nf_t3D12	0,15	-0,1	0,13	-0,18	0,02	0,17	-0,25	0,26	-0,28	0,21	-0,14	-0,05	-0,11

TAV. 12

CROSS CORRELAZIONI DEI DATI DI BILANCIO CON LE DIFFERENZE
DODICESIME DEGLI IMPEGNI IN CONTO COMPETENZA DEL TITOLO I
(spese correnti) (F_S_E_CP_IMP_T1D12)

Lag	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
f_s_e_cp_dis_t1D12	0,09	-0,07	0,07	0,08	-0,17	-0,17	0,25	-0,08	0,04	0,08	-0,01	0,22	-0,53
f_s_e_cp_dis_f_t1D12	-0,2	-0,01	0,15	0,2	-0,2	-0,26	0,15	0,04	0,13	-0,12	0,16	-0,08	-0,01
f_s_e_cp_dis_nf_t1D12	0,14	-0,07	0,04	0,04	-0,14	-0,12	0,24	-0,09	0,02	0,11	-0,04	0,26	-0,58

APPENDICE 2

Equazioni di raccordo ed identità contabili modello mensile

VARIABILI CONTEMPORANEE

STANZIAMENTI

I dati di stanziamento di bilancio (variabili S_P_CP_LB) derivano dalla parte annuale e sono esogeni

variazione stanziamenti in corso d'anno

:EQ_STAN_T1_BM

:EQ_STAN_T2_BM

:EQ_STAN_T3_BM

passaggio da differenze dodicesime a flussi

$$(2) \quad f_{s_e_cp_t1} = f_{s_e_cp_t1d12} + f_{s_e_cp_t1(-12)}$$

$$(3) \quad f_{s_e_cp_t2} = f_{s_e_cp_t2d12} + f_{s_e_cp_t2(-12)}$$

$$(4) \quad f_{s_e_cp_t3} = f_{s_e_cp_t3d12} + f_{s_e_cp_t3(-12)}$$

$$(5) \quad f_{s_e_cp_pr} = f_{s_e_cp_t1} + f_{s_e_cp_t2}$$

$$(6) \quad f_{s_e_cp} = f_{s_e_cp_t1} + f_{s_e_cp_t2} + f_{s_e_cp_t3}$$

equazioni di creazione di variabili di *stock* - totale stanziato (nuovi stanziamenti)

$$(7) \quad s_{e_cp_t1} = dummy1b * (f_{s_e_cp_t1}) + dummy11b * (s_{e_cp_t1(-1)} + f_{s_e_cp_t1})$$

$$(8) \quad s_{e_cp_t2} = dummy1b * (f_{s_e_cp_t2}) + dummy11b * (s_{e_cp_t2(-1)} + f_{s_e_cp_t2})$$

$$(9) \quad s_{e_cp_t3} = dummy1b * (f_{s_e_cp_t3}) + dummy11b * (s_{e_cp_t3(-1)} + f_{s_e_cp_t3})$$

$$(10) \quad s_{e_cp} = s_{e_cp_t1} + s_{e_cp_t2} + s_{e_cp_t3}$$

totale stanziamenti di competenza a fine di ciascun mese

$$(11) S_TOT_CP = S_P_CP_LB + S_E_CP$$

$$(12) F_S_TOT_CP = D(S_TOT_CP)$$

$$(13) S_TOT_CP_T1 = S_P_CP_LB_T1 + S_E_CP_T1$$

$$(14) F_S_TOT_CP_T1 = D(S_TOT_CP_T1)$$

$$(15) S_TOT_CP_T2 = S_P_CP_LB_T2 + S_E_CP_T2$$

$$(16) F_S_TOT_CP_T2 = D(S_TOT_CP_T2)$$

$$(17) S_TOT_CP_T3 = S_P_CP_LB_T3 + S_E_CP_T3$$

$$(18) F_S_TOT_CP_T3 = D(S_TOT_CP_T3)$$

$$(19) S_TOT_CP_PR = S_TOT_CP_T1 + S_TOT_CP_T2$$

$$(20) F_S_TOT_CP_PR = D(S_TOT_CP_PR)$$

IMPEGNI

impegni in conto competenza

:eq_imp_t1_ver3

:eq_imp_t2_ver1

:eq_imp_t3_ver1

impegni in conto residui (attualmente posti eguali a zero)

$$(21) s_e_rs_imp_t1 = 0$$

$$(22) s_e_rs_imp_t2 = 0$$

$$(23) s_e_rs_imp_t3 = 0$$

$$(24) s_e_rs_imp = 0$$

passaggio da differenze dodicesime a flussi

$$(25) f_s_e_cp_imp_t1 = f_s_e_cp_imp_t1d12 + f_s_e_cp_imp_t1(-12)$$

$$(26) f_s_e_cp_imp_t2 = f_s_e_cp_imp_t2d12 + f_s_e_cp_imp_t2(-12)$$

$$(27) f_s_e_cp_imp_t3 = f_s_e_cp_imp_t3d12 + f_s_e_cp_imp_t3(-12)$$

$$(28) f_s_e_cp_imp_pr = f_s_e_cp_imp_t1 + f_s_e_cp_imp_t2$$

$$(29) f_s_e_cp_imp = f_s_e_cp_imp_t1 + f_s_e_cp_imp_t2 + f_s_e_cp_imp_t3$$

equazioni di creazione di variabili di *stock* - totale impegnato in conto competenza

$$(30) \quad s_e_cp_imp_t1 = dummy1b * (f_s_e_cp_imp_t1) + dummy11b * (s_e_cp_imp_t1(-1) + f_s_e_cp_imp_t1)$$

$$(31) \quad s_e_cp_imp_t2 = dummy1b * (f_s_e_cp_imp_t2) + dummy11b * (s_e_cp_imp_t2(-1) + f_s_e_cp_imp_t2)$$

$$(32) \quad s_e_cp_imp_t3 = dummy1b * (f_s_e_cp_imp_t3) + dummy11b * (s_e_cp_imp_t3(-1) + f_s_e_cp_imp_t3)$$

$$(33) \quad s_e_cp_imp = s_e_cp_imp_t1 + s_e_cp_imp_t2 + s_e_cp_imp_t3$$

totale impegni a fine di ciascun mese

$$(34) \quad S_TOT_IMP_T1 = S_E_CP_IMP_T1 + S_E_RS_IMP_T1$$

$$(35) \quad S_TOT_IMP_T2 = S_E_CP_IMP_T2 + S_E_RS_IMP_T2$$

$$(36) \quad S_TOT_IMP_T3 = S_E_CP_IMP_T3 + S_E_RS_IMP_T3$$

$$(37) \quad S_TOT_IMP_PR = S_TOT_IMP_T1 + S_TOT_IMP_T2$$

$$(38) \quad S_TOT_IMP = S_E_CP_IMP + S_E_RS_IMP$$

PAGAMENTI

previsione dei flussi di pagamento Banca d'Italia in differenze dodicesime

conto competenza

:EQ_CP_PAG_T1_ver3

:EQ_CP_PAG_T2_ver1

:EQ_CP_PAG_T3_ver1

equazioni di accumulo *stock* pagamenti sui residui

$$(39) \quad s_st_rs_acc_al_t1 = s_e_rs_acc_al_t1_Z(-12) - s_e_rs_pag_t1(-1) * dummy11b$$

$$(40) \quad s_st_rs_acc_al_t2 = s_e_rs_acc_al_t2_Z(-12) - s_e_rs_pag_t2(-1) * dummy11b$$

$$(41) \quad s_st_rs_acc_al_t3 = s_e_rs_acc_al_t3_Z(-12) - s_e_rs_pag_t3(-1) * dummy11b$$

conto residui

:EQ_RS_PAG_T1_ver6

:EQ_RS_PAG_T2_ver6

:EQ_RS_PAG_T3_BM

$$(42) \quad f_{s_e_rs_pag_pr} = f_{s_e_rs_pag_t1} + f_{s_e_rs_pag_t2}$$

$$(43) \quad f_{s_e_rs_pag} = f_{s_e_rs_pag_t1} + f_{s_e_rs_pag_t2} + f_{s_e_rs_pag_t3}$$

passaggio da differenze dodicesime a flussi

competenza

$$(44) \quad f_{s_e_cp_pag_t1} = f_{s_e_cp_pag_t1d12} + f_{s_e_cp_pag_t1(-12)}$$

$$(45) \quad f_{s_e_cp_pag_t2} = f_{s_e_cp_pag_t2d12} + f_{s_e_cp_pag_t2(-12)}$$

$$(46) \quad f_{s_e_cp_pag_t3} = f_{s_e_cp_pag_t3d12} + f_{s_e_cp_pag_t3(-12)}$$

$$(47) \quad f_{s_e_cp_pag_pr} = f_{s_e_cp_pag_t1} + f_{s_e_cp_pag_t2}$$

$$(48) \quad f_{s_e_cp_pag} = f_{s_e_cp_pag_t1} + f_{s_e_cp_pag_t2} + f_{s_e_cp_pag_t3}$$

residui

$$(49) \quad f_{s_e_rs_pag_t1d12} = f_{s_e_rs_pag_t1} - f_{s_e_rs_pag_t1(-12)}$$

$$(50) \quad f_{s_e_rs_pag_t2d12} = f_{s_e_rs_pag_t2} - f_{s_e_rs_pag_t2(-12)}$$

$$(51) \quad f_{s_e_rs_pag_t3d12} = f_{s_e_rs_pag_t3} - f_{s_e_rs_pag_t3(-12)}$$

$$(52) \quad f_{s_e_rs_pag_prd12} = f_{s_e_rs_pag_t1d12} + f_{s_e_rs_pag_t2d12}$$

$$(53) \quad f_{s_e_rs_pagd12} = f_{s_e_rs_pag_t1d12} + f_{s_e_rs_pag_t2d12} + f_{s_e_rs_pag_t3d12}$$

$$(54) \quad f_{s_e_cs_pag_t1d12} = f_{s_e_cp_pag_t1d12} + f_{s_e_rs_pag_t1d12}$$

$$(55) \quad f_{s_e_cs_pag_t2d12} = f_{s_e_cp_pag_t2d12} + f_{s_e_rs_pag_t2d12}$$

$$(56) \quad f_{s_e_cs_pag_t3d12} = f_{s_e_cp_pag_t3d12} + f_{s_e_rs_pag_t3d12}$$

equazioni di raccordo tra flussi di pagamenti Banca d'Italia e pagamenti di bilancio

:eq_cp_raccordo_bki_t1

:eq_cp_raccordo_bki_t2

:eq_cp_raccordo_bki_t3

$f_{cp_pag_bki} = f_{cp_pag_bki_t1} + f_{cp_pag_bki_t2} + f_{cp_pag_bki_t3}$

$f_{cp_pag_bki_pr} = f_{cp_pag_bki_t1} + f_{cp_pag_bki_t2}$

:eq_rs_raccordo_bki_t1

:eq_rs_raccordo_bki_t2

:eq_rs_raccordo_bki_t3

(57) $f_{rs_pag_bki} = f_{rs_pag_bki_t1} + f_{rs_pag_bki_t2} + f_{rs_pag_bki_t3}$

(58) $f_{rs_pag_bki_pr} = f_{rs_pag_bki_t1} + f_{rs_pag_bki_t2}$

equazioni di creazione di variabili di *stock* - totale pagato -

(59) $s_{e_cp_pag_t1} = dummy1b * (f_{s_e_cp_pag_t1}) + dummy11b * (s_{e_cp_pag_t1}(-1) + f_{s_e_cp_pag_t1})$

(60) $s_{e_cp_pag_t2} = dummy1b * (f_{s_e_cp_pag_t2}) + dummy11b * (s_{e_cp_pag_t2}(-1) + f_{s_e_cp_pag_t2})$

(61) $s_{e_cp_pag_t3} = dummy1b * (f_{s_e_cp_pag_t3}) + dummy11b * (s_{e_cp_pag_t3}(-1) + f_{s_e_cp_pag_t3})$

(62) $s_{e_cp_pag_pr} = dummy1b * (f_{s_e_cp_pag_pr}) + dummy11b * (s_{e_cp_pag_pr}(-1) + f_{s_e_cp_pag_pr})$

(63) $s_{e_cp_pag} = s_{e_cp_pag_t1} + s_{e_cp_pag_t2} + s_{e_cp_pag_t3}$

(64) $s_{e_rs_pag_t1} = dummy1b * (f_{s_e_rs_pag_t1}) + dummy11b * (s_{e_rs_pag_t1}(-1) + f_{s_e_rs_pag_t1})$

(65) $s_{e_rs_pag_t2} = dummy1b * (f_{s_e_rs_pag_t2}) + dummy11b * (s_{e_rs_pag_t2}(-1) + f_{s_e_rs_pag_t2})$

(66) $s_{e_rs_pag_t3} = dummy1b * (f_{s_e_rs_pag_t3}) + dummy11b * (s_{e_rs_pag_t3}(-1) + f_{s_e_rs_pag_t3})$

(67) $s_{e_rs_pag_pr} = dummy1b * (f_{s_e_rs_pag_pr}) + dummy11b * (s_{e_rs_pag_pr}(-1) + f_{s_e_rs_pag_pr})$

(68) $s_{e_rs_pag} = s_{e_rs_pag_t1} + s_{e_rs_pag_t2} + s_{e_rs_pag_t3}$

$$(69) \quad s_e_cs_pag_t1 = s_e_cp_pag_t1 + s_e_rs_pag_t1$$

$$(70) \quad s_e_cs_pag_t2 = s_e_cp_pag_t2 + s_e_rs_pag_t2$$

$$(71) \quad s_e_cs_pag_t3 = s_e_cp_pag_t3 + s_e_rs_pag_t3$$

$$(72) \quad s_e_cs_pag_pr = s_e_cp_pag_pr + s_e_rs_pag_pr$$

$$(73) \quad s_e_cs_pag = s_e_cp_pag + s_e_rs_pag$$

equazioni di creazione di variabili di *stock* - totale pagato - Banca d'Italia

$$(74) \quad s_cp_pag_bki_t1 = dummy1b * (f_cp_pag_bki_t1) + dummy11b * (s_cp_pag_bki_t1(-1) + f_cp_pag_bki_t1)$$

$$(75) \quad s_cp_pag_bki_t2 = dummy1b * (f_cp_pag_bki_t2) + dummy11b * (s_cp_pag_bki_t2(-1) + f_cp_pag_bki_t2)$$

$$(76) \quad s_cp_pag_bki_t3 = dummy1b * (f_cp_pag_bki_t3) + dummy11b * (s_cp_pag_bki_t3(-1) + f_cp_pag_bki_t3)$$

$$(77) \quad s_cp_pag_bki_pr = dummy1b * (f_cp_pag_bki_pr) + dummy11b * (s_cp_pag_bki_pr(-1) + f_cp_pag_bki_pr)$$

$$(78) \quad s_cp_pag_bki = s_cp_pag_bki_t1 + s_cp_pag_bki_t2 + s_cp_pag_bki_t3$$

$$(79) \quad s_rs_pag_bki_t1 = dummy1b * (f_rs_pag_bki_t1) + dummy11b * (s_rs_pag_bki_t1(-1) + f_rs_pag_bki_t1)$$

$$(80) \quad s_rs_pag_bki_t2 = dummy1b * (f_rs_pag_bki_t2) + dummy11b * (s_rs_pag_bki_t2(-1) + f_rs_pag_bki_t2)$$

$$(81) \quad s_rs_pag_bki_t3 = dummy1b * (f_rs_pag_bki_t3) + dummy11b * (s_rs_pag_bki_t3(-1) + f_rs_pag_bki_t3)$$

$$(82) \quad s_rs_pag_bki_pr = dummy1b * (f_rs_pag_bki_pr) + dummy11b * (s_rs_pag_bki_pr(-1) + f_rs_pag_bki_pr)$$

$$(83) \quad s_rs_pag_bki = s_rs_pag_bki_t1 + s_rs_pag_bki_t2 + s_rs_pag_bki_t3$$

$$(84) \quad f_tot_pag_bki_t1 = f_cp_pag_bki_t1 + f_rs_pag_bki_t1$$

$$(85) \quad f_tot_pag_bki_t2 = f_cp_pag_bki_t2 + f_rs_pag_bki_t2$$

$$(86) \quad f_tot_pag_bki_t3 = f_cp_pag_bki_t3 + f_rs_pag_bki_t3$$

$$(87) \quad f_tot_pag_bki_pr = f_cp_pag_bki_pr + f_rs_pag_bki_pr$$

$$(88) \quad f_tot_pag_bki = f_cp_pag_bki + f_rs_pag_bki$$

$$(89) \quad s_tot_pag_bki_t1 = s_cp_pag_bki_t1 + s_rs_pag_bki_t1$$

$$(90) \quad s_tot_pag_bki_t2 = s_cp_pag_bki_t2 + s_rs_pag_bki_t2$$

$$(91) \quad s_tot_pag_bki_t3 = s_cp_pag_bki_t3 + s_rs_pag_bki_t3$$

$$(92) \quad s_tot_pag_bki_pr = s_cp_pag_bki_pr + s_rs_pag_bki_pr$$

$$(93) \quad s_tot_pag_bki = s_cp_pag_bki + s_rs_pag_bki$$

creazione della variabile di residui di nuova formazione

$$(94) \quad S_E_RS_COM_T1 = S_E_CP_IMP_T1 - S_E_CP_PAG_T1$$

$$(95) \quad S_E_RS_COM_T2 = S_E_CP_IMP_T2 - S_E_CP_PAG_T2$$

$$(96) \quad S_E_RS_COM_T3 = S_E_CP_IMP_T3 - S_E_CP_PAG_T3$$

$$(97) \quad S_e_RS_COM_PR = S_E_RS_COM_T1 + S_E_RS_COM_T2$$

$$(98) \quad S_E_RS_COM = S_E_CP_IMP - S_E_CP_PAG$$

$$(99) \quad S_E_RS_STAN = S_TOT_CP - S_E_CP_IMP$$

$$(100) \quad S_E_RS_STAN_T1 = S_TOT_CP_T1 - S_E_CP_IMP_T1$$

$$(101) \quad S_E_RS_STAN_T2 = S_TOT_CP_T2 - S_E_CP_IMP_T2$$

$$(102) \quad S_E_RS_STAN_T3 = S_TOT_CP_T3 - S_E_CP_IMP_T3$$

$$(103) \quad S_E_RS_STAN_PR = S_E_RS_STAN_T1 + S_E_RS_STAN_T2$$

BIBLIOGRAFIA

- ARTONI R. - BERNARDI L. - LIPPI M., «Analisi delle serie storiche e bilancio dello Stato», *Rivista di Diritto Finanziario e Scienza delle Finanze*, 1985.
- ASTEGIANO L. - COLETTA S. - FERRARI S. - FREZZA F. - LOMUSCIO S. - PERUGINI L. - STATI A., «Il data warehouse ed i sistemi informativi direzionali della Ragioneria Generale dello Stato», *Quaderni Consip*, n. IX, 2006.
- DIEBOLD F.X. - MARIANO R.S., «Comparing Predictive Accuracy», *Journal of Business & Economic Statistics*, American Statistical Association, vol. 13(3), 1995.
- FORECASTING GOVERNMENT CASH FLOWS IN THE UK, pagina internet: <http://www.mj-w.net/>
- GIARDA P., «Il monitoraggio mensile della spesa del settore statale», in *La riforma dell'amministrazione pubblica tra vincoli di bilancio ed obiettivi di efficienza*, Commissione Tecnica per la Spesa Pubblica, Ministero del Tesoro, del Bilancio e della Programmazione Economica, 1997.
- GIOVANNINI A., «L'utilizzo dell'analisi delle serie storiche per la previsione del fabbisogno di cassa del settore statale», in BERNARDI L. (a cura di), *Informazioni e previsioni per la finanza pubblica*, Il Mulino, 1991.
- LEAL T. - PEDREGAL D.J. - PÉREZ J.J., «Short-Term Monitoring of the Spanish Government Balance with Mixed-Frequencies Models», Banco de España, *Working Papers*, no. 0931, 2009.
- MOULIN L.M. - SALTO A. - SILVESTRINI VEREDAS D., «Using Intra Annual Information to Forecast the Annual State Deficits. The Case of France», *Core Discussion Paper*, no. 48, IAP Statistics Network, 2004.
- NEWBOLD P. - HARVEY D.I., «Forecast Combination and Encompassing», in CLEMENTS M.P. - HENDRY D.F. (eds.), *A Companion to Economic Forecasting*, Blackwell Publishers, 2002.
- ONORANTE L. - PEDREGAL D.J. - PÉREZ J.J. - SIGNORINI S., «The Usefulness of Intra-Annual Government Cash Budgetary Data for Fiscal Forecasting in the Euro Area», *Journal of Policy Modeling*, Elsevier, vol. 32(1), pages 98-119, January 2010.
- PEAR J., *Casuality*, Cambridge University Press, 2000.
- PEDREGAL D.J. - PEREZ J.J., «Should Quarterly Government Finance Statistics Be Used for Fiscal Surveillance in Europe?», *ECB Working Papers*, no. 937, 2008.
- PISAURO G., «Le tecniche budgetarie ed il sistema dei controlli», in *Nuovo sistema dei controlli sulla spesa pubblica*, Banca d'Italia, 1994.
- POZZUOLI S. - ERCOLI R., «Quanto sono credibili le previsioni di finanza pubblica di medio e lungo termine in Italia? Un modello previsionale del fabbisogno del settore statale», *SIEP Working Paper*, 2008.
- RAGIONERIA GENERALE DELLO STATO, «I principali saldi di finanza pubblica: definizioni, utilizzo, raccordi », Servizio studi, *Strumenti e Metodi*, n. 1, 2008.

- RAGIONERIA GENERALE DELLO STATO, «La riforma della contabilità e della finanza pubblica. Novità, riflessioni e prospettive», Servizio studi, *Strumenti e metodi*, n. 2, 2010.
- SPRITES P. - GLYMOURE C. - SCHEINES R., *Causation, Prediction and Search*, Springer - Verlag, New York DC, 1993.
- ZANARDI A., «La tesoreria statale: informazioni di consuntivo, previsioni gestionali e proposte di intervento strutturali», in BERNARDI L. (a cura), *Analisi e modelli per la gestione della finanza pubblica*, Bologna, Il Mulino, 1991.

On the Fuzzy Boundaries between Public and Private in Health-Care Organization and Funding Systems

Alessandro Petretto*
University of Florence

The paper proposes a theoretical survey of health-care organization and funding systems, by underlining and discussing the combination public versus private provision and production of social assistance services. The survey then examines the present features of various National health services, focusing on the efficiency and equity of public intervention and on the industrial organization of the institutional design of health care. Hence, the paper treats the cost-benefit of vertical integration of structures, devoted respectively to purchasing and providing health-care services, it considers the effects of managed competition in quasi-markets and, finally, it deals with supplementary insurance systems.

[JEL Classification: I1; H1; H4].

Keywords: health-care organization; health-care funding; social insurance; private and public provision.

* <alessandro.petretto@unifi.it>, Department of Social Sciences. Paper presented at the Special session “Markets and institutions”, of 7th Congress of the Italian Association for the History of Political Economy (STOREP), *The Shifting Boundaries between Public and Private in Economics*, Trent May 30th - June 1st 2010. I thank all the participants to the session for the helpful discussion. I also thank the anonymous referees for their valuable comments and suggestions.

1. - Introduction

The paper presents a mainly theoretical survey of health-care organization and financing models with a specific glance to the boundaries between public and private involvement in this sector. Indeed, health-care supplies an exemplary “case study”, enlightening all the possible ways of designing and structuring the public intervention in welfare economics. In industrialised countries, the public intervention in health care is realised by several regimes of organization and financing, and none has an absolute supremacy over the others in terms of productive and allocative efficiency, equity and fulfilment of the implicit social rights. Therefore, for the economists, health-care organization is a very ticklish and scientifically stimulating topic.

In analysing and evaluating the various organizational and funding systems we cannot leave out some preliminary considerations about health-care policy in the main industrialised countries. First, from the supply side, the sector is characterised by a long run costs inflation process, due to the shifting up of medical technological frontier and to the consequent dynamics of quality of care (Cutler, 2002). Moreover, from the demand side, there is an expenditure shift due to population ageing and other demographic processes (Zweifel *et al.*, 2009, ch. 1). Second, in health-care sector there is a huge number of economic agents, with different and often conflicting preferences and interests, and consequently payoff functions. Therefore, it is illusionary to treat this complex non-cooperative game like a unitary and centralised decision-making, as if it would be a cooperative game. Third, the outcome of any chosen organization is highly influenced by the flow of incentives among the agents. In this respect, for instance, fully planned *command & control* systems, which are not as well *incentive-compatible*, can result deeply inefficient and even inequitable.

Indeed, according to the modern economic theory, health-care systems are successful in attaining their aims if they can optimally combine the incentives of physicians, patients, providers, and funding entities, for producing and supplying health services at a quantitative and qualitative level coherent with the effective need, containing production costs and optimally allocating public and private resources among the agents themselves. Within each health-care organization there are several principal/agent relationships with asymmetric information whose framework influences dramatically the decision making outcome. Consequently, it is crucial to have a system of incentives and delegations able to boost all the informed agents to choose strategies coherently with the objectives of the uninformed organization (Zweifel *et al.*, 2009, ch. 11).

The distribution of incentives among economic agents, with different and independent pay-offs, roughly reminds a notion of competitive market, characterised by complex structures of demand and supply. In this regard, by many scholars of disciplines not in the field of economics the connection of market and competition notions to health care is considered as improper because not compatible with the social nature of the implicit rights. One of the aim of this work is to explain how the competition in health care can be compatible, not only with efficiency, but also with equity concerns. However, market and competition must be considered as independent notions from privatization of health insurers and providers. Indeed, in some health-care system (HCS), competition is fully running within the public sector of the economy, in the so called “internal markets” (Le Grand, 2007).

Consequently, we are going to show that the boundary between private and public in health care not only is tough to disentangle, but it is neither an effective dimension for identifying and evaluating a HCS. However, we have to underline that the approach of this paper belongs exclusively within a economic theory context. Thus, we are not able to properly consider that the differences of HCSs, and particularly the boundaries between private and public, can be explained also by historical differences across countries and consequently by differences on individual preferences, life styles and values.

The plan of the work is as follows. In section 2, the distinction between public (private) provision and public (private) production of health services is examined. From these two concepts, the design of the various *organizational* and *financial models* in industrialised countries obtains, given for granted that organizing industrially and funding a health-care system are two different issues and two different economic problems, although, of course, linked each other. We also attempt to classify organically the several models, in order to include them in a restricted number of conceptual containers. In section 3, we are following a typical *Industrial Organization* (I-O) approach, enlightening the cost-benefit of separation or integration of conveying demand of care agencies (insurance companies, health districts) from productive and supplying entities (hospitals, clinics, fund-holders and medical laboratories). In section 4, we examine the various financing and social insurance systems, with a specific analysis of the so called multi-pillar regime which seems to be the prevailing mixed public/private system in Europe. In section 5, we present some European systems comparisons, according to the dimensions analysed in the paper (separation *versus* integration, different means of financing) and we concludes with some summarizing, although of course not conclusive, comments.

2.- Health-Care Specificity and Government Intervention for Provision and Production of Services

2.1 *Health Services and Individual Well-Being*

Health services, like general practitioner and specialised medical examinations, diagnostic tests, admissions to hospital, medicines, are devoted to treat a state of sickness, aiming at a adequate level of individual health (Culyer, 1989). According to the classical notion coined by Amartya Sen, health is a *functioning*, extensively influencing, as liberty, nourishment and basic education, the individual level of *well-being* (Basu and Lopez-Calva, 2010; Fleurbaey, 2009). However, the individual health level does not derive only from the consumption of health services, but mainly from the efficiency of a peculiar household production process, depending on his/her personal life style. The latter process, on the other hand, works in a social context given by several meaningful variables, as the environmental conditions, the society level of knowledge and culture, the distribution of wealth, etc.

With reference to microeconomics of consumption, health services are private goods, since they are both rival and excludable, but are not direct consumption commodities, as a loaf of bread or a glass of wine. They are, instead, intermediate goods, working as inputs in the household productive process determining the level of “health functioning” (Hurley, 2000; Zweifel *et al.*, 2009, ch. 3; Anderson and Grossman, 2009).

In formal terms, let define as $\mathbf{m}=(m_k, k=1\dots M)$ the vector of M health services to treat a status of sickness, indicated by s , a measure of need. Then, let define with H the level of health reached by the individual, a measure of human capital stock (Grossman, 2000; Becker, 2007). Consequently, the individual i well-being may be represented as follows:

$$(1) \quad U^i = U^i(x_p, H^i)$$

$$(2) \quad H^i = h^i(s_p, \mathbf{m}_i)$$

According to (1), the level of utility achieved by individual i depends on the direct consumption of a private composite commodity x_p , the *numeraire*, and by the functioning “health” H^i . According to (2), i “health functioning” depends on

the initial state of health, s_p , combined with the use of M health services, as described by the production function $h^i(\cdot)$, embedding as well the effects of environmental and social conditions and the life style.

Any health service j has a positive, not-increasing, marginal productivity,

$$\frac{\partial h^i(\cdot)}{\partial m_j} \equiv h_j^i(\cdot) > 0, \frac{\partial^2 h^i(\cdot)}{\partial m_j^2} \equiv h_{jj}^i(\cdot) \leq 0, j = 1, \dots, M$$

and the sign of cross-derivative h_{jl}^i depends instead on the relationship of complementarity or substitutability between services j and l . Notice as the marginal productivity h_j^i formalizes the notion of “appropriateness” of j on health care of individual i , which is specified by the physician along the medical protocols.

In this particular context, the consumer equilibrium is given by the following condition

$$(3) \quad SMS_{j,x}^i \equiv \frac{U_H^i h_j^i}{U_x^i} \equiv SMS_{H,x}^i h_j^i = P_j^i$$

According to (3), the *marginal benefit* of care throughout the service j , $SMS_{j,x}^i$ is a measure of willingness to pay, given by the marginal utility of health (in terms of *numeraire*),

$$SMS_{H,x}^i \equiv \frac{U_H^i}{U_x^i}$$

weighted with the appropriateness index h_j^i .

In equilibrium $SMS_{j,x}^i$ must be equated to the out-of-pocket individual’s *cost* of service, P_j^i , as a fraction of the producer price (marginal cost): $P_j^i = \eta_j^i MC_j; \eta_j^i \leq 1$

P_j^i is the opportunity cost for the individual i , so the lower is η_j^i , the lower is the strength of the scarcity signal sent to agent i . The size of the opportunity cost specifically reflects the health-care financing system. In general, the fraction η_j^i is less than one because the service is not paid at “full price”. According to the prevailing system, the financing of the $(1-\eta_j^i)MC_j$ may derive from insurance *premia*, or from social contributions, or from the general tax system funding a National Health Service (NHS). In the latter case, P_j^i is near to 0 as it may, at the most, include a moderate *co-payment*, (sub-sections 2.2.1, 2.2.2 and section 4).

2.2 *From the Rationale of Public Intervention to Health-Care Systems Typologies*

The modern theory of *Public Economics* singles out a set of reasons for the public intervention in the organization and provision of health services. The motivations follow both efficiency and equity concerns (Hurley, 2000; Zweifel *et al.*, 2009, ch. 5).

2.2.1 *Efficiency and Public Provision of Health Care*

First, although, as said, health services are private goods, nevertheless they often produce *positive externalities*, as their benefits spread also to people not directly consuming them. Vaccinations and the large part of preventive medicine provide the most emblematic cases. By consuming only according to private economic calculus it would determine a social under-provision of the service, thus a public provision for internalizing the effects becomes socially desirable.

Second, health services are often *merit goods*, for whom a social (paternalistic) preference is adding to, or even substituting, the private one. The merit good argument is linked to imperfect (myopic) individual evaluation of the benefit of a commodity consumption. When a health service has this feature, its consumer price should be subsidized, until, if necessary, the total exemption (sub-section 2.2.2 and Appendix, A.1).

Individuals have *imperfect information* on the features and the expected benefits of treatments. In general, health services are *experience goods*, a specific category of commodities whose quality is, according to I-O theory, *ex-ante* unknown, and can be perceived only *ex-post* (Bar-Isaac and Tadelis, 2008; Belleflamme and Peitz, 2010). In health care, it is the services productivity and efficacy to be unknown before the consumption. Sometimes health services tend to be also *credence goods*, when their quality is never ascertained. In this informational framework the crucial agency relationship between the physician (the more informed agent) and the patient (the uninformed principal) cannot be leaved to a private contract, unable to bound the former from inducing the latter to consume what and how he wants for personal aims.

This argument might justify specific contractual relationships between the state and general practitioners, in order to increase the confidence by patients and also to reduce search costs (Mc Guire, 2000; Dranove and Satterthwaite, 2000). Such contracts have generally a private nature, in terms of a convention with the unions of the sector, which disciplines the extent of practitioners' duties and their rewards.

The individual, facing the possibility of health-care spending, given uncertainty and risk aversion, has a clear incentive to buy insurance. However, in this context we have the well known cases of *market failure*, due to asymmetric information between insurer and insured. In particular, the phenomena of *ex-post Moral Hazard* (MH) given by “the third party purchasing”, TPP, context (Nyman, 1999; Pauly, 2000; Zweifel and Manning, 2000; Chalkley and Khalil, 2005), and of *Adverse selection* (AS) (Zweifel *et al.*, 2009, ch. 5 and ch. 6), generally imply the non-existence or the inefficiency of insurance market equilibria. In these cases a *social insurance* can guarantee, at some condition, welfare-improving equilibria w.r.t private second best Pareto-constrained equilibria (Cutler e Zeckausser, 2000; Zweifel *et al.*, 2009, ch. 5 and ch. 6).

In presence of MH, it can be shown that a risk-sharing equilibrium with partial coverage can be improved in a NHS organization by a higher coverage and by a system of controls, provided the decentralised entities are effectively constrained by “hard” budget constraints. However, a system of co-payments may still be desirable for risk-sharing also in a NHS. In case of AS, it can be proved that a *Rothschild - Stiglitz separating equilibrium* (RSSE) can be improved by substituting it with an uniform average premium, a *Community rating insurance* (CRI) with cross-subsidization. Further, this average premium can be also substituted by a system of social health contributions, as payroll taxes, or by specific taxes dedicated to funding a NHS (section 4).

A little bit of formalization on private insurance failure can be useful later on. Let us start with a *ex-post* MH-TPP context. With $y = Y - \pi - E + I = Y - \pi - E + (1-k)E$, $0 \leq k \leq 1$, we represent the individual disposable income, equal to gross income Y less the premium π and the health service expenditure E plus the benefit of insurance (reimbursement), $I = (1-k)E$, where k is the coinsurance rate. $u(y)$ is the status dependent wealth utility¹, with the usual risk-averse hypotheses, and $EU(y)$ the expected utility, where the “sick status” has probability p and the “healthy status” $(1-p)$. The optimal second best coinsurance rate $k^* > 0$ (with a loading factor d) may be given as follows²:

¹ This representation of preference may derive from utility function (1) and (2) by assuming that there is only a treatment and two health status (“healthy” and “sick”) and that the consumer, by purchasing the treatment, can precisely re-establish the healthy status before sickness (HOEL M., 2007).

² The condition is obtained in many theoretical contexts. See for instance CUTLER D.M. and ZECKAUSER R.J. (2000); PAULY M. (2000); ZWEIFEL P. *et al.* (2009, page 237-240).

$$(4) \quad \frac{1 - k^*}{k^*} = \frac{\theta}{\varepsilon_{Ek}}$$

where

$$\theta \equiv 1 - \frac{u_y^s}{(1+d)EU'(y)} < 0; \quad EU'(y) \equiv pu_y^s + (1-p)u_y^b \quad \text{and} \quad \varepsilon_{Ek} < 0$$

is the elasticity of E w.r.t. k . Condition (4) manages a trade-off between the aim of risk-sharing and the aim of controlling the level of over-spending in E , and it is similar to Ramsey inverse elasticity formula of optimal commodity taxation. The level of optimal coinsurance rate can be translated in a optimal co-payment, positively related to price elasticity; so it should be high for service with high elasticity and zero for services with zero elasticity (services for chronic pathologies and diseases)³. The alternatives to the co-payment, for controlling the MH-TPP over-spending effect, may be the consumption rationing by lengthening the waiting lists (Gravelle and Siciliani, 2008), harder budget constraints and taxes increase. All these alternative might be even more distortionary than the co-payment.

Now let us have a look to a standard AS context. Given two types of individuals with high, h , and low, l , probability of illness, $p_h > p_l$, and with fractions of population respectively of λ and $(1-\lambda)$. Let π_i , $i=h,l$, be the premium for a contract giving a reimbursement I_i of the health service expenditure E , with $I_i \leq E$. As said, the RSSE, with the share of low risks sufficiently small in order to guarantee its existence, implies a structure of premiums and insurance coverage as follows

$$(5) \quad \pi_h = p_h E; \quad \pi_l = p_l I_l \quad \text{with} \quad I_l < E$$

while the social insurance Pareto-superior (both types are better off), in terms of mandatory pooling solution, may work with a CRI p^a as follows

³ The co-payment could be, of course, also personalised, *i.e.* graduated in terms of the income and the need of the patient.

$$(6) \quad \pi_b = p^a E = \pi_l; p^a = \lambda p_b + (1 - \lambda) p_l$$

where there is an implicit cross-subsidization taxes-transfers structure from l to b :

$$t_l = (p_l - p^a)E = -\lambda (p_b - p_l)E < 0; t_b = (p_b - p^a)E = (1 - \lambda) (p_b - p_l)E > 0; \lambda t_b + (1 - \lambda)t_l = 0$$

The insurance market failure arises also for those individuals whose probability of illness is near to one (elderly people, poor and socially excluded individuals, chronic invalid individuals) which could be cream-skimmed by private insurers. The coverage of these high risks can be guaranteed only by a social universal insurance system.

2.2.2 Equity and Public Provision in Health Care

Equity in health care pursues a notion of *equity of outcomes*, aiming at guaranteeing the opportunity to reach a given level of health to all individuals, independently on their economic, social and territorial conditions and status (Williams and Cookson, 2000; Wagstaff and Van Doorslaer, 2000). The *final objective* is to allow *human capabilities*, such as the recovery of physical functionality, the absence of complications and a adequate life expectation, and to realise at least a decent level of *health functioning*. In formal term, using (1) and (2), the health-care system should satisfy these two conditions:

$$(7) \quad s_i = s_u \Rightarrow m_i = m_u \text{ and then } H^i = H^u \text{ iff } h^i(\cdot) = h^u(\cdot)$$

$$(8) \quad s_i < s_u \Rightarrow m_i \geq m_u$$

According to (7), if two individuals, i and u , have equal initial health status, they should have the possibility to access to the same vector of services and then to be potentially able to reach the same level of health. The sole differences should depend on the parameters of the household production function and then on the life style. According to (8), the individual i , with a worse initial health status than

individual u , should accede to a vector of services with at least one component strictly greater. Thus, this principle of health equity tends to reduce the differences in the initial level of health along the basic *equality of opportunity* idea of "...more resources for less talent...". However, the ways to actually reach, or approximate, this final result depend on some feasible intermediate objectives. Indeed, several notions of universalistic provision of health care are actually pursued by the various systems, especially for specifying the prevailing funding criteria.

First, we have the simple notion of *Equality of per-capita expenditure*, when the funding of decentralised public bodies is designed to equate *per-capita* public spending, possibly weighted with socio-demographic variables. Second, we have the *Equality in satisfying standardised needs*, when the funding is referred to the notion of Essential levels of health care (ELC) and standardised costs per treatment. Third, it is frequently evoked the notion of *Equality of access to services*, essentially meaning that the providers are obliged to guarantee an equal treatments to patients, independently on individual risk (no cream-skimming) and income (no wealth discrimination). Finally, we may remind the concept of *Equality of individual payments*, when a uniform per-patient cost for health-care treatments is required (no price discrimination). These intermediate objectives are actually attainable, but, could be conflicting each other and, in any case, they obtain only second best equity conditions, as they are imperfect and constrained respect to the first best one (fulfilment of a adequate level of "functioning health").

Particularly meaningful is the second notion, referring to the criterion of guaranteeing a essential package of care covered by public funding. For explaining this in Appendix (A.1) we'll follow the merit good argument along a simple formal model (Schroyen, 2005, 2010).

2.2.3 *Industrial Configuration Failure and Public Production of Health Care*

The previously mentioned equity and efficiency concerns offer the rationale for public intervention in terms of *public provision* of health services, *i.e.* by funding them outside price-setting, independently on the nature and ownership of the chosen providers. However, for some treatments, it may occur situations of "industrial configuration failure" where is socially desirable to join together *public provision* and *public production* as well. When it is convenient to localize services provision in urban areas where the private returns to invested capital are higher and there are economies of scale, local monopolies can easily take place for extracting spatial rents. Elsewhere, with lower expected returns, it may instead turn out a scarce network of providers and then an excess-demand of services.

In these cases, the public sector should be engaged not only to funding the services, but also to providing them directly, in order to support a fair territorial distribution of supply. The inadequacy of supply and the rationale of public production also occurs when, in some territories, the network of private providers, although numerous ones, cannot supply high-quality services. In these cases of inefficient industrial structure, public production, if well organised, is justified both for efficiency and equity aims.

2.3. Health-Care Systems and Coverage of Risk Disease

The distinction between public (private) provision and production is crucial in order to analyse and classify the several models of health care. Indeed, the production argument allows us to talk about *health-care organization* of supply, while the provision argument to talk about *health-care financing* of medical expenditures. Both these two items contribute to define the features of a HCS.

We may distinguish two polar organizational cases: the *Public system* and *Private insurance system*. As far as the first one is concerned, there is the purest version, the so called *Beveridgean* model, which considers full public provision and production of health services, both financed by general taxation. Essentially, the English and Italian NHSs, before the reforms of the last two decades, were following this model. Also the so called *Bismarkian* model is a public model. It is the original social insurance model financed by health contributions going directly to public Health Funds. In principle, both *Beveridgean* and *Bismarkian* models are universalistic ones.

As a *Private insurance system* we can mention that one presently still working in US. In this system there is a prevailing structure of private insurances policies and only some public programs for specific categories of patients and limited fractions of population: the *Medicare* for elderly people (over 65 years-old) and *Medicaid* for poor people (selected by means testing). This health-care system is coherent with the concept of “residual welfare state” implying a limited coverage of every social risk⁴.

The two polar models are now mainly academic ones and statistical outliers. For instance, all European countries are now applying *mixed systems*, where public and private programs are acting simultaneously for guaranteeing a universalistic

⁴ For *Census Bureau* about 46 millions of American citizens are not assured for illness risks. They are mostly less than 65 years old individuals and with a average income, so they are not covered by either *Medicare* or *Medicaid*.

provision of health services. Also the American system, according to the Barack Obama's reform which introduces a set of innovations w.r.t. the previous *Private insurance system*⁵, is going to be transformed in a mixed one (Sullivan, 2010).

The long and deep discussion on this reform clearly has shown that, in order to evaluate the different systems, it is crucial to precise what is meant with degree of *health-care coverage* of each one. However, this notion is not unique as it may be referred to: (i) the *extent* of coverage, *i.e.* the share of population whom health care is guaranteed; (ii) the *depth* of coverage, *i.e.* the number and the features of services included in the insured package; (iii) the *highness* of coverage, *i.e.* the fraction of treatment costs directly financed by the insurer or the NHE and then not directly paid by patients. This classification must be taken into account for avoiding confusions in looking for what is simply named a universalistic system.

In order to distinguish the variety of mixed systems we have to consider several features. First, the *criteria* according to which the general practitioner (the agent who makes the order of purchasing the service) is assigned to each household. Then, the individual degree of choice of the provider by which to obtain health-care services, as diagnostics and specialist treatments and hospital admissions, is meaningful. In this respect, there are systems allowing a high individual freedom of choice and other ones with a rigid assignment by public administration, but, of course, there are several intermediate situations. Another relevant distinction refers to the ways of paying the providers, *i.e.* if it is allowed an *ex-post* coverage of the production costs, or it is fixed an *ex-ante* budget, or it is organized a system of prospective standardised tariffs for each treatment (Appendix, A.2). Further, the mixed systems may be differentiated according to the forms of organizing the supply of drugs and to the ways their prices are established, and the co-payments and coinsurance rates structure is designed.

According to these *criteria*, the literature in health economics usually distinguishes three specific types of mixed systems. The first one is known as the *reim-*

⁵ First, Obama's reform introduces some measures for protecting the insured individuals, or enrolled at a *Medicare* o *Medicaid*, against rent seeking behaviours of insurance companies, like forbidding both *cream skimming* and the rule of fixing a maximum amount of care, while foreseeing a limit for *out-of-pocket* expenditure. Second, the reform tends to create a publicly regulated market, where several insurance companies competitively supply bundles of homogeneous services, according to some common rules. Further, all the American citizens will be obliged to buy a health insurance, consequently individuals or firms unable to pay for the premiums will receive adequate tax allowances and benefits. This *new insurance exchange model* will be financed by general taxation and the expected efficiency gains from a higher competition respect to current situation where local insurance monopolies prevail in many states.

bursement model, where insurance companies or decentralised health districts (or regional governments) reimburse the patients expenditures, after they have paid, at administered prices, the services to public as well private providers. The second system is the *integrated model*, where the health district builds up with the providers a unique connected public structure. The patients do not pay for the treatments which are financed by taxes, but they have a limited or no choice where to receive them. The internal hospitals and clinics are financed at costs of inputs or, sometimes, with reference to a fixed budget. Finally there is the *contractual model* where the funding body, the insurer or the health district is separated from the providers which, whether public or private ones, are committed and rewarded according to a procurement contract. This is based on prices that in most cases are fixed *ex-ante* and standardised along the system of Diagnostic related groups (DRG). Consumers-patients generally have, according to the variety of contractual systems, freedom of choice of the provider and the general practitioner they desire to engage. *Quasi-markets (QM)* are a specific typology of the third model, often called managed or internal competition, applied in the last decades in some European countries (section 3.2).

In the in block diagram of Graph 1, we summarize how *production of health services* can be developed according to the various organizational systems while, in Graph 2, we summarize how *provision of health services* can be carried out, according to the various funding systems. In the blocks, some countries/systems combinations are inserted as reference.

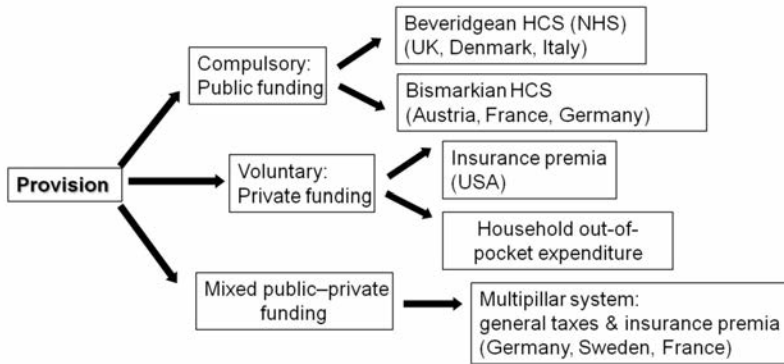
GRAPH 1

CLASSIFICATION OF HCSs ACCORDING TO PRODUCTION ORGANIZATION FEATURES



GRAPH 2

CLASSIFICATION OF HCSs ACCORDING TO PROVISION AND FINANCING FEATURES



The next two sections, 3 and 4, will expand such a double classification within respectively the production/organization and provision/financing frameworks.

3. - Health-Care Industrial Organization of Production: Separation *versus* Integration

3.1 *The Cost-Benefit of Vertical Separation (Integration) in Health-Care Industry*

The vertical separation of purchasing structures from producing and supplying ones is a necessary but not sufficient condition for creating competition in health-care industry. Thus, this institutional design issue must be treated *per se*, by looking at cost-benefit of vertical integration in terms of production costs level, economies of scale and scope, network and coordination economies. Within this context, health economics can well consider the *pros* and *cons* of different organizational systems. Indeed, the transactions between a buyer, *e.g.* a health district, and seller, *e.g.* a hospital, can be realized by a market contractual exchange (in case of separation) or by a internal transfer within a unitary body (in case of integration). Therefore, for modelling the determinants of this choice, we have to consider the usual alternative between hierarchy and market, by comparing the administrative costs of managing a complex structure (the *firm*), with the costs for signing incomplete contracts with the providers (the *market*)⁶. Further, a va-

⁶ For general surveys of this topic see GIBBONS R. (2005); BOLTON P. and DEWATRIPONT M. (2005, ch. 10).

riety of industrial issues matter, like the degree of complementarity of the purchasing and production assets, the *rent-seeking* behaviours of the several agents, those making the choice and those applying it for the relevant activities, the existence of sunk costs and irreversible investments and the consequent hold-up issue. Of course, all the transaction costs concerns, as the contractual size and complexity, and the time and costs requested for settling the eventual controversies on trial are also relevant (Williamson, 2005). Thus, by limiting the opportunistic behaviours and *ex-ante* and *ex-post* contract inefficiency, the integration will be preferable to a system based on market exchanges and transactions. This will happen when its advantages can exceed the static and dynamic inefficiencies due to huge bureaucratic centralised structures, typical of public administration.

We can formalize the different typologies of I-O schemes as follows. Let us employ these notations:

A = set of institutions, such as national or regional authorities, devoted to regulating a public and/or private production,

B = set of agencies conveying the users demand and needs, and

C = set of the suppliers of treatments and final services.

Now, we have

TABLE 1

PRODUCTION ORGANIZATION SYSTEMS		
Pure integrated model	$A \cup B \cup C \equiv \Omega$	the union of the three sets, Ω , assembles all the health-care activities
Separated QM model	$A \cap B \cap C \equiv \emptyset$	the intersection of the three sets is empty because each set interact with the others by contractual exchanges
Semi-integrated model	$(A \cup B) \equiv \Psi \cap C \equiv \emptyset$	<i>A</i> and <i>B</i> are joined in a union set Ψ , but the intersection of the latter with <i>C</i> is once again empty: Ψ buys services from <i>C</i>

In a *de-integrated (separated) model*, $A \cap B \cap C \equiv \emptyset$, as the Dutch one, there is a limited degree of industrial concentration. The tasks of expressing, on the behalf of patients, the demand and of establishing the appropriateness of services and treatments are delegated to a specific Authority, a public Sponsor of citizens. The providers are government or non-profit institutions, but they must be, in any case, appropriately *ex-ante* selected as “reliable providers”. The selection procedure aims at controlling and limiting the provider power in the negotiation due to asymmetric information and at assuring that the several necessities and urgencies of users are suitably fulfilled.

In a *public-public integrated model*, $A \cup B \cup C \equiv \Omega$, like that one prevailing in some European regimes, there is a strong industrial integration, that, in the polar version, unifies in a whole structure, a “Local health firm”, as the ASL in Italy, all the functions of planning, demand rationing, financing, production and supply of services. In milder versions of the model, $(A \cup B) \equiv \Psi \cap C \equiv \emptyset$, some hospital firms (AO) may be separated from the ASL, still remaining publicly owned. This *semi-integrated* configuration is wholly working in UK, where there is a *Health District Authority* (HD) separated from the *Trust hospitals* (TH), which patients can almost freely choose as favourite.

The main objective of the integration is to improve the capacity of coordination by the planner, by limiting the conflicting interests among the involved agents. This kind of conflict is, instead, considered beneficial by the advocates of the de-integrated model. Further, the integrated model is aimed at limiting the opportunistic manipulation of demand, which is often not objectively determinable, mainly on the social-assistance components. However, the integrated model has all the shortcomings and defects of the centralised and complex structures, at high level of bureaucratization and high administrative costs (Dixit, 2002), and it suffers the politicians interference, in all the ways it can occur. Moreover, the model does not give adequate incentives toward the internal efficiency of the institutions (Le Grand, 2003, 2007), whose lack becomes the main cause, together with the so called “soft budget constraint syndrome”, of the wide sunk deficits of the decentralised bodies (Rodden *et al.*, 2001; Wildasin, 2004). Consequently, on looking at the *cost-benefit* of separation *vs.* integration, considerations of *Political economy* should be taken into account. The application of new *Political economy* models (Besley, 2007) to health economics allow to contemplate the behaviours and the conflicting relationships among citizens and politicians, the public providers and the interests groups, as the pharmaceutical industry and doctors profession (Zweifel *et al.*, 2009, ch. 13).

In conclusion, we summarize the *pros* and *con* of institutional design based on separation as follows.

TABLE 2

PROS AND CONS OF THE SEPARATION IN INDUSTRIAL HEALTH ORGANIZATION

<i>Benefits</i>	<i>Costs</i>
positive incentives from the interests conflict matching the demand by a ASL (HD) and the supply by the providers, as a AO (TH)	a limited exploitation of economies of scale and scope (no gains from increasing returns and complementarities in cost structure)
positive incentives from competition-quality-choice conduct arising from purchasing contracts	high transaction costs for signing and implementing highly incomplete contracts
more transparent accountancy procedures and better performance measurement	lack of whole corporate view of the three management functions and consequently difficulty to effectively control demand and supply
more management specialization on the three main functions and responsibilities: <ul style="list-style-type: none"> • <i>Insurance</i> (risk and need perception) • <i>Production</i> (measurement and control of costs and returns) • <i>Purchasing</i> (demand input controls) 	phenomena of demand induction from providers mainly boosted to increase the revenues
controls of costs and quality by Health Authority, acting as “aware bidder contractors” and not as a “blind purchasers”	

3.2. *Separation Plus Competition: The Quasi-Markets*

A clear separation of structures, with , is typical of *Quasi*-markets (QM), which are, as said, the most advanced form of the contractual model⁷. The features and the cost-benefit of such a model can be summarized as follows.

⁷ See MAYNARD A. (1994) and JONES P.R. e CULLIES J.C. (1996) for two formal analyses of the working of an emblematic QM, like the English one during Margaret Thatcher’s era. For a systematic evaluation of the today NHS, after Tony Blair partial revisions of the latter, see the OCDE Report by SMITH P. e GODDARD M. (2009).

TABLE 3

MANAGED COMPETITION AND *QUASI*-MARKETS IN HEALTH CARE

<i>Advantages</i>	<i>Disadvantages</i>
limiting the productive role of the state, which could better specialize its effort working only as purchaser of services, on the behalf of the citizens enrolled in the NHS	possibly boosting <i>cream-skimming</i> procedures by providers
enhancing a beneficial competition among providers (mainly public ones). It is a <i>monopolistic competition</i> on the quality, as the prices of treatments are generally fixed	requiring complex <i>ex-ante</i> and <i>ex-post</i> performance controls by a public body, or a specific Authority, able also to properly regulate the internal market with plans and hard budgets, efficient bidding and contractual activities
pushing up the <i>voice</i> of patients-consumers through the role of sponsor played by the insurers	enhancing, contrary to what expected, both private and public health expenditure, owing to phenomena of moral hazard and of inducing distorted demand by highly competitive providers, engaged to acquire as many customers as possible
allowing the <i>exit</i> of patients-consumers, giving them the freedom of choosing the preferred provider	lack of incentive to efficiency if, in some areas, competition among hospitals is simply not possible, so local monopolies actually arise (<i>industrial configuration failure</i>)
boosting cost-containing and quality enhancing actions	together with the incentive to increase observable features of quality, there may be the incentive to decrease the not observable ones ⁸
boosting R&D activities in medical industries, as the hospitals, being “residual claimant”, can reinvest the “profits”	freedom of choice is somewhat misleading given the imperfect information of patients and the potentially non-benevolent pay-off and behaviour of the physician prescribing the treatments.

From Table 3, it turns out that managed competition may be socially desirable in some institutional and economic contexts but in other ones it actually does not work. Indeed, only a limited group of European countries have pushed their HCS reforms in toward a pure contractual-QM model. We may say that now in

⁸ PROPPER C. *et AL.* (2008) have shown that Trust hospitals competition in NHS, although has reduced average waiting lists (observable feature), has increased other non observables relevant features of quality, like death rates.

The Netherlands, UK, Germany, Sweden, after several succeeding reforms in the nineties, *quasi*-markets organized in potentially pro-competitive systems are prevailing. In countries like Spain and Italy there are instead HCSs with only some elements of managed competition. Indeed, although they intended to introduce internal competition features, actually gave up to fully apply them, emphasizing instead their limits. Further, there are countries, like France, Denmark, Finland, that, even if aware of the defects of the purely planned systems in terms of incentive to efficiency, have introduced only very few or no pro-competitive items⁹.

In the models where it is $A \cap B \cap C \equiv \emptyset$ or $(A \cup B) \equiv \Psi \cap C \equiv \emptyset$, the different techniques for rewarding the providers are crucial¹⁰, as imply different cost-containing and quality-enhancing incentives (Ma, 1994). The simple analytic model presented in Appendix (A.2) may enlighten these features.

3.3. Ownership Structures and Investment Task Assignments in the Hospitals

The trade-off between cost-containment and quality can influence also the institutional design and the ownership options for health structures. As regard, particularly meaningful are recent extensions of the theory of property rights with incomplete contracts, treating the issue of delegation of responsibility on assets running, according to the incentives by private as well public managers to productive efficiency and quality of services (Hart, 2003; Sadka, 2007). The segmentation of a public firm in different branches, with out-sourcing of some purchasing and selling activities, *e.g.* the separation of hospitals from the public health authorities, could be justified in terms of the assignment of investment tasks, *i.e.* the convenience of transferring the responsibility for certain elements of the treatments to private hands, while maintaining other elements in public hands.

Hoppe and Schmitz (2010) have recently developed a model where contracts on the privatization of infrastructures projects, like new hospitals, do not only specify the transfer of ownership rights, but also assign the responsibility, regarding design construction, maintenance and modernization of the structure itself. Thus they consider a set of combinations public *vs.* private activities which appear

⁹ Notice as the model may even change among regions or states of a federal country, with decentralised health care, as we may say it happens in Italy, where in Lombardy there is a QM system and in Tuscany a somewhat command & control system.

¹⁰ The topic has been analysed by a large and complex literature, surveyed, among others by CHALKLEY M. and MALCOMSON J. (2000); DRANOVE D. and SATTHERWAITE M.A. (2000); BOADWAY R. *et AL.* (2004) and ZWEIFEL P. *et AL.* (2009, ch. 10).

very promising for analysing health-care industrial organization. Let take into account a situation where, at date 0, the government (G), in our case a health district (HD), and a manager (M) of a hospital write a contract that specifies a volume q of treatments, with features described *ex-ante*, and a payment T from G to M . When the manager provides the treatments he incurs costs qC_0 , while the health district's benefit is given by qB_0 , where $B_0 > C_0$. The parties also agree on an ownership structure and an investment task assignment.

The *ownership structure* $o \in \{M, G, J, N\}$ determines who is in control of the hospital assets and equipments. Under private ownership ($o=M$), the manager has the right to modify the assets in order to implement innovations that may improve quality or reduce the production cost, enhancing x -efficiency. Under public ownership ($o=G$), the HD controls the essential infrastructures. There are also two kinds of partnerships. Under $o=J$, both parties have veto power, so that no one can implement any innovations without the partner's consent, while, under $o=N$, each party has the right to modify the assets to implement innovations. The *task assignment* $A \in \{M, G, MG, GM\}$ determines who is in charge of the two types of non-contractible investments, for reducing costs of providing treatments and enhancing quality, that can be made at date 1. The first type of innovation reduces the manager cost per unit, but at the same time it also lowers the quality and thus reduce the government's benefit. The second type of investment results in a innovation that improves the treatment's quality, so that the HD's benefit per unit increase, but also increases of the hospital's cost per unit. The assignment $A=MG$ (resp. $A=GM$) means that the hospital manager (resp. the HD) is in charge of cost-containment investment task and the government (resp. the manager) is responsible for quality-enhancing investment task. The party in charge bears the associated investment costs. Finally, at date 2, the parties may renegotiate the quantity of the treatment to be provided, the decisions whether or not to implement the innovations, and the payment. The negotiations are modelled as Nash bargaining solutions, according to which the renegotiation *surplus* is divided between the parties.

Let us concentrate on Hoppe and Schmitz propositions on the cost-benefit of ownership structure and tax assignment, leaving aside the propositions regarding the volume of treatments, and let consider the case $o=G$, where the ownership of the hospital is public one (proposition 5ii). This case emerges as optimal when the negative side effect of the quality innovation is negligible. It is shown that, if M has a larger bargaining power than G , the preferred task assignment is $A=MG$, then M should be responsible for the cost investment and G for the quality in-

vestment. While, if G has a larger bargaining power than M , the preferred task assignment is $A=G$, then G should be responsible for both investments. Further Hoppe and Schmitz (proposition (6)) show that a partnership with no veto power ($\theta=N$) can be optimal only if the parties' bargaining strengths do not become too asymmetric. When one party's bargaining power becomes large it becomes impossible to find the quantity of production that balances the incentives for making both types of investments.

Consequently, Hoppe and Schmitz results suggest, first, that partnerships between the public and the private sector in running a hospital should be carried out when the bargaining power is relatively balanced and if the side effects of cost and quality innovations are relatively unimportant. Second, they imply that public (private) ownership should prevail if the side effects of cost (quality) innovations are strong.

In this respect, it is worth to notice that the private bargaining power is inversely related to the degree of competition among hospitals. However, the government's bargaining power is likely to be weak *ex-post* since finding an alternative suppliers, during the renegotiation, is usually difficult. Moreover, it is empirically proved that the effect on hospital production costs of improving quality of treatments is quite high, given the needed sophisticate equipments. While, if the hospital is far from x -efficiency, a innovation reducing costs could imply relatively low side effects on quality.

3.4 *The Alternative of Non-Profit Organization for Hospitals and Nursing Homes*

The ownership issue for hospitals and other facilities is not restricted to public/private dichotomy only. An intermediate solution is to build up a non-profit firm. Non-profits organizations generally espouse goals other than profit maximization, such as, fairness, education, preservation of values, etc. For analysing the strategies chosen by an hospital management to pursue these ends we have to recall that its activity has a multi-product nature.

For a non-profit firm some of products are "preferred", *i.e.* mission-oriented services, although privately unprofitable goods, and other are "non-preferred", *i.e.* minor services, although profitable revenue-goods. An efficient management of the hospital should consider this feature as an opportunity, as producing at the maximum profits the "non-preferred" products may provide means to the ends pursued by the "preferred" products (Weisbrod, 2006). In other words some hospitals, as well as nursing homes, although legally defined as non-profit firms,

with the consequent fiscal advantages, may seek to maximize profits derived from some but not all of its activities. Such a cross-subsidization could be beneficial from a second best perspective tending to maximize the implicit multidimensional objective function.

Several empirical researches has examined mixed industries, mainly hospital, facilities for the mentally handicapped, nursing homes, showing behavioural differences in many dimensions: efficiency, mortality rates, satisfaction of staff members and among patient's families. In particular, (Weisbrod, 2006) has shown that observed behavioural differences across institutional forms reflect two specific perspectives. On one hand, different institutional forms face diverse constraints. For instance, w.r.t. profit firms, non-profit ones of course face a "nondistribution constraint" limiting the size of profit. While, w.r.t. governmental agencies, the non-profit institutions have access to volunteer labour and private donations of money. Moreover, according to the different institutions, Weisbrod (2006) has also shown that diverse organization objective functions arise, as they are manifest in managerial reward systems. In case of non-profit firms, the empirical analysis shows that differences in reward system may arise from strong incentives in the form of performance-based bonus to optimize their organizational objective functions, according to the several ends they pursue.

Grabowsky *et al.* (2009) examine the increasing market of home care services in US. These are mainly provided by private for profit agencies, reimbursed at cost-based payment systems, but recently also by an increasing number of non-profit firms. They show that the non-profit status is associated with shorter enrollment in home health care. Moreover, they show that non-profit home health-care agencies, because they lack a residual claim, are less sensitive to a change towards a prospective payment system. With stronger incentive-correlated ways of reimbursement, patients at non-profit (profit) agencies have a lower (larger) decrease in length of home health care and lower (greater) increase in discharges without the goals of care having been met. This empirical result can be theoretically sustained by the model offered in Appendix (A.2).

Thus, in case of a non-profit enterprise running an hospital or a agency, the simple model of Hoppe and Schmitz (2010) of previous section 3.3 should be extended in order to admit more complex contracts between G , the health district, and the manager, M . This should specify an *ex-ante* multi-product profile of treatments, associated with the several goals to be pursued, other than efficiency. The features of the payment rules will be more compound and, at the same time, constrained by much more relevant asymmetric information nuisances.

4. - Health-Care Provision: Compulsory *versus* Voluntary Ways of Financing

4.1 *Health Insurance Framework for Funding Systems*

We may say that the logic of insurance applies more or less to all health-care funding systems, also to the strictly public ones. Indeed, a treatment is, in any case, a risky event whose coverage should be guaranteed to risk averse individuals. The financing systems are generally distinguished according to the type of interaction among the three main players in the health-care game - the *patient*, the *provider* and a *regulating entity*¹¹. Thus, a significant distinction comes from the way the patients (and the tax-payers) finance the institution/insurer. More specifically, we may distinguish as follows:

TABLE 4

FUNDING HEALTH-CARE EXPENDITURE

<i>Instruments</i>	<i>Insurance systems</i>
insurance premiums based on individuals risk and administrative costs	voluntary private insurance policies
community rating insurance (CRI), based on average risk	compulsory public insurance
health contributes (payroll-taxes) paid by employees and employers	compulsory social insurance by categorical funds of non-profit enterprises
general taxes devoted to health-care spending	standard National Health Service (NHS)

It is clear that the pure insurance mechanism is running only in case of voluntary private insurance, where each policy breaks even and premiums are “fair”. In case of insurance with CRI an explicit cross-subsidization from low risks to high risks is at work. In the other cases, subsidization is still present, although not explicitly. In the third one, in fact, the premium is arranged as a fraction of labour income, possibly with upper limits to the payment. Thus, as the risk of sickness is inversely correlated to the labour productivity (skill) and then to the income, also in this case, low risks tend to subsidize in the average the high risks. In the last case, the NHS, the insurance mechanism is of course quite mild, but actually it remains in aggregative terms. Indeed, the amount of resources coming from taxation and devoted to funding health care must cover *ex-ante* the aggregate

¹¹ This is often called the *Medical care triad*, see CUTLER D.M. and ZECKAUSER R.J. (2000).

value of insured risks, measured by the *Essential levels of Health Care* (ELC). However, in this case, the logic of insurance is inevitably violated if, removing such a constraint, the deficits are *ex-post* bailed-out by the government.

When the insurance coverage is funded by general taxation, trade-offs arise between redistributive and insurance aims. In particular, being risks, as said, inversely correlated to income, the social coverage rate may reach redistribution objectives, helping the parallel role played by progressive taxation. Indeed, it has been shown that the former can moderate the distortion effects of the latter (Rochet, 1991; Cremer and Pestieau, 1996)¹². We are going to analyse this issue in the following sub-section, here we propose a synthetic picture of the main OECD countries funding systems.

TABLE 5

FINANCING SYSTEMS OF HEALTH CARE IN OECD COUNTRIES

NHS supported by general taxation (<i>Beveridgean</i> models)	Australia, Canada, Denmark, Finland, Iceland, Ireland, Italy, New Zealand, Norway, Portugal, Spain, Sweden, UK
Social insurance systems financed by mandatory payroll contributes to Social Insurance Funds (<i>Bismarkian</i> models)	Austria, Belgium, France, Germany, Hungary, Japan, Luxembourg, The Netherlands (before 2006 reform), Poland
Systems of private insurance, although compulsory, financed by CRI	Switzerland, The Netherlands (after 2006 reform)
Mixed social systems, as a combination of the three previous cases	South Korea, Greece, Turkish, Slovakia, Czech Republic
Systems with prevailing voluntary private insurance	US, with <i>Medicaid</i> e <i>Medicare</i>
Mixed private systems financed by direct payments by patients	Mexico

Source: OCDE, *Health Data*, year 2009.

Mixed systems are prevailing in the groups of European welfare tradition countries. In particular in the first group, the Beveridgean NHSs, beside general taxation there are also social contributions or earmarked taxes on wages. This happens in UK but also in Italy. In the Bismarkian social insurance systems of

¹² For a analyses of European reforms in terms of integration or separation between redistribution and insurance aims see BREYER F. and HAUFLER A.(2000).

the second and third group, the insurance funds are frequently integrated with resources from taxation system. US and Mexico systems are somewhat outliers. Quite interesting is the cross analysis of separation issue and funding issue to understand the complexity of health-care systems. For instance, the UK NHS, though funded by general taxation, is characterized by an internal market with high industrial de-integration and large freedom of choice by patients (Smith and Goddard, 2009), while, in the classical US private insurance system are working the *Health Maintenance Organizations* which follow a integrated model of industrial organization.

Several European countries have a health-care funding system structured in two or three “pillars”, but with meaningful differences in the size of each one (Mossialos and Thomson, 2004; Zanette and Ricatti, 2006; Gechert, 2009). A first pillar is devoted to socially cover a uniform basic package of services (ELC) and a second one is devoted to a private integrative insurance for the remaining services. The existence of a partial social coverage in most NHSs is signalled by the several forms of rationing publicly provided services and treatments, according to their “appropriateness” and cost-effectiveness. The pillars may be three when a second pillar can be inserted within the previous two ones, with publicly regulated capitalization health funds. This second insurance setting, mainly organised by public municipalities and regions through “reliable” companies, plays the fundamental role of limiting the size of the purely private health spending remaining in the third pillar.

The second pillar may refer to *supplementary insurance*, when it provides a double coverage to services in the statutory package. It could be the case to offer a perceived quality advantage over the care secured by the statutory package, *e.g.* in the form of reducing waiting times or access to superior facilities. The integrative insurance may be also *complementary* when offers full or partial cover for services excluded or not fully covered by the statutory package, like Long Term Care expenditures (in Germany), and for covering the liability for co-payments levied on ELC services (in France).

4.2 On the Theoretical Foundation of a Multi-Pillar Funding System

The rationale and the features of these mixed systems have been widely investigated by theoretical health economics, following two streams of literature. The first one straight considers the social convenience of a public community rating insurance plus a supplementary private insurance. The second one analyses the more complex and more realistic case of three-pillar systems with distortionary

income taxation financing the social component. We postpone this subject to the successive sub-section.

We may say that the topic has been originally introduced in the theory of health insurance by Wilson (1977) seminal contribution. Wilson moves from a classical adverse selection *Rothschild-Stiglitz separating equilibrium* (RSSE), where, since the low risks are rationed, there may be scope for Pareto-improvements. The main result is that an improvement can be achieved by introducing compulsory insurance coverage. Wilson has modelled a health system where there is a mandatory public policy, just breaking even because lying in the pooling line pertinent to the population at large, and individuals are allowed to buy supplementary health insurance. Given this framework, it turns out that the pair of contracts (compulsory plus voluntary part) Pareto-dominates the RSSE without social insurance (Zweifel *et al.*, 2009, page 177). As shown also by Dahlby (1981), the new mixed equilibrium gives a full (partial) coverage insurance to high (low)-risks, as in RSSE, but now both contracts are preferred to the previous ones. The high-risk types benefit from cross-subsidization within the public insurance contract, while low-risks types are better off given the relaxation of the rationing restriction, allowing an increased overall coverage as compared to the absence of mandatory public insurance with partial coverage. This mixed two-pillar solution is even Pareto-superior to a one-pillar public insurance with a uniform community rating for all individual. In fact, by allowing an appropriate risk-selection, it reduces the effect of the tax burden to low-risks which have to pay for their insurance and for cross-subsidizing high-risks.

Hansen and Keiding (2002) extend Wilson-Dahlby approach by providing a welfare comparison between three alternative health insurance regimes: (i) a unregulated voluntary market regime, where consumers either buy full coverage or no coverage; (ii) a compulsory and uniform universal regime with community rating; (iii) a compulsory, universal community-rated regime that allows private supplementary insurance. The novelties of their analysis are twofold: first, in order to make the welfare comparison between the community-rated insurance (CRI) regime and the voluntary private insurance one, they derive the level of the compulsory insurance as the equilibrium choice in a median voter model of political economy. Second, they base their comparison on either the Hicksian compensation principle (where the winners could compensate the losers), or a utilitarian average utility concept. These welfare criteria appear to be more meaningful than the simple Pareto-criterion, considered by Wilson-Dahlby approach, since compulsory insurance regimes are likely to make low risks worse off, assuming that the level of

mandatory coverage regimes exceed the level that would be chosen voluntarily by low risks. Hansen and Keiding show that their voluntary full coverage reaches a higher average utility than the compulsory insurance determined by the median voter, under plausible hypotheses on the distribution of risks. Low risks are better off under the compulsory insurance status because they purchase insurance at reasonable cost, while they remain uninsured in the voluntary setting. High risks, instead, are worse off with compulsory coverage because they have less coverage than they would choose to buy in the voluntary market equilibrium, but this drawback is reduced if they can purchase risk-rated supplementary coverage. Hence it is welfare improving to permit supplementation to a compulsory CRI scheme.

However, this result is not fully robust: changing slightly the context, we may reach different outcomes. Kifmann (2002), for instance, shows that if insurers that offer the basic CRI benefits are allowed also to offer the supplementary ones, the regime can only benefit low risks at the expense of high risks. The finding depends on Kifmann's assumption that the community-rated insurers offer supplementary benefits that disproportionately attract low risks. Consequently a separating equilibrium is reached where the high risks buy the basic package from community-rated insurers, while the low risks buy the supplementary coverage from other insurers. This has the effect of reducing the transfer from low to high risks, as desired by the government, and leads the low risks to consume benefits worth less than their cost, because they avoid the cross-subsidy to high risks. This result seems more plausible if the supplementary package covers specific services that only appeal to relatively healthy individuals.

Danzon (2002), by comparing, in a special issue of *Journal of Health Economics*, the contributions by Hansen and Keiding and Kifmann, concludes that «...the welfare effects of permitting supplementation of compulsory, community-rated insurance depend critically on who buys the supplementary coverage and how the supplementation affects the rating of the basic coverage. Supplementary coverage is more likely to be welfare improving if it is bought primarily by high risks, for whom the basic coverage is presumably suboptimal. This result is more likely if the supplementary benefit is offered by separate insurers».

Different insurance regimes have impacts not only in terms of efficiency but also in terms of wealth vertical redistribution. In this respect, Leach (2010), from a pure redistributive perspective given by a social aversion to the inequality¹³,

¹³ There is a strictly concave *ex-post* social welfare function over agents who are *ex-ante* identical, differing *ex-post* in the state of their health.

looks for the optimal health-care system among a fully private insurance, a fully public (social) insurance and a mixed insurance (public insurance with a supplementary private one). By comparing the first two extreme solutions, the author shows that with the optimal public health-care system, people who are relatively unhealthy, with and without treatment, receive more health care, and people who are relatively healthy, with and without treatment, receive less health care. However, the aggregate quantity of health care under the optimal public regime might be either greater or less than under private one. When there is a mixed system, with both public and private insurance, the latter covering treatment for types that are not covered by the former, it may happen that, in an equality-averse society, the role of private insurance is almost negligible. The parallel system of private health insurance tends in fact to reduce welfare with respect to an optimally designed entirely public system.

The first stream of literature is, other than fairly inconclusive, not completely satisfactory in modelling a multi-pillar system. Even if we derive, as in Hansen and Keiding (2002), a social preference for a mixed system, where a public compulsory insurance, covering a package of essentials, is integrated by a private policy topping up the remaining services, we have no indication on the optimal composition of the total insurance, *i.e.* the number and the size of the pillars. The most convincing analyses for analysing multi-pillar health insurance systems come from contributions modelling the social insurance in the first pillar as a redistribution device financed by a distortional income tax¹⁴. This kind of model reflects the phenomenon of frequent interaction between social and private insurance now characterizing many sectors of welfare states in industrialised countries (Chetty and Saez, 2010).

4.3 *Three-Pillar System with Optimal Income Taxation*

Blomqvist and Johansson (1997) and Selden (1997) considered a multi-pillar public/private health insurance system, but without an income-dependent contribution to social insurance with a variable labour supply. Petretto (1999) extends the two contributions in this way. He models a NHS with a supplementary private insurance, but where low-risk or rich people are not allowed to opt-out of the NHS. Thus a three-pillar system is modelled so that the health services purchase by an individual is matched by three funding components: the social in-

¹⁴ As said, in the previous sub-section, this issue has been firstly analysed by ROCHET J.-C. (1991) and CREMER H. and PESTIEAU P. (1996), but without investigating a multi-pillar setting.

surance, the supplementary private insurance policy and the out-of-pocket expenditure. The social insurance is financed by a linear income tax, whose structure is optimally chosen, together with the social insurance coverage, by the government. This faces n types of individuals distinguished by two parameters, the probability of illness and skills (wage in efficiency units). There is *ex-post* moral hazard with regard to health expenditure and adverse selection from the government perspective, as it cannot observe individual's skill parameter. Moreover, although the probabilities of illness are however known, the government cannot differentiate its tax policy between high-risk and low-risk individuals, but only according to their personal incomes.

In this framework, the individual's equilibrium private coinsurance rate follows the standard condition of optimal insurance with *ex-post* moral hazard. This requires, as in condition (4) in section 2, to equalising, at the margin, the *gain of risk-sharing* with the deadweight loss due to *moral hazard effect* for each individual. The private insurance contract, signed by individual i , includes a coinsurance rate $k_i(\cdot)$, which is a "reaction function" of social policy instruments, *i.e.* the payroll tax rate t and the social insurance rate, α . Thus $P_i=(1-\alpha)k_i(\alpha,t)$ may be considered as the individual i "unit price" for health expenditure E_i or his unit cost of the service out-of-pocket consumption.

Thus with

$$(9) \quad \frac{\partial P_i}{\partial \alpha} = -k_i(\alpha, t) + (1 - \alpha) \frac{\partial k_i}{\partial \alpha}$$

we measure the change of this price due to social insurance, which is negative for a sufficiently inelastic reaction of private insurance to its social counterpart (crowding-out effect).

As far as the optimal social insurance rate α^* is concerned, this is given by equalizing the sum of two marginal social benefits (the l.h.s) with a marginal social cost (the r.h.s) as follows (Petretto, 1999, formula (19) and (20))¹⁵:

$$(10) \quad SRSG+SRG=SMHC$$

¹⁵ The terms with "bar" are expected values and the terms not identified with i are average values in the population.

where

$$\begin{aligned}
 SRS&G = \sum_i Cov_p(\bar{b}_i, -\frac{\partial P_i}{\partial \alpha_i} E_i) \\
 SR&G = n[\bar{b}\bar{\xi}_E - 1]\bar{E} \\
 SM&HC = (1 - \alpha) \sum_i \bar{b}_i(1 - k_i) \frac{\partial \bar{E}_i}{\partial \alpha} + \alpha n \frac{\partial \bar{E}}{\partial \alpha}
 \end{aligned}$$

On the l.h.s., *SRS*G represents the social-risk sharing gain from a marginal increase of social insurance coverage, α , and is given by a measure of the benefit for the whole society from “buying” a further insurance policy with uniform reimbursement rate for all members. α is higher, the higher is the sum of individuals’ gains of risk sharing, referring to out-of-pocket health spending. The gain from risk-pooling is somewhat socialised in two ways: first, by following a utilitarian welfare function rule, *i.e.* summing up the individual covariances between marginal utility of income and out-of-pocket health expenditure; second, by considering, as in the classical optimal taxation framework, each individual’s “social” (*i.e.* measured in terms of government revenue) expected marginal utility of income, \bar{b}_i . The term *SR*G represents the social redistribution gain stressing the role of α^* for vertical equity aims. The higher is the *distributional characteristics* of health spending, $\bar{\xi}_E = Cov(\bar{b}_i, \bar{E}_i)$, the higher is the optimal rate α^* , as in the standard model of many-consumers optimal commodity taxation, with a linear income tax. Indeed social insurance coverage works like a subsidy applied to health expenditure at a uniform rate. On the r.h.s., *SM*HC represents a social moral hazard cost as a weighted sum of two individual’s moral hazard effects: the first comes from private budget constraints and it is linked to the sum of individual effects in terms of expected health expenditure increase; the second is determined by government budget constraint and is related to the effect on total public expenditure due to social health insurance.

Boadway *et al.* (2006) extend Petretto (1999) by adding adverse selection on the distribution of risks to *ex-post* moral hazard. In their model, private insurers are well informed about households productivity so can offer separate insurance policies to persons of different income classes. But, since risk class cannot be observed, insurance policies must be designated to separate high risk from low-risk persons, as in *Rothschild-Stiglitz model*. In other words, within each productivity class, there is an insurance market equilibrium separating the two risk subclasses

where insurers offer policies (premium and coverage) for L (low-risk) and H (high-risk) households of productivity class i , choosing their most preferred policy. Insurance companies are then better informed than government, which can observe neither risk class nor productivity. This gives advantage to private insurers thereby making the case for social insurance as strong as possible.

The most plausible circumstances contemplated by the model actually design situations where there is room for private insurance (the second pillar), given the presence of partial social insurance (the first pillar). Neither high-risk persons or low-risk persons have full insurance given *ex-post* moral hazard and then, as in Petretto (1999), there is an out-of-pocket component of health expenditure (the third pillar). As far as the choice of the optimal social coverage rate, Boadway *et al.* find an expression extending in many respects condition (15) by Petretto. Given the adverse selection context, it is analogous to the standard expression for linear income tax rate, combining equity and efficiency concerns, except that it includes additional terms reflecting the various margins of distortion of the economy. Boadway *et al.* (2006, formula (12)-(17)) are in fact able to explicit α^* as follows:

$$(11) \quad \alpha^* = \frac{[Eq] + [LD] + [AS]}{[MH]}$$

where

$[Eq]$ = *equity term*, i.e. the covariance over all types between the marginal expected social valuation of income and health-care spending,

$[LD]$ = *efficiency term*, i.e. the indirect effect of social insurance on labour supply distortion,

$[AS]$ = *adverse selection term*, for fulfilling self-selection constraints, and

$[MH]$ = *moral hazard term*, linked to the compensated total change in the demand for health care spending w.r.t. α .

As in Petretto (1999), at the optimum of a three stages game, for plausible hypotheses, a interior solution arises, such that $0 < \alpha^* < 1$. This result implies that, for designing a mixed insurance system, a uniform social insurance rate is justified on the basis of equity as well efficiency considerations. Therefore, the first pillar, although essential and probably with a significant size, in general, does not guarantee a full social coverage: some services must be rationed, looking at their “appropriateness”, and further expenditures must be financed with supplementary insurance and co-payment (out-of-pocket).

Table 6 summarises the logic of mixed insurance system, by representing the sequential decision-making structure implying three pillars, to be solved by backward induction.

TABLE 6

THE THREE-STAGE, THREE-PILLAR, GAME WITH ASYMMETRIC INFORMATION

STAGE 1	Optimal social insurance rate, $0 < \alpha^* < 1$ and optimal income tax rate t^* , with a trade-off equity-efficiency
STAGE 2	Supplementary insurance, given α and t , with optimal coinsurance rate
STAGE 3	Out-of-pocket expenditure (co-payment), given the two insurances

4.4 *Toward a General Theory of Optimal Taxation with Endogenous Private Insurance*

The logic of multi-pillar funding may be extended to the more general issue of looking for the optimal amount of redistribution taxation and social insurance. Chetty and Saez (2010) has recently developed a quite general model, with moral hazard and adverse selection, where social and private insurance, in terms of linear contracts based on individual income, interact each other. They extend Petretto (1999), by considering several combinations of optimal tax rates, and Boadawy *et al.* (2006) too, by admitting private and public insurers to have the same informational set, being in other words unable to observe both risks and abilities.

Chetty and Saez contribution is mainly devoted to correct the standard optimal income tax formula with the presence of a private insurance. In this respect, their various tax formula extend the one provided by Petretto (1999, (21) and (22)), applying to many forms of social insurance, not only health insurance, and given the adverse selection in a continuum of individuals. However, just for giving a sketch, let us ignore the latter and denote with t_s the “tax rate” chosen by the government and with t_p the “tax rate” in the private insurance contract. The two rates give the extent respectively of social and private insurance.

Let z denote the earning of type- n individual and $Z = \int zf(n)dn$ the average value in the population. The individuals are risk-averse so would like to insure themselves against the risk of having low-income realizations (*e.g.* due to a treatment expenditure) and $w = (1 - t_p)z + t_p Z$ is the net-of-private insurance income. Government taxation applies to the net incomes w^{16} , thus $c = (1 - t_s)w + t_s W$, $W = \int wf(n)dn$,

¹⁶ Here we are presenting the case where private insurance is not optimally chosen. For the extension to a multi-stage game where also consumers optimize the choice of insurance see proposition (2) of CHATTY R. and SAEZ E. paper (2010).

is the final, after the two insurances, disposable income. Let τ denote the total tax rate, defined as $(1-\tau)=(1-t_p)(1-t_s)$, and now $c=(1-\tau)z+\tau Z$. If the private insurer and the government cooperated to set τ , to maximise social welfare, the resulting contract would be identical to that described in the standard optimal taxation setting. However, as pointed out in Petretto (1999) and Boadway *et al.* (2006), private insurers take the government contract t_s as given when they chose t_p , and $t_p(t_s)$ is the corresponding reaction function, analogous to $k_i(.)$ previously used.

Chatty and Saez show that, taking the function $t_s \rightarrow t_p(t_s)$ as given, the optimal government tax formula is as follows (u' is marginal utility of income and U' is the average in the population):

$$(12) \quad \frac{t_s}{1-t_s} = -t_p + \frac{1-r}{\epsilon_{W,1-t_s}} \frac{-Cov(w, u')}{WU'}$$

$$r = - \frac{d \log (1-t_p)}{d \log (1-t_s)}$$

denotes the empirically observed rate at which public insurance crowds out private insurance. If $r=0$, there is not crowd-out. If $r=1$, there is perfect crowd.

$$\epsilon_{\bar{W},1-t_s} = \frac{d \log W}{d \log (1-t_s)}$$

denotes the elasticity of average, post-insurance, income w.r.t. the net-of-tax rate $(1-t_s)$, taking into account the endogenous response of private insurance t_p to a change in t_s .

According to (12), private insurance affects the formula for optimal tax rate in two ways. First, the term $-t_p$ on the r.h.s. reflects the reduction in the optimal level of government taxation given the presence of private insurance. Second, the standard inverse elasticity term

$$\frac{1}{\epsilon_{W,1-t_s}} \frac{-Cov(w, u')}{WU'}$$

is multiplied by $(1-r) < 1$, making the optimal government tax rate, and then social insurance, once more smaller. Such an integration designs a multi-pillar structure for many social protection sectors, and then also for health care.

5. - Some Empirical Comparisons and Concluding Remarks

In this work we surveyed the HCSs, trying to follow an approach coherent with economic theory. A specific aim was to disentangle the boundary between public and private institutional involvement, even if we find out that this boundary is not always meaningful for enlightening the differences among systems. Let us give some synthetic sketches of the arguments carried on.

First of all, on examining the features of public intervention in health care, it is necessary to distinguish the motives for public provision from those for public production. The two notions do not coincide and the two forms of public intervention should not be examined jointly. Further, when public production is preferred, this can be integrated or separated from the purchasing of health services made by a public entity or an insurer, on the behalf of the citizens enrolled in the NHS. The choice of this configuration may be usefully analysed in terms of the I-O and property rights theory of cost-benefit of vertical integration of a firm. However, we do not achieve a theoretical conclusive answer to this important institutional question, that must be therefore verified according to the prevailing socio-economic context.

The health-care organization systems are a lot, a sort of continuum, but they are independent on the funding systems. For instance the NHS of UK, is almost fully public in funding (Table 7). However, it is a managed competition system with the separation of public Trust hospitals from the District health authorities. Moreover, the freedom by patients, and general practitioners, in choosing the providers and the producers of services is allowed.

HCSs with strong private insurances, like in US, are quite costly and lack adequate coverage of sickness risks. Consequently, they are now everywhere integrated with social insurance, as the recent Obama's reform proposes. However, also a fully centralised command & control system finds it hard to reach the planned equity aims, given the lack of information by the planner and the conflicting individual incentives of the many involved agents. These systems have been as well generally dropped out by the reforms toward mixed systems carried on in most European countries in the last three decades.

Among the mixed systems, the managed competition in health care is a widely applied regime in Europe. I might reach appreciable results in terms of efficiency and equity. However, it needs a fairly territorial distribution of good-quality providers and it calls for complicated and costly mechanisms of contracts regulation and *ex-ante* and *ex-post* controls.

TABLE 7

HEALTH EXPENDITURE BY FINANCING AGENT IN 5 EU COUNTRIES

	Sweden	France	Germany	Italy	UK
General Government	81.5	78	76.9	77.9	84.1
<i>a.</i> Territorial government	81.5	5.5	8.7	77.7	84.1
<i>b.</i> Social security funds		72.5	68.2	0.2	
Private Sector	18.5	22	23.1	22.1	15.9
<i>a.</i> Private insurance	0.2	13.3	9.3	1.0	1.1
<i>b.</i> Household out-of-pocket	16.7	72	13.1	19.7	10.5
<i>c.</i> Non-profit institutions					
serving households	0.2	0.1	0.4	1.4	3.7
<i>d.</i> Corporations	1.4	1.4	0.3		
	100	100	100	100	100

Source: OCDE, *Health data*, year 2009.

As said, the issue separation *versus* integration in production has no relevance on the boundaries between private and public financing means. If we pick up two countries where there are two different production organizations, an integrated and bureaucratic one as in France, and a separated one as in The Netherlands, we see, somewhat surprising, that the structure linking financing agencies to providers is not so different (Table 8). In both countries, general government finances the same level of hospitals expenditure, about one third, although hospitals in France are internal to the Public Administration (a vertical integrated system), while Dutch hospitals, even if public ones, are outside the P.A. (a classical universalistic marketed HCS). A greater difference there is in residential care facilities industry, as in the separated system the public expenditure is higher in percentage because the externalized agencies are more numerous.

TABLE 8

PROVIDERS PER FINANCING INSTITUTIONS: EXPENDITURE COMPOSITION (%)
PROVIDERS

	<i>Hospitals</i>	<i>Residential care facilities</i>	<i>Ambulatory health care</i>	<i>Retail sale of medical goods</i>
France <i>integrated model for production</i>				
<i>a. General Government</i>	32.7	7	19.6	13.5
<i>b. Private sector</i>	2.8	-	7.3	6.2
Netherlands <i>separated model for production</i>				
<i>a. General Government</i>	33.1	22.7	13.4	9.6
<i>b. Private sector</i>	0.6	0.1	9.1	3.8

Source: OCDE, *Health data*, year 2009.

All the health-care funding systems, the NHSs too, tend to follow the logic of insuring individual risks, and fulfilling, at the same time, an aggregate financial equilibrium. In this respect, the theory shows that a mixed system with of a social insurance, for an essential package of services, and a supplementary and integrative private insurance might represent an efficient configuration.

In many European HCSs the second pillar in health-care is financed by integrative insurance funds, even public ones (Table 7). In some countries, like Italy, there are somewhat polar systems where the wideness of the first pillar, given the increasing comprehensiveness of the ELC packages, has limited the expansion of the second pillar, actually preventing the achievement of a meaningful role by integrative insurance funds. However, in these cases, the large extent of essential package in the first pillar has not limited the size of the third pillar which has been instead increasing in the time. In Table 7 we see that in Italy, where the integrative insurance is still more limited than in France or Germany (1.0 respect to 13.3 and 9.3) the household out-of-pocket expenditure is much higher (19.7 against respectively 7.2 and 13.1).

A system with an increasing range of treatments and services included in the essential package and strictly funded with central government taxation, although justified by equity concerns, could become, in the long run, unsustainable for the financial equilibria and for the level of aggregate tax burden. In perspective, the sustainability problems could be limited by widening the second pillar, together with a system of co-payments, even income-related. Despite several institutional problems, and often large administrative and transaction costs (Geichert 2010), this mixed system could assure the feature of universality with a high degree of public provision of health-care services.

APPENDIX

A.1. Essential Levels of Health Care and Co-Payment

Let

$$(A.1) \quad TWP_i \equiv \phi_i(m_i, Y_i, \bar{v}^i)$$

be the function of the total willingness to pay by i for the treatment m_i , given his income Y_i and utility level $U^i(\cdot) = \bar{v}^i$ (the desired well-being). Consequently the marginal willingness to pay is as follows:

$$(A.2) \quad MWP_i = - \frac{\partial \phi_i(m_i, Y_i, \bar{v}^i)}{\partial m_i}$$

If the government is committed to guarantee to every citizen the access to a ELC of m , it is reasonable to think that the government gives to the treatment a higher marginal evaluation than that one recognised by the individual himself:

$$(A.3) \quad MWP_i^G = MWP_i + \omega_i(m_i)$$

Therefore, according to government evaluation, the amount of *numeraire* the individual *should* be willing to spend for buying the treatment m_i and reaching the desired utility level \bar{v}^i , is:

$$(A.4) \quad TWP_i^G \equiv \phi_i^G(m_i, Y_i, \bar{v}^i) = \phi_i(m_i, Y_i, \bar{v}^i) - \int_{m_0^G}^m \omega_i(\chi) d\chi$$

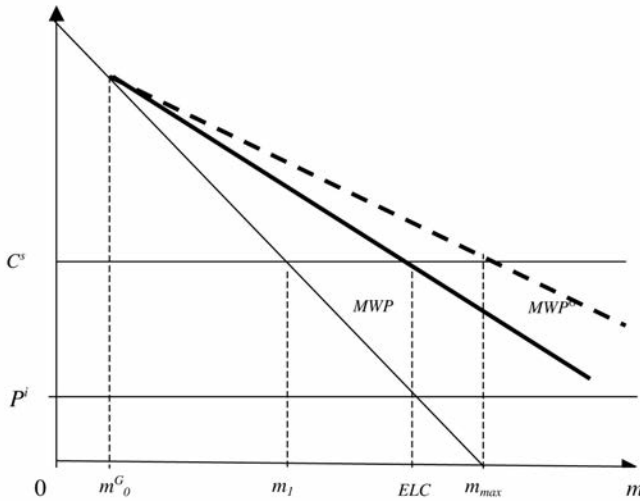
m_0^G is the level of the service behind which the public marginal evaluation is deviating from that one of the individual as such. It can be shown that to (A.4) it corresponds the level of well-being

$$U^G(m_i, Y_i) = U^i[m_i, Y_i + \int_{m_0^G}^m \omega_i(\chi) d\chi]$$

achievable with a lump sum subsidy to i . Alternatively, the consumer price for the service i has to face must be lower than the producer price, in order to push his consumption over the level he otherwise would choose, until the specified essential level.

GRAPH 1A

SOCIAL SERVICE AND ACCESS TO THE ESSENTIAL LEVEL OF HEALTH CARE



The expression $-\int_{m_0^G}^m \omega_i(\chi) d\chi$ represents the area of Graph 1a between the two curves of marginal evaluation, by the government and by the individual, thus the area of public responsibility for the service provision. In the horizontal axis, $m_1 > m_0^G$ is the market solution where the private MWP is equal to the marginal (standardised) cost C . ELC is the consumption socially guaranteed, given by crossing the bold curve MWP^G with C^s line, which corresponds, at the same time, to the level at which private MWP curve is crossing the co-payment line $P^i < C^s$. Notice that the co-payment is zero when the MWP^G is the dotted curve, $ELC = m_{max}$ and the whole expenditure is borne by the government.

A.2 Tariffs for Health Services and Incentives in Separated and Semi-Integrated Models

Let a tariff for a hospital treatment be given by this linear function:

$$(A.5) \quad T = a + b C(q, e, \mu); \quad 0 \leq b \leq 1$$

where $C(\cdot)$ is the production cost of the volume of treatments q ; $C_q \equiv \frac{\partial C}{\partial q} > 0$ is the marginal cost. $e \in [e_{\min}, e_{\max}]$ is the cost-containment effort carried on by the hospital manager and $C_e \equiv \frac{\partial C}{\partial e} < 0$ its marginal effect on the level of cost, with $\frac{\partial |C_e|}{\partial e} < 0$, for regulatory concerns. $\mu \in [\mu_{\min}, \mu_{\max}]$ represents the quality of the treatment and $C_\mu \equiv \frac{\partial C}{\partial \mu} > 0$ its marginal cost, positive as clearly higher quality requires more resources, with now $\frac{\partial C_\mu}{\partial \mu} > 0$. Let us suppose that a HD or a ASL plan establishes a given volume of output to be provided, so $q = q^*$ is inserted in the contract signed with the hospital (TH or AO). Both e and m are instead not observable by the HD and then are not contractible, while production costs of the TH are only *ex-post* observable, *i.e.* when realised. Let then define with $\psi(e, \mu)$ the objective function of the manager of the hospital, where $\psi_e \equiv \frac{\partial \psi}{\partial e} < 0$ represents the marginal cost of cost-containment effort (disutility), with $\frac{\partial |\psi_e|}{\partial e} > 0$, and $\psi_\mu \equiv \frac{\partial \psi}{\partial \mu} > 0$ the marginal benefit of quality in terms of prestige and reputation of the manager, with $\frac{\partial \psi_\mu}{\partial \mu} < 0$.

In a separated or semi-integrated system, a *residual claimant* provider will tend to choose e and μ to maximize the following difference function:

$$(A.6) \quad R = T - C(q^*, e, \mu) - \psi(e, \mu) = a - (1 - b) C(q^*, e, \mu) + \psi(e, \mu)$$

In the case of a fee-for-service tariff (*cost-plus contract*), as in both reimbursement and integrated models, we have $b=1$. Therefore, it is

$$T=a+C(q^*,e,\mu) \text{ and } R=a+\psi(e,\mu)$$

i.e. the residual is independent on production costs, so there is no incentive to contain them. Indeed, theoretically, the optimal level of the hospital manager effort is a corner solution such that $e^*|_{b=1}=e_{\min}$. However, being fully insured, the manager does not exploit the informative monopoly and might give up cream-skimming procedures and provide high quality treatments. Therefore, it might well happen a further corner solution such that $\mu^*|_{b=1}=\mu_{\max}$.

In the case of a fixed per treatment tariff (*fixed-price contract*), *i.e.* $b=0$, as in pure QM contractual model, where there is a prospective DRG payment, it is

$$T=a \text{ and } R=a-C(q^*,e,\mu)+\psi(e,\mu).$$

Since the residual is now decreasing with the treatment production cost, the hospital has an effective incentive to contain it. The optimal level of the effort is given by the condition $|C_e|=|\psi_e|$ implying that $e^*|_{b=0}>e^*|_{b=1}=e_{\min}$. However the hospital, bearing the full firm risk, may be induced to cream-skim high illness risks and to restrain the quality level of the service; indeed, in this case, the optimal quality level is given by condition $C_\mu=\psi_\mu$, and $\mu^*|_{b=0}<\mu^*|_{b=1}=\mu_{\max}$. In more concrete terms, a fixed-price contract, without a reliable *ex-ante* selection of providers and effective controls on standards, may entail a worsening of treatment quality, with for instance a higher rate of patients mortality.

This simple model seems to suggest that, in order to reach an adequate standard of quality but satisfying a fixed budget constraint, a mixed systems of *cost and risk-sharing* payment, with $0<b<1$, might be preferable. The optimal levels e^0 and μ^0 are now given by the conditions $(1-b)|C_e|=|\psi_e|$ and $(1-b)C_\mu=\psi_\mu$, according to which it obtains

$$e^*|_{b=0} > e^0 > e^*|_{b=1} = e_{\min} \text{ and } \mu^*|_{b=0} < \mu^0 < \mu^*|_{b=1} = \mu_{\max}$$

Actually, a incentive risk-sharing criterion may be pursued by organising a system of budgeted plans based on a fixed volume of treatments, some proxy indexes of quality, and standardised and verifiable costs, with *ex-post* revenues abatements in presence of non-fulfilments of the objectives.

BIBLIOGRAPHY

- ANDERSON A. - GROSSMAN M., «Health and the Household», *Review of Economics of the Household*, no. 7, 2009, pages 219-226.
- BAR-ISAC H. - TADELIS S. «Seller Reputation», *Foundation and Trends in Microeconomics*, no. 4, 2008, pages 273-351.
- BASU K. - LOPEZ-CALVA L.-F., «Functioning and Capabilities», in ARROW K.J. - SEN A.K. - SUZUMURA K. (eds.), *Handbook of Social Choice and Welfare*, vol. 2, North Holland, Amsterdam, 2010.
- BECKER G.S., «Health as Human Capital: Synthesis and Extensions», *Oxford Economic Papers*, no. 59, 2007, pages 379-410.
- BELLEFLAMME P. - PEITZ M., *Industrial Organization, Markets and Strategies*, Cambridge University Press, New York, 2010.
- BESLEY T., *Principled Agents, the Political Economy of Good Government*, Oxford University Press, 2007.
- BLOMQUIST A. - JOHANSSON P.-O., «Economic Efficiency and Mixed Public/Private Insurance», *Journal of Public Economics*, no. 66, 1997, pages 505-516.
- BOADWAY R. - LEITE-MONTEIRO M. - MARCHAND M. - PESTIEAU P., «Social Insurance and Redistribution with Moral Hazard and Adverse Selection», *Scandinavian Journal of Economics*, no. 108(2), 2006, pages 279-298.
- BOADWAY R. - MARCHAND M. - SATO M., «An Optimal Contract Approach to Hospital Financing», *Journal of Health Economics*, no. 23, 2004, pages 85-110.
- BOLTON P. - DEWATRIPONT M., *Contract Theory*, MIT Press, Cambridge, Mass., 2005.
- BREYER F. - HAUFLE A., «Health Care Reform: Separating Insurance from Income Redistribution», *International Tax and Public Finance*, no. 7, 2000, pages 445-461.
- CHALKLEY M. - KHALIL F., «Third Party Purchasing of Health: Patient Choice and Agency», *Journal of Health Economics*, no. 24, 2005, pages 1132-1153.
- CHALKLEY M. - MALCOMSON J., «Government Purchasing of Health Services», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- CHATTY R. - SAEZ E., «Optimal Taxation and Social Insurance with Endogenous Private Insurance», *American Economic Journal: Economic Policy*, no. 2, 2010, pages 85-114.
- CREMER H. - PESTIEAU P., «Redistributive Taxation and Social Insurance», *International Tax and Public Finance*, no. 3, 1996, pages 281-295.
- CULYER A., «The Normative Economics of Health Care Finance and Provision», *Oxford Review of Economic Policy*, no. 5, 1989, pages 34-58.
- CUTLER D.M., «Equality, Efficiency, and Market Fundamentals: The Dynamics of International Medical-Care Reform», *Journal of Economic Literature*, vol. XL, 2002, pages 881-906.

- CUTLER D.M. - ZECKAUSER R.J., «The Anatomy of Health Insurance», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- DAHLBY, B., «Adverse Selection and Pareto Improvements Through Compulsory Insurance», *Public Choice*, no. 37, 1981, pages 548-558.
- DANZON P.M., «Editorial on Welfare Effects of Supplementary Insurance. A Comment», *Journal of Health Economics*, no. 21, 2002, pages 923-926.
- DIXIT A., «Incentives and Organizations in the Public Sector: An Interpretative Review», *Journal of Human Resources*, vol. 37, 2002, pages 696-727.
- DRANOVE D. - SATTERTHWAITTE M.A., «The Industrial Organization of Health Care Markets», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- FLEURBAEY M., «Beyond the GDP: The Quest for a Measure of Social Welfare», *Journal of Economic Literature*, vol. XLVII, no. 4, 2009, pages 1029-1076.
- GECHERT S., «Supplementary Private Health Insurance in Selected Countries: Lessons for EU Governments», *CESifo Economic Studies*, no. 56(3), 2010, pages 444-464.
- GIBBONS R., «Four Formal(izable) Theories of Firm?», *Journal of Economic Behavior & Organization*, no. 58, 2005, pages 200-245.
- GRABOWSKY D.C. - HUSKAMP H.A. - STEVENSON D.G. - KEATING N.L., «Ownership Status and Home Health Care Performance», *Journal of Aging & Social Policy*, no. 21(2), (2009)pages 130-143.
- GRAVELLE H. - SICILIANI L., «Optimal Quality, Waits and Charges in Health Insurance», *Journal of Health Economics*, no. 27, 2008, pages 663-674.
- GROSSMAN D., «The Human Capital Model», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- HANSEN B.O. - KEIDING H., «Alternative Health Insurances Schemes: A Welfare Comparison», *Journal of Health Economics*, no. 21, 2002, pages 739-756.
- HART O., «Incomplete Contracts and Public Ownership: Remarks and an Application to Public-Private Partnership», *Economic Journal*, no. 113, 2003, pages C69-C76.
- HOEL M., «What Should (Public) Insurance Cover?», *Journal of Health Economics*, no. 26, 2007, pages 251-262.
- HOPPE E.I. - SCHMITZ P.W., «Public versus Private Ownership: Quantity Contracts and the Allocation of Investment Tasks», *Journal of Public Economics*, no. 94, 2010, pages 258-268.
- HURLEY J., «An Overview of the Normative Economics of the Health Sector», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- JONES P.R. - CULLIS J.G., «Decision Making in Quasi-Markets: A Pedagogic Analysis», *Journal of Health Economics*, no. 15, 1996, pages 187-208.

- KIFMANN M., «Community Rating in Health Insurance and Different Benefit Packages», *Journal of Health Economics*, 2002, pages 719-737.
- LEACH J., «Ex-Post Welfare under Alternative Health Care Systems», Department of Economics, Mc Master University, *Working Paper* 2007-05, forthcoming in *Journal of Public Economic Theory*, 2010.
- LE GRAND J., *Motivation, Agency and Policy*, Oxford University Press, 2003.
- .-, *The Other Invisible Hand, Delivering Public Services Through Choice and Competition*, Princeton University Press, 2007.
- MA A., «Health Care Payment Systems: Cost and Quality Incentives», *Journal of Economics and Management Strategy*, no. 3(1), 1994, pages 93-112.
- MC GUIRE T., «Physician Agency», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- MAYNARD A., «Can Competition Enhance Efficiency in Health Care? Lessons from the Reform of UK National Health Service», *Social Science and Medicine*, vol. 39, no.10, 1994.
- MOSSIALOS E. - THOMSON S., *Voluntary Health Insurance in the European Union*, Brussels, World Health Organization, European Observatory on Health Systems and Policies, 2004.
- NYMAN J., «The Economics of Moral Hazard Revisited», *Journal of Health Economics*, no. 18, 1999, pages 811-824.
- PAULY M., «Insurance Reimbursement», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- PETRETTO A., «Optimal Social Health Insurance with Supplementary Private Insurance», *Journal of Health Economics*, no. 18, 1999, pages 727-745.
- PROPPER C. - BURGESS S. - GOSSAGE D., «Competition and Quality: Evidence from the NHS Internal Market 1991-1999», *The Economic Journal*, no. 118, 2008, pages 138-170.
- ROCHET J.-C., «Incentive, Redistribution and Social Insurance», Geneva, *Papers of Risk and Insurance*, no. 16, 1991, pages 143-165.
- RODDEN J. - ESKELAND G. - LITVAK J. (eds.), *Decentralization and the Challenge of Hard Budget Constraints*, The MIT Press, Cambridge, 2001.
- SADKA E., «Public-Private Partnerships: A Public Economics Perspective», *CESifo Economic Studies*, no. 53(3), 2007, pages 466-490.
- SCHROYEN F., «An Alternative Way to Model Merit Goods Arguments», *Journal of Public Economics*, no. 89, 2005, pages 957-966.
- .-, «Operational Expressions for Marginal Cost of Indirect Taxation when Merit Arguments Matter», *International Tax and Public Finance*, no. 17, 2010, pages 43-51.
- SELDEN T.M., «More on Economic Efficiency of Mixed Public/Private Insurance», *Journal of Public Economics*, no. 66, 1997, pages 517-523.

- SMITH P. - GODDARD M., «The English National Health Service: An Economic Health Check», OCDE, Economics Department, *Working Paper*, no. 717, ECO/WK (2009) 58, 2009.
- SULLIVAN S.D., «US Health Care Reform: A Work in Progress», *European Journal of Health Economics*, no. 11, 2010, pages 117-118.
- WAGSTAFF A. - VAN DOORSLAER E., «Equity in Health Care Finance and Delivery», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- WEISBROD B.A., «Why Private Firms, Governmental Agencies and Non-Profit Organizations Behave Both Like and Differently. Application to the Hospital Industry», Department of economics, *Working Paper*, Northwestern University, Evanston, Ill., 2006.
- WILDASIN D.E., «The Institutions of Federalism: Toward an Analytical Framework», *National Tax Journal*, vol. LVII (2), 2004, pages 247-272.
- WILLIAMS A. - COOKSON R., «Equity in Health», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam, 2000.
- WILLIAMSON O.E., «Economics of Governance», *American Economic Review*, no. 95(2), 2005, pages 1-18.
- WILSON C., «A Model of Insurance with Incomplete Information», *Journal of Economic Theory*, no. 12, 1977, pages 167-207.
- ZANETTE A. - RICATTI M., «Health Care System in the Industrialised Countries and the Role of Private Insurance», *European Papers on the New Welfare*, no. 6, 2006, pages 60-80.
- ZWEIFEL P. - MANNING W., «Moral Hazard and Consumer Incentives in Health Care», in CULYER A.J. - NEWHOUSE J.P. (eds.), *Handbook of Health Economics*, North Holland, Amsterdam. (2000),
- ZWEIFEL P. - BREYER F. - KIFMANN M., *Health Economics*, Springer, Heidelberg, 2009.

Alcune implicazioni fiscali di scenari alternativi alla proibizione della cannabis

Marco Rossi*

Università La Sapienza, Roma

We evaluate some budgetary implications of having cannabis regulated as tobacco was in Italy in 2007-2008. Depending on which scenario would have followed this shift of regime, the fiscal revenues from sale and income tax would have been between 7,5 and 13 billions of euros, moreover 830 million of euros would have been saved from jailing expenses. We also discuss about other implications such as health and productivity costs and the prohibition-induced crime.

[JEL Classification: 5E; 6D].

Keywords: prohibition; pigouvian tax; fiscal budget.

* <rotorossi@yahoo.com>, Facoltà di Sociologia, Scienze Politiche e Comunicazione, Dipartimento di Scienze Sociali ed Economiche.

L'autore desidera ringraziare i *referee* anonimi per gli utili suggerimenti.

1. - Introduzione

Sulla base di un approccio pragmatico, il dibattito sul controllo degli stupefacenti sviluppatosi negli ultimi anni non nasconde che il proibizionismo ha causato degli effetti indesiderati, tra cui il più importante è stato lo sviluppo di un lucrativo mercato nero e della criminalità ad esso connessa¹. Recentemente sono stati proposti dei sistemi di controllo alternativi alla semplice proibizione: il rapporto presentato dall'incaricato speciale A. Grover all'Assemblea Generale dell'ONU² propone infatti di decriminalizzare il consumo di droga e di utilizzare invece gli strumenti della Convenzione Internazionale sul controllo del Tabacco anche per il controllo degli stupefacenti. In linea con questa proposta, cerchiamo di valutare l'importo di alcune implicazioni fiscali degli scenari che si sarebbero potuti instaurare se nel biennio 2007-2008 la normativa che disciplinava il mercato ed il consumo delle sigarette in Italia fosse stata applicata anche al mercato della cannabis³. In particolare ipotizziamo che:

- a) la produzione, manifattura, distribuzione all'ingrosso ed al dettaglio della cannabis fosse stata svolta tramite un sistema di operatori autorizzati dai Monopoli di Stato (Leggi n. 331 del 1990, n. 92 del 2001 e n. 248 del 2005; D.Lgs. n. 261 del 1990 e n. 203 del 2005);
- b) il prezzo di vendita al dettaglio della cannabis, fissato dall'autorità, contenesse una gravosa *sin tax* pari al 75,5% (Leggi n. 825 del 1965, n. 76 del 1985 e successive modificazioni; D.lgs. n. 24 del 2004 e n. 168 del 2004 e Decreto Direttoriale 15 ottobre 2004, art. 1);
- c) la vendita ed il consumo di cannabis fossero stati sottoposti a limitazioni ed ogni attività promozionale della cannabis fosse stata vietata (D.lgs. n. 165 del 1962, n. 184 del 2003, n. 300 del 2004; Leggi n. 3 del 2003, n. 574 del 1985 e n. 52 del 1983).

Si avvisa subito il lettore che la scelta sulla legalizzazione della cannabis non può limitarsi all'analisi delle sue eventuali implicazioni economiche, di cui quelle fiscali sono solo una componente. La legalizzazione della cannabis può avere altre implicazioni per la collettività, a partire da quelle di salute e sicurezza pubblica

¹ UNODOC, *World Drug Report 2009*, Executive Summary, page 18.

² "Report of the Special Rapporteur on the Right of everyone to the enjoyment of the highest attainable standard of physical and mental health", par. D, page 23.

³ KILMER B. *et AL.* (2010) stimano i possibili effetti della legalizzazione della cannabis sul bilancio statale della California.

per finire con quelle diplomatiche⁴. La scelta dell'operatore pubblico sulla legalizzazione della cannabis non può prescindere da queste implicazioni extra-economiche. Il peso che l'operatore pubblico attribuisce a ciascuna di queste possibili implicazioni dipende dal sistema di valori tramite il quale la collettività percepisce i costi ed i benefici delle scelte pubbliche. Ad esempio, l'operatore pubblico delle società occidentali ritiene che i danni causati dal consumo di alcolici siano inferiori a quelli che deriverebbero dalla sua proibizione. Ed infatti l'imposizione del proibizionismo degli alcolici negli Usa ("Dry Act" del 1920) non colse l'obiettivo di ridurre significativamente il consumo, ma indusse degli effetti indesiderati così gravi (tra in cui in particolare lo sviluppo di una criminalità attiva nel mercato nero) da indurre l'operatore pubblico a rinnegare la sua scelta proibizionista ("Repeal Act" del 1933). Viceversa, in alcuni paesi islamici il divieto del consumo di alcolici è rispettato e la sua imposizione non induce rilevanti effetti indesiderati. Evidentemente per queste collettività la scelta proibizionista è coerente con il sistema di valori tramite il quale esse percepiscono i costi ed i benefici del consumo di alcolici. Pertanto, il tema della legalizzazione della cannabis andrebbe affrontato includendo nella funzione obiettivo dell'operatore pubblico tutte le eventuali implicazioni, anche extra-economiche, di questa scelta, ed attribuendo a ciascuna di queste un peso coerente con il sistema di valori prevalente nella società italiana. Il complesso tema della compatibilità della legalizzazione della cannabis con il sistema di valori diffuso nella società italiana esula dagli obiettivi di questo lavoro⁵, il quale si limita a valutare alcune implicazioni fiscali di due scenari alternativi alla proibizione, cercando, in tal modo, di fornire un parziale contributo all'adozione di scelte pubbliche "evidence-based".

In generale, le implicazioni fiscali connesse all'uso della cannabis derivano sia dai costi dovuti all'uso della cannabis in sé, sia nei costi sostenuti dalla società in risposta a questo consumo⁶. I primi consistono nei costi sanitari per i trattamenti e

⁴ L'Italia ha ratificato la convenzione ONU del 1988, che la invita ad attribuire natura di reato, ai sensi del proprio ordinamento giuridico interno, al possesso, l'acquisto e la coltivazione di stupefacenti illegali destinati al consumo personale.

⁵ Ci limitiamo ad osservare che la proibizione del consumo di cannabis non è rispettata da una significativa frazione della popolazione italiana. Se il rispetto della norma è un segnale della sua accettazione, questa vasta diffusione del consumo illegale di cannabis induce a dubitare sulla compatibilità dell'attuale proibizionismo con il sistema di valori diffuso nella società italiana.

⁶ Per una rassegna ed una discussione dei costi sociali connessi al consumo di droga: PACULA R.L. *et AL.* (2009).

le malattie dovute all'uso della cannabis, nel costo della perdita di produttività dovuta all'uso di cannabis, ed infine nelle esternalità sopportate dalle società a causa di questo consumo (incidenti di guida, crimine indotto dal traffico illegale di cannabis, ecc.). A questi costi, dovuti al consumo di cannabis in sé, si aggiungono le spese per l'applicazione della normativa proibizionista (polizia, magistratura, carceri, ecc.), il costo della perdita di produttività dovuta alle sanzioni subite a causa dell'uso di cannabis, ed i costi delle politiche di prevenzione. Questo studio non offre una esaustiva valutazione dell'opportunità sociale del cambio di regime ipotizzato, e nemmeno di tutte le sue possibili implicazioni fiscali. Il nostro studio si limita ad una parziale valutazione del costo opportunità delle entrate fiscali non riscosse a causa del proibizionismo e delle spese carcerarie sostenute per la detenzione dei soggetti condannati per violazione della normativa vigente. Circa le altre implicazioni fiscali di cui sopra, questo studio si limita invece ad una loro introduttiva discussione, ma non offre alcuna stima del loro eventuale importo. L'omissione della stima di queste altre poste contabili, in particolare degli effetti della legalizzazione della cannabis sulla criminalità e sulla produttività, fa sì che i nostri risultati siano solo una parziale valutazione delle possibili implicazioni fiscali che avrebbe potuto avere l'equiparazione normativa della cannabis alle sigarette in Italia.

Per stimare la base imponibile su cui applicare la tassazione è necessario ipotizzare quali potrebbero essere le conseguenze della legalizzazione della cannabis sul suo consumo. Da quando la politica della proibizione della cannabis è stata adottata a livello internazionale⁷, non si sono registrati casi di legalizzazione, ma in alcuni paesi il possesso/consumo di cannabis è stato depenalizzato⁸. L'analisi di questi casi indica che la depenalizzazione della cannabis si è tradotta in due diversi modelli. Nel primo caso, la depenalizzazione si è limitata a legalizzare gli strumenti utili alla soddisfazione di quei consumi interni definibili di "sussistenza", cioè quel consumo di cannabis che si sarebbe in ogni caso realizzato a prescindere dallo *status* legale di questa sostanza. La coltivazione domestica della cannabis (Alaska, Australia), o la vendita di suoi modesti quantitativi unitari tramite punti di vendita autorizzati (Olanda), sono stati gli strumenti utilizzati per la soddisfazione di questo consumo di "sussistenza". In questi casi non si sono registrati aumenti nella diffusione del consumo di cannabis tra la popolazione.

⁷ "Convenzione Unica sugli Stupefacenti" (Onu, New York, 1961) "Convenzione sulle Sostanze Psicotrope" (Onu, Vienna, 1971).

⁸ Per una rassegna sui principali risultati della letteratura sugli effetti della depenalizzazione: PACULA R.L. (2010).

Nel secondo caso, cioè in Olanda dalla seconda metà degli anni Ottanta alla fine degli anni Novanta, l'offerta di cannabis è stata guidata da fini commerciali, tanto da rendere questo Paese, ed Amsterdam in particolare, una assai famosa meta del "turismo dello spinello". In questo caso la diffusione nel consumo di cannabis tra la popolazione olandese è cresciuta di un terzo, ed inoltre si stima che circa l'80% della cannabis venduta in Olanda sia consumata da stranieri. Sulla base di quest'analisi dei casi di depenalizzazione, il nostro modello valuta le implicazioni fiscali di due scenari alternativi. Nel primo scenario l'equiparazione tra sigarette e cannabis è finalizzata alla soddisfazione di quei consumi definibili di "sussistenza", ossia quella domanda di cannabis che si sarebbe comunque rivolta al mercato illegale italiano. Nel secondo scenario, la legalizzazione assume una finalità pienamente commerciale. Nel primo scenario, le conseguenze sul consumo di cannabis sarebbero state nulle; nel secondo è possibile ipotizzare che, almeno nel breve termine, il consumo di cannabis sarebbe aumentato di circa un terzo. In conclusione, sulle base delle stime disponibili sulla diffusione e sul prezzo della cannabis nel mercato illegale italiano, abbiamo stimato che nel biennio 2007-2008, a seconda degli scenari ipotizzati, l'erario avrebbe potuto riscuotere una somma compresa tra circa 7 e 10 miliardi di euro dalle imposte sulle vendite di cannabis, ed una somma compresa tra 0,5 e 3 miliardi di euro da imposte sul reddito. A questi importi si può aggiungere un risparmio di spesa carceraria stimata in circa 840 milioni di euro nel biennio 2007-2008. Altre entrate sarebbero potuto derivare dall'indotto e, in uno scenario commerciale tipo "Amsterdam", dal cosiddetto "turismo dello spinello". Sempre dal punto di vista delle eventuali entrate, un'implicazione fiscale di particolare rilievo avrebbe potuto derivare dalla riduzione del crimine, sia quello legato al mercato illegale della cannabis, sia della criminalità in generale. Nello scenario in cui la legalizzazione avesse indotto una maggiore diffusione del consumo di cannabis, questi effetti positivi sarebbero comunque stati compensati da un aumento dei costi sanitari e da una riduzione della produttività dei consumatori.

L'articolo consiste nella presente introduzione, il secondo paragrafo illustra il quadro normativo vigente e la sua applicazione, il terzo il mercato illegale della cannabis in Italia, il quarto l'analisi teorica sugli effetti della legalizzazione sul consumo, il quinto l'analisi empirica dei casi di depenalizzazione della cannabis, nel sesto si formulano delle ipotesi operative, nel settimo si offre una valutazione delle imposte non riscosse sulle vendite di cannabis, nell'ottavo si valutano le mancate entrate delle imposte sul reddito di produttori e distributori, nel nono si discutono le implicazioni della proibizione/legalizzazione sulla criminalità e, in

particolare, si stima la spesa carceraria, nel decimo i costi sanitari del consumo di cannabis, nel paragrafo 11 si discute di cannabis e produttività e nell'ultimo paragrafo si offrono delle riflessioni conclusive. Riferimenti bibliografici, appendice statistica e appendici sui casi di depenalizzazione in Olanda, Alaska ed Australia completano il lavoro.

2. - Il quadro normativo e la sua applicazione

2.1 *La normativa*

La legislazione italiana in materia di stupefacenti, inclusa la cannabis, è il "Testo Unico sulla Droga", D.p.r. del 3 ottobre 1990, n. 309. In seguito al referendum popolare del 1993, ne è stata abrogata la parte che classificava il consumo come reato penale. Successive modifiche sono consistite nella Legge n. 350 (24 dicembre 2003), Legge n. 251 (5 dicembre 2005), ed infine nella Legge n. 49 (21 febbraio 2006). L'attuale legge italiana punisce sia la vendita che il consumo degli stupefacenti in generale, inclusa la cannabis. Le medesime sanzioni sono applicate senza distinzione tra stupefacenti, ma per ciascuna droga sono previste specifiche soglie quantitative al fine distinguere tra possesso finalizzato al consumo o allo spaccio. Le sanzioni previste per il reato di traffico di stupefacenti sono sia di natura pecuniaria (multa da 26mila a 260mila euro), sia detentiva (da 6 a 26 anni di reclusione). Il consumo di stupefacenti è invece punito per via amministrativa tramite l'applicazione di varie sanzioni quali il ritiro del passaporto, della patente di guida, del permesso di soggiorno, ecc. per un periodo compreso tra un mese ed un anno, sono inoltre previste altre sanzioni amministrative fino a due anni. La nuova normativa sugli stupefacenti del 2006 è stata percepita come restrittiva, in particolare per la cannabis. Essa, infatti, inasprisce le sanzioni per lo spaccio di droga, ma anche il possesso, depenalizzato, è adesso oggetto di sanzioni amministrative percepite come gravose⁹, inoltre, la soglia che differenzia il possesso di cannabis finalizzato al consumo da quello finalizzato allo spaccio è ritenuta modesta rispetto ad altre sostanze illecite (come la cocaina).

⁹ In realtà la depenalizzazione non sempre implica una maggiore permissività: si sostiene infatti che subire sanzioni amministrative, quali il ritiro della patente, permesso di soggiorno, ecc., possa essere più gravoso per il condannato delle sanzioni penali (SCHAFER C. e PAOLI L., 2006).

2.2. Applicazione della normativa da parte delle Autorità

Dal 1999 al 2008 il numero delle operazioni concernenti la cannabis¹⁰ è sceso dalle circa 11mila all'anno del periodo 1999-2002 alle circa 9mila del periodo 2003-2008, con un peso percentuale sul totale delle operazioni antidroga sceso, all'incirca, dal 50% al 40% (Tav. 2.2.1). Queste operazioni hanno condotto ad ingenti sequestri di cannabis (Tav. 2.2.3). In totale, nel periodo 2004-2008, sono state sequestrate circa 113 tonnellate di hashish (22 all'anno, il massimo 24 nel 2008), circa 18 tonnellate di marijuana (3,6 all'anno, il massimo quasi cinque nel 2007) e circa due milioni e trecento mila piante di cannabis (440mila all'anno, il massimo un milione mezzo nel 2007). Alla riduzione di operazioni, dal 1999 al 2008, è corrisposta una riduzione di segnalazioni all'Autorità Giudiziaria per violazione della normativa sulla cannabis, che sono, infatti, diminuite da circa 14mila all'anno del periodo 1999-2002 a circa 11mila all'anno del periodo 2003-2008, con un peso percentuale sul totale dei segnalati antidroga sceso, all'incirca, dal 40% al 30% (Tav. 2.2.4). Tra le segnalazioni del 2008, 12.454 hanno riguardato il traffico, 285 l'associazione e 66 altro. Queste segnalazioni hanno condotto a 9.170 arresti, 3.022 domiciliari e 73 irreperibili¹¹. Ogni anno, dal 2004 al 2008, circa 21 mila detenuti sono stati presenti nelle carceri italiane per violazioni della normativa sugli stupefacenti (art. 73). Usando come indice di ponderazione la frequenza che la cannabis ha avuto nell'insieme delle segnalazioni per violazione della legge sugli stupefacenti (Tav. 2.2.4), stimiamo che ogni anno circa settemila soggetti siano stati detenuti per violazioni della normativa sulla cannabis (Tav. 2.2.5).

3. - Il mercato (illegale) della cannabis in Italia

Nonostante l'introduzione di una normativa più stringente nel 2006, la percentuale di italiani adulti che hanno fatto uso di cannabis almeno una volta all'anno è cresciuta dall'11,2% stimato nel 2005 al 14,6% del 2008¹². In Italia la diffusione del consumo di cannabis è ben al di sopra della media europea e mondiale. Si stima, infatti, che quasi un terzo della popolazione italiana adulta (età

¹⁰ Le operazioni sono la somma delle operazioni concernenti i derivati della cannabis: hashish e marijuana, escluse le operazioni in cui sono state rinvenute piante di cannabis (circa 900, sia nel 2007 sia nel 2008).

¹¹ OEDT, *National Report 2009: Italy*, part B: Selected Issues - 11.3 Offences - tab. 11.9.

¹² UNODOC, *World Drug Report, 2008/9*, Statistical annex, tab. 3.5.1.3.

15-64) abbia consumato cannabis almeno una volta nella vita (uno su sette nell'ultimo anno)¹³. Sulla base di queste stime è possibile ipotizzare che in Italia circa 5,46 milioni di italiani abbiano consumato cannabis nel 2007 e circa 5,82 nel 2008 (Tav. 3.1). Per stimare la quantità di cannabis consumata annualmente è possibile moltiplicare il numero stimato dei consumatori per il loro consumo medio *procapite*. Sono disponibili diverse stime del consumo medio annuale di cannabis relativo alla popolazione in età lavorativa che ha fatto uso di cannabis nei dodici mesi precedenti. Bouchard (2008) stima in 94 grammi il consumo annuo dei consumatori canadesi del Quebec; Kilmer e Pacula (2009) stimano un consumo medio di 93 grammi per i consumatori statunitensi; Slack *et al.* (2008) un consumo annuo di 98 grammi per i consumatori neozelandesi; mentre l'Unodoc (2006) stima il consumo medio mondiale in 116 grammi all'anno *procapite*. Moltiplicando la media (non ponderata) delle stime del consumo annuale (circa 100 grammi) per il numero stimato di consumatori italiani, otteniamo una stima del consumo di cannabis in Italia intorno alle 1100 tonnellate nel biennio 2007-2008. I prezzi tipici al dettaglio dei principali derivati della cannabis registrati nel mercato nero italiano nel biennio 2007-2008 sono stati compresi tra un minimo di 6,6 euro al grammo per la marijuana ed un massimo di 8,8 per l'hashish. Sulla base di questi dati, abbiamo calcolato il prezzo tipico medio al grammo della cannabis come media ponderata tra prezzo tipico dell'hashish e della marijuana, usando come indice di ponderazione la percentuale che hashish o marijuana hanno avuto sul totale dei sequestri di cannabis. Il risultante prezzo tipico medio della cannabis, è stato di circa 8,1 euro al grammo nel 2007 e di circa 8,6 euro al grammo nel 2008 (Tav. 3.2). In conclusione, moltiplicando la stima della quantità di cannabis consumata nel biennio 2007-2008 in Italia per il suo prezzo tipico al dettaglio registrati nel mercato nero italiano, abbiamo stimato una spesa per l'acquisto di cannabis di quasi 9,5 miliardi di euro nel biennio 2007-2008.

4. - Analisi teorica degli effetti della legalizzazione sul mercato della cannabis

L'analisi teorica degli effetti del proibizionismo suggerisce che una legalizzazione della cannabis ne farebbe aumentare i consumi. MacCoun (1993) ha identificato sei diversi meccanismi tramite i quali un cambiamento nella normativa proibizionista, o nella sua applicazione, può influenzare il consumo di cannabis.

¹³ OEDT, *National Report 2009: Italy*, part B: Selected Issues, tab. 11.3.

- 1) La proibizione in sé costituisce un ostacolo simbolico che disincentiva il consumo di cannabis da parte dei cittadini rispettosi della legge. La rimozione del divieto consentirebbe anche a questi cittadini di consumare cannabis senza violare la legge. Da questo punto di vista la legalizzazione indurrebbe un aumento del consumo di cannabis. L'entità di questo effetto dipenderebbe dalla propensione della popolazione italiana a rispettare questa legge.
- 2) La proibizione in sé può arricchire di motivazioni il consumo di cannabis. Se questo divieto è percepito da gruppi sociali "devianti" come un tabù, la cannabis diventa per loro una sorta di "frutto proibito", ed il suo consumo acquista anche un valore simbolico. La legalizzazione, rimuovendo il tabù, indurrebbe una riduzione di quei consumi motivati da ragioni simboliche. In realtà tale simbolismo, rintracciabile nelle prime fasi di diffusione del consumo, è ormai superata da un utilizzo delle droghe prevalentemente consumistico (Meneghini, 1994), perciò, da questo punto di vista, la rimozione del divieto non dovrebbe avere conseguenze assai rilevanti.
- 3) La proibizione implica l'eventuale imposizione di sanzioni a carico dei soggetti che la violano. Adottando un modello di consumatore razionale (Becker *et al.*, 2006), il rischio di incorrere in queste sanzioni si traduce in un aumento del costo atteso del consumo di cannabis che ne riduce il consumo. La legalizzazione, rimuovendo queste sanzioni, indurrebbe un aumento dei consumi. In realtà, poiché il rischio atteso di subire sanzioni per il consumo di cannabis è quasi nullo, la legalizzazione non dovrebbe incidere significativamente sul calcolo dei costi attesi dei consumatori. Il consumo di cannabis ha, infatti, delle caratteristiche ne rendono difficile sia l'indagine che l'accertamento. In primo luogo, dal punto di vista giuridico, il consumo di cannabis appartiene alla categoria dei reati privi di vittima, reati per i quali manca la denuncia della parte offesa. Inoltre, tecnicamente, il consumo di cannabis non richiede particolari attrezzature ed è facilmente occultabile.
- 4) La proibizione della cannabis può attivare dei fattori di controllo sociale di tipo informale a danno dei consumatori che si aggiungono al rischio di sanzioni formali. La legalizzazione, rimuovendo anche queste sanzioni, indurrebbe quindi un aumento dei consumi. L'entità di questo effetto dipende dalla capacità della normativa di influenzare il sentimento sociale della popolazione cui è applicata. Al proposito si ricorda che la popolazione italiana si è formalmente espressa a favore della depenalizzazione del consumo di droga: il 18 e 19 aprile 1993, un referendum popolare abrogativo ha raccolto quasi 35 milioni di voti a favore su circa 37 milioni di votanti.

- 5) L'applicazione della normativa proibizionista riduce la possibilità di acquisto della cannabis, relegandone gli scambi nel mercato nero. La sua legalizzazione aumenterebbe sia la quantità che l'accessibilità dei punti di vendita, inducendo pertanto un aumento dei consumi. L'entità di questo effetto dipenderebbe dalla differenza che introdurrebbe rispetto alla attuale diffusione del mercato nero della cannabis. Al proposito, i dati sulla diffusione dello spaccio mostrano che questa sostanza è già facilmente reperibile sul territorio italiano¹⁴, il quale non solo si presta facilmente alla coltivazione della cannabis¹⁵ ma da cui, soprattutto, è agevole il contrabbando dal Marocco (uno dei principali produttori mondiali) e dall'Est¹⁶.
- 6) L'applicazione della normativa proibizionista, riducendo l'offerta di cannabis, dovrebbe farne aumentare il prezzo di vendita e così ridurre il consumo. La legalizzazione della cannabis ne aumenterebbe l'offerta, ridurrebbe il prezzo e pertanto farebbe aumentare il consumo di cannabis. La notevole differenza tra i prezzi all'ingrosso e quelli al dettaglio della cannabis nel mercato illegale italiano conferma che il prezzo di vendita della cannabis contenga un forte premio per il rischio, connesso allo spaccio di questa sostanza¹⁷. La legalizzazione, rimuovendo le sanzioni connesse al traffico eliminerebbe questa componente, riducendo sensibilmente il prezzo di vendita della cannabis. In assenza di correttivi fiscali, questa riduzione del prezzo indurrebbe un aumento dei consumi.

5. - Analisi dell'evidenza empirica sugli effetti della depenalizzazione della cannabis

Da quando, a partire dagli anni sessanta, la politica della proibizione della cannabis si è diffusa a livello internazionale, non si sono registrati casi di legalizzazione, ma in alcuni paesi il possesso/consumo di cannabis è stato depenalizzato. In generale pare che il consumo di cannabis sia più sensibile al suo prezzo che al suo *status* legale, ovvero i cambiamenti di regime paiono essere stati efficaci nella misura in cui sono riusciti ad influenzare il prezzo della cannabis¹⁸. I casi di depenalizzazione più stu-

¹⁴ OEDT, *National Report: Italy 2009*, part B: Selected Issues, par. 11.2 e DCSA (2008).

¹⁵ Il significativo numero di piante sequestrate in Italia (Tav. 2.2.2) conferma l'idoneità del territorio italiano alla coltivazione della cannabis.

¹⁶ OEDT, *National Report: Italy 2009*, part B: Selected Issues, par. 11.2, e DCSA (2008).

¹⁷ UNODOC, *World Drug Report 2006/9*, Statistical annex, tab. 3.4.3.

¹⁸ Per una rassegna sui principali risultati della letteratura sugli effetti della depenalizzazione: PACULA R.L. (2010).

diati sono il caso olandese, in cui tale politica è stata adottata sin dagli anni Settanta, e i più recenti casi in Alaska ed Australia. L'evoluzione di queste vicende (illustrata nelle appendici A.1 e A.2) indicano che la depenalizzazione della cannabis si è tradotta in due diversi modelli. Nel primo caso, la depenalizzazione si è limitata a legalizzare gli strumenti utili alla soddisfazione di quei consumi interni definibili di "sussistenza", cioè quel consumo di cannabis che si sarebbe comunque realizzato a prescindere dallo *status* legale di questa sostanza. La coltivazione domestica della cannabis (Alaska, Australia), o la vendita di suoi modesti quantitativi unitari presso luoghi di consumo autorizzati (Olanda), sono stati gli strumenti utilizzati alla soddisfazione di questo consumo di "sussistenza". In questi casi non si sono registrati aumenti nella diffusione del consumo di cannabis tra la popolazione. Nel secondo caso, cioè in Olanda dalla seconda metà degli anni Ottanta alla fine degli anni Novanta, l'offerta di cannabis è stata guidata da fini commerciali: il numero e l'attività dei *coffee-shops* sono cresciuti fino a rendere questo Paese, ed Amsterdam in particolare, la più famosa meta europea del "turismo dello spinello". In questo caso la diffusione nel consumo di cannabis tra la popolazione olandese è cresciuta di un terzo; si stima inoltre che circa l'80% della cannabis venduta in Olanda sia consumata da stranieri. La depenalizzazione non ha invece mai avuto effetti sul prezzo della cannabis. Per spiegare questo risultato Pacula *et al.* (2007) sostengono che la funzione d'offerta della cannabis sia inclinata positivamente, in questo caso la rimozione delle sanzioni può aver indotto un aumento dell'offerta che è stato compensato da un simultaneo aumento della domanda, con effetti nulli sul prezzo. Mentre questa ipotesi potrebbe giustificare il caso olandese, negli altri casi si può ipotizzare che il mantenimento del divieto di commercio non abbia modificato il premio per il rischio contenuto nel prezzo della vendita (illegale) della cannabis.

6. - Prezzo d'offerta e quantità venduta della cannabis "legalizzata": formulazione degli ipotetici scenari

Per quanto concerne quale sarebbe stato il prezzo d'offerta della cannabis in regime di legalizzazione formuliamo l'ipotesi che esso sarebbe stato amministrativamente fissato ad un livello esattamente pari a quello affermatosi nel mercato illegale ($P_L = P_B$). Questa è una ipotesi semplificatrice, che ci consente di ignorare quale fosse l'elasticità della domanda di cannabis rispetto al suo prezzo¹⁹, ed inoltre im-

¹⁹ PACULA R.L. *et al.* (2000) discutono l'argomento dell'elasticità della cannabis rispetto al suo prezzo.

plica che la funzione d'offerta di cannabis sarebbe stata infinitamente elastica, ossia in grado di soddisfare qualunque livello di domanda al prezzo regolamentato. Per quanto concerne la domanda di cannabis, sulla base dell'analisi teorica e dell'evidenza empirica ipotizziamo che si sarebbe potuti instaurare due diversi scenari:

- A) Scenario "commerciale": come accaduto tra la fine degli anni Ottanta e l'inizio degli anni Novanta in Olanda, anche in Italia la legalizzazione si sarebbe potuta tradurre in una piena commercializzazione della cannabis, soddisfatta tramite una rete di distributori simili ai *coffe-shops* olandesi. In questo caso la diffusione del consumo annuale di cannabis sarebbe cresciuta fino a coinvolgere circa il 19% della popolazione italiana adulta, e, in proporzione, la quantità legalmente consumata (Q_L) sarebbe stata circa un terzo maggiore di quella già consumata illegalmente (Q_B);
- B) Scenario di "sussistenza": se invece la legalizzazione si fosse limitata a far emergere dal mercato illegale i consumi interni, qui definiti di "sussistenza", la diffusione del consumo annuale di cannabis sarebbe rimasta ai livelli già registrati e la distribuzione di cannabis sarebbe stata eseguita tramite l'esistente rete di tabaccherie.

6.1 Ripartizione dei ricavi delle vendite

La nostra ipotesi di legalizzazione estende al mercato della cannabis la normativa fiscale applicata nel mercato italiano delle sigarette: per un valore di vendita delle sigarette pari a 100 euro, 58,5 euro sono versate alle casse dell'erario a titolo di accisa fissa e 17 euro a titolo di Iva, 10 euro vanno al rivenditore e 14,5 al produttore²⁰. Applicando la suddetta ripartizione degli incassi al mercato della cannabis avremmo avuto che, per ogni grammo di cannabis venduto nel 2007 al prezzo al dettaglio di 8,1 euro al grammo, lo stato avrebbe potuto incassare circa 6 euro (Graf. 6.1.1). La quota destinata ai produttori e distributori di cannabis sarebbe stata di circa 2 euro al grammo, un importo simile ai prezzi all'ingrosso registrati sul mercato nero italiano²¹. Sebbene questo ricavo sia modesto rispetto al prezzo al dettaglio della cannabis registrato nel mercato nero, si ricorda che tale prezzo contiene una enorme componente di premio per il rischio, in particolare

²⁰ Si noti che, nonostante l'imposizione di una tassazione così elevata e nonostante che in Italia sbocchino alcune delle principali rotte del contrabbando (come quella balcanica e quella nordafricana), piuttosto che un terminale di consumo, l'Italia è un crocevia di transito del commercio illegale che si indirizza soprattutto verso i mercati nord europei.

²¹ UNODOC, *World Drug Report 2008 e 2009*, Statistical annex, tab. 3.4.3.

dello spaccio al dettaglio²². Una volta depurato il prezzo d'offerta da questa componente, 2 euro al grammo dovrebbero essere sufficienti a coprire i costi di produzione e distribuzione della cannabis. Lo studio di Caulkins (2010) sui costi di produzione della cannabis in California, una regione climaticamente e geomorfologicamente simile all'Italia, conferma che due euro al grammo sono più che sufficienti a coprire i costi di produzione, i quali sono stati stimati in circa 20/35 dollari per oncia di prodotto finito.

La sostenibilità della ripartizione dei ricavi qui proposta si basa sull'ipotesi di fondo che l'attuale mercato nero della cannabis sia perfettamente concorrenziale²³. L'assunzione della concorrenzialità implica che il prezzo d'offerta sia uguale al costo marginale (che include il premio per il rischio di spaccio al dettaglio). In questo caso, la supposta ripartizione dei ricavi implicherebbe che un'impresa otterrebbe lo stesso profitto marginale (nullo) sia rimanendo nel mercato nero sia emergendo da esso. Tradizionalmente si tende a supporre che in caso di indifferenza nei profitti attesi (incluso il premio per il rischio), gli operatori preferiscano rispettare le legge, cioè emergere dal mercato nero. Inoltre, l'emersione dall'illegalità sarebbe particolarmente conveniente per le imprese produttrici di cannabis: non solo consentirebbe loro di ridurre i costi di produzione, depurandoli dalle componenti di rischio, ma anche di sfruttare le economie di scala che caratterizzano la coltivazione di questa pianta. Adesso in Italia, al fine di occultare agli investigatori la persecuzione di un illecito, la coltivazione all'aperto di cannabis avviene in piantagioni generalmente di piccola dimensione, sparse sul territorio e spesso in luoghi difficilmente accessibili. Si tratta di una tecnica di produzione chiaramente costosa e poco efficiente. La storia della coltivazione della canapa in Italia (diffusa fino agli anni cinquanta) suggerisce che si potrebbero adottare tecniche più efficienti qualora fosse mutato il contesto legale. Anche la coltivazione di cannabis in serra gode di economie di scala, ma, sempre per ragioni di riservatezza, la coltivazione su scala domestica è la più diffusa in Italia²⁴.

²² Una conferma che il prezzo al dettaglio della cannabis contiene una enorme componente di premio per il rischio proviene dall'analisi comparata presentata nel *World Drug Report 2010* dell'UNODOC (par. 2.4.3 "Prices"). La regressione dei prezzi al dettaglio su quelli all'ingrosso registrati nel 2008 suggerisce che questa componente di premio per il rischio connesso allo spaccio sia stata pari al 129% per la marijuana (fig. 167) ed al 110% per l'hashish (fig. 170).

²³ Sebbene i media preferiscano rappresentare il mercato nero delle droghe come oligopolistico, l'analisi empirica conferma le ragioni teoriche circa la spontanea natura concorrenziale di questi mercati illegali, in particolare di quello della cannabis (REUTER P.H., 1983).

²⁴ Nel 2008 in Italia operavano circa 150 "Grow Shops", negozi specializzati nella vendita dei materiali utili alla coltivazione domestica della cannabis (www.dolcevitaonline.it).

7. - Stima delle imposte (non riscosse) sulle vendite di cannabis in Italia

L'applicazione di una *sin tax*²⁵ pari al 75,5% del valore stimato della vendita della cannabis avrebbe comportato la riscossione per l'erario nel biennio 2007-2008 di una imposta compresa tra 7 e 10 miliardi di euro, a seconda degli scenari ipotizzati²⁶. In dettaglio: circa 3,359 miliardi di euro nel 2007 e circa 3,775 nel 2008 nell'ipotesi di "sussistenza" ($Q_L=Q_B$), circa 4,477 miliardi di euro nel 2007 e circa 5,302 nel 2008 nell'ipotesi "commerciale" ($Q_L=4/3 Q_B$). Inoltre, nell'ipotesi che la legalizzazione avesse indotto uno scenario di tipo commerciale, le vendite di cannabis sarebbero potute aumentare non solo a causa della diffusione del consumo tra la popolazione italiana ma anche per gli acquisti dei cosiddetti "turisti della spinello". L'Ufficio del Turismo Olandese afferma che nel 2007 un turista straniero su quattro (cioè circa 1 milione e 120mila turisti) hanno visitato un *coffee-shop* in Olanda, di questi il 10% ha affermato che questa era la ragione della sua visita (Van Laar *et al.*, 2010). L'ipotesi che parte della vendita di cannabis olandese soddisfi anche i cittadini dei paesi limitrofi è inoltre confermata dalla correlazione inversa tra la distanza di questi paesi dall'Olanda e percezione della disponibilità di cannabis ((MacCoun, 2010a, Graf. 5). Stimando che questi "turisti dello spinello" acquistino un grammo a testa, è possibile stimare in circa 1125kg la quantità vendita, per un controvalore al dettaglio compreso tra 311 e 560 milioni di euro all'anno, cioè circa l'80% della produzione olandese è in realtà destinata alla vendita a stranieri (MacCoun, 2010a). In questo studio non si offre alcuna stima di quale sarebbero state le entrate fiscali dalle vendite di cannabis a turisti stranieri in Italia. Si nota tuttavia che queste entrate sarebbero state maggiori quanto più lo scenario italiano fosse stato simile al modello olandese appena descritto.

²⁵ La scelta dell'aliquota del 75,5%, identica a quella imposta sui tabacchi lavorati, è solo per ragioni di semplicità. In realtà non sappiamo se essa corrisponda alla "aliquota ottimale" suggerita da BECKER G.S., GROSSMAN M. e MURPHY K.M. (2006). Cioè se essa corrisponda a quella aliquota che, se applicata al prezzo di vendita della cannabis, indurrebbe un livello di consumo compatibile con la massimizzazione vincolata del benessere sociale.

²⁶ L'ipotesi implicitamente adottata nella nostra simulazione, è quella di una completa emersione delle vendite di cannabis dal mercato nero in seguito alla sua legalizzazione. Nella misura in cui parte degli scambi fossero restati nel mercato nero, i nostri calcoli sovrastimano gli introiti che l'erario avrebbe potuto riscuotere dalle imposte sulle vendite di cannabis.

8. - Produzione e distribuzione: le imposte (non riscosse) sul reddito degli impiegati

Il proibizionismo implica che l'erario non possa riscuotere l'imposta sul reddito personale ottenuto da traffici dichiarati illeciti. Solamente nel caso in cui i trafficanti siano condannati, i proventi del traffico di droga sono confiscati, altrimenti i redditi dei trafficanti sono esentasse. La legalizzazione consentirebbe invece all'erario di riscuotere le imposte sul reddito dei produttori e distributori di cannabis. Nello scenario di "sussistenza", in cui i consumi non sarebbero aumentati e la vendita eseguita tramite le esistenti tabaccherie, la ricaduta occupazionale sarebbe stata modesta e le entrate sul reddito sarebbero state complessivamente circa 530 milioni di euro nel biennio 2007-2008. Nello scenario "commerciale", in cui i consumi sarebbero aumentati e la vendita eseguita tramite una nuova rete di *coffee-shops* stile Amsterdam, le conseguenze occupazionali sarebbero state assai rilevanti e le entrate sul reddito sarebbero state complessivamente circa 3,4 miliardi di euro nel biennio 2007-2008. Altre entrate sarebbero inoltre potute derivare dall'imposta sul reddito dei lavoratori impiegati nell'indotto della cannabis.

8.1 *Le imposte (non riscosse) sul reddito degli impiegati nella produzione di cannabis*

Una semplice stima delle entrate da imposte sul reddito dei produttori di cannabis deriva dall'applicazione dell'aliquota Irpef minima sulla quota di proventi destinati ai produttori. Nel biennio 2007-2008 questi redditi sarebbero stati pari al 14,5% dei ricavi della vendita di 9,5-12,6 miliardi di euro di cannabis, cioè 1,4-1,8 miliardi di euro a seconda degli scenari ipotizzati. Applicando l'aliquota Irpef del 23% su questi redditi²⁷, l'erario avrebbe potuto incassare circa 315-420 milioni di euro nel biennio 2007-2008. Altre entrate sarebbero poi potute derivare dalle imposte sul reddito dei lavoratori impiegati nell'indotto della produzione di cannabis. Per quanto concerne le ricadute occupazionali dirette, occorre ricordare che quella della cannabis è una coltivazione stagionale ad elevata intensità di impiego del fattore lavoro. Si stima che la produzione di una oncia di prodotto finito in California richieda l'impiego di una quantità di fattore lavoro compresa tra le 2,2 e le 3,3 ore, a seconda del grado di meccanizzazione del processo, cioè circa 100 ore al kg. nell'ipotesi di elevata meccanizzazione del processo produttivo

²⁷ Abbiamo ipotizzato che i lavoratori addetti alla produzione della cannabis fossero stati pagati come i loro colleghi operai del settore agricolo stagionale, cioè 15.262 euro all'anno nel 2007 e 15.337 nel 2008 (ISTAT, *Retribuzioni*, cat. A011. - "Coltivazione di colture agricole non permanenti").

(Caulkins, 2010). Poiché il clima e la morfologia della California è assai simile a quella dell'Italia, usiamo questa stima anche per calcolare la quantità di lavoro che sarebbe stata necessaria per produrre la cannabis consumata in Italia. Moltiplicando la quantità che sarebbe stata consumata in Italia (1.100 o 1.500 tonnellate a seconda degli scenari ipotizzati) per l'impiego del fattore lavoro stimato, otteniamo che nel biennio 2007-2008 sarebbero state necessarie circa 110-150 milioni di ore lavoro per soddisfare la domanda italiana di cannabis. Ipotizzando che ogni lavoratore impiegato fornisca 2.000 ore lavorative annuali, ciò significherebbe l'impiego di circa 55-75mila lavoratori all'anno.

Rimangono tuttavia dei dubbi sulla completa emersione delle imprese produttrici di cannabis dal mercato nero, soprattutto in ragione della tradizionale diffusione nel settore agricolo italiano del lavoro nero e di altre forme di evasione fiscale. Si avvisa pertanto il lettore che i nostri calcoli, omettendo questo fattore di correzione, possono sovrastimare le entrate fiscali che sarebbero potute derivare dalle imposte sul reddito dei produttori di cannabis.

8.2 *Le imposte (non riscosse) sul reddito degli impiegati nella distribuzione di cannabis*

L'intensità dell'impiego del fattore lavoro nella distribuzione della cannabis dipende dagli scenari che si possono ipotizzare in conseguenza della sua legalizzazione. Se fosse prevalsa una finalità commerciale, come in Olanda, la distribuzione si sarebbe potuta avvalere di una rete di *coffee-shops*. Simili ai caffè italiani, i *coffee-shops* sono locali confortevoli (dotati di impianti di intrattenimento e ristorazione), finalizzati alla vendita di cannabis da consumare *intra muros*. Attualmente in Olanda ne sono operativi circa settecento (MacCoun, 2010a), ed impiegano circa 3.400 lavoratori. Essi sono sufficienti a distribuire cannabis a 900mila consumatori: in media un *coffee-shop* ogni 125 consumatori²⁸. In uno scenario commerciale di questo genere, è verosimile che, in seguito alla sua legalizzazione, la diffusione del consumo di cannabis aumenti di circa un terzo, per cui in Italia avremmo avuto circa 7,3 milioni di consumatori nel 2007 e circa 7,8 nel 2008. Assumendo che in Italia ci fosse stata una medesima intensità di impiego del fattore lavoro, nei circa 60mila *coffee-shops* italiani sarebbero stati impiegati circa 280mila addetti nel 2007 e 300mila nel 2008. Ipotizzando una loro retribuzione annuale pari a quella riscossa dagli operai nel settore della di-

²⁸ Considerando che nel 2005 il tasso di diffusione del consumo di cannabis tra la popolazione olandese adulta è stato del 5,4%, è possibile stimare che nel 2005 siano stati quasi 900 mila i consumatori di cannabis in Olanda.

stribuzione²⁹, questi avrebbero riscosso un reddito da lavoro di circa 10 miliardi di euro nel biennio 2007-2008. Applicando il secondo scaglione Irpef (27%), la retribuzione di questi lavoratori avrebbe implicato la riscossione di circa tre miliardi di euro nel biennio 2007-2008 da imposte sul reddito degli addetti alla distribuzione della cannabis. Se la legalizzazione fosse invece stata finalizzata alla soddisfazione dei consumi di “sussistenza” tramite una rete di venditori da asporto autorizzati, cioè le attuali tabaccherie, l’impiego del fattore lavoro nella distribuzione della cannabis sarebbe stato sensibilmente inferiore che nello scenario precedente e la diffusione del consumo non sarebbe aumentata. Si noti, infatti, che in Italia quasi tutti i consumatori di cannabis sono anche tabagisti³⁰, pertanto la legalizzazione della cannabis avrebbe permesso ai fumatori di cannabis di acquistare in tabaccheria anche la materia prima (cannabis) oltre agli strumenti per il consumo (cartine e sigarette). Le ricadute occupazionali sarebbero state quindi modeste. Una semplice stima delle entrate da imposte sul reddito dei distributori di cannabis deriva dall’applicazione dell’aliquota Irpef minima sulla quota di proventi destinati ai distributori. Nel biennio 2007-2008 questi redditi sarebbero stati pari al 10% dei ricavi della vendita di 9,5 miliardi di euro di cannabis, cioè 950 milioni di euro. Applicando l’aliquota Irpef del 27% su questi redditi, l’erario avrebbe potuto incassare circa 255 milioni di euro nel biennio 2007-2008. Altre entrate sarebbero poi potute derivare dalle imposte sul reddito dei lavoratori impiegati nell’indotto della distribuzione di cannabis.

9. - Implicazioni sul costo della criminalità cannabis-correlata

La diffusione della cannabis, la più popolare tra le droghe illegali, ha causato un significativo numero di reati sistemici e di violazioni della normativa³¹. In generale, la proibizione del commercio di un bene ha come effetto indesiderato la creazione di un mercato illegale alimentato da coloro che comunque desiderano

²⁹ Cioè 18.722 euro all’anno nel 2007 e 19.380 nel 2008. ISTAT, *Retribuzioni*, cat. G472. – “Commercio al dettaglio prodotti agricoli, alimentari e tabacco in esercizi”.

³⁰ In Italia la modalità assolutamente dominante dell’assunzione di cannabis consistente nella miscela dei suoi derivati (hashish o marijuana) con del tabacco, utile alla combustione dei derivati stessi.

³¹ Nella maggior parte dei paesi europei i reati segnalati riguardano l’uso o la detenzione personale, in particolare di cannabis, il cui peso nell’ambito dei reati droga-collegati variava nel 2005 tra il 42% ed 74% (OEDT, *Bollettino Statistico 2007* - tab. DLO-3).

acquistare quel bene a prescindere dal suo *status* legale. In altri termini la proibizione della cannabis ha avuto come effetto indesiderato lo sviluppo di un mercato illegale, in cui si commettono sia violazioni della normativa che reati sistemici. Il proibizionismo è infatti incompatibile con la tutela giudiziaria dei diritti di proprietà sui beni o le attività vietate. I soggetti coinvolti nel mercato nero devono quindi risolvere le loro controversie in via extra-giudiziale, ricorrendo alla violenza come strumento residuale. Al proposito, le organizzazioni criminali realizzano delle economie di scopo impegnandosi anche nel narcotraffico, opportunità di profitto che pare essere stata sfruttata dalle organizzazioni criminali italiane (DCSA, 2010). Inoltre, il proibizionismo può aver indirettamente indotto un aumento nella criminalità, nella misura in cui l'utilizzo di risorse investigative, giudiziarie, ecc. per l'applicazione della normativa proibizionista sia stato incompatibile con l'utilizzo degli stessi strumenti per il controllo di altre attività criminali. Se ammettiamo che l'applicazione del proibizionismo abbia sottratto risorse utili ad altri compiti di polizia, è ipotizzabile che ciò possa aver indotto un aumento generalizzato della criminalità. In conclusione, la legalizzazione della cannabis, qualunque scenario avesse indotto, avrebbe comunque ridotto il crimine indotto dal proibizionismo: sia direttamente, rimuovendo il mercato illegale della cannabis, sia indirettamente, consentendo che forze dell'ordine e magistratura potessero utilizzare le risorse liberate dalla persecuzione dei reati di cannabis verso la persecuzione della criminalità in generale³². Le presumibili conseguenze erariali sarebbero potute consistere in un aumento delle entrate, il quale sarebbe stato proporzionale al danno subito dall'economia italiana a causa del crimine indotto dalla proibizione della cannabis. Sebbene questi introiti sarebbero potuti essere significativi, questo lavoro si limita a valutare una implicazione fiscale del proibizionismo della cannabis, cioè la spesa carceraria sostenuta in Italia nel biennio 2007-2008 per la detenzione dei soggetti condannati per violazione della normativa sulla cannabis. Non abbiamo invece valutato le risorse impegnate dalle forze dell'ordine e dalla magistratura per l'applicazione della normativa proibizionista. Ribadiamo che la rimozione del proibizionismo in sé non avrebbe implicato una riduzione delle risorse pubbliche già destinate a forze dell'ordine e magistratura, ma piuttosto un utilizzo delle stesse per i restanti compiti di queste istituzioni. Ad eccezione delle spese carcerarie, l'effetto in sé della legalizzazione sarebbe stato quindi nullo per i capitoli di spesa destinati a forze dell'ordine e magistratura.

³² MIRON J.A. (1997) sostiene che l'evidenza empirica corrobora l'ipotesi che il proibizionismo induca un aumento della criminalità.

9.1 *Stima della spesa carceraria*

Una stima della spesa carceraria implicata dal proibizionismo della cannabis è ottenibile moltiplicando il costo medio annuo di detenzione per il numero di detenuti che scontano condanne per reati di cannabis. Sulla base dei dati del Ministero della Giustizia (DAP) e della Corte dei Conti, è possibile stimare che il costo medio annuale di ogni singolo detenuto sia stato di 72.431 euro nel 2007 e di 55.502 nel 2008³³. Il numero di detenuti per violazioni della normativa sulla cannabis è stato stimato sulla base dei detenuti *ex art. 73*, usando come indice di ponderazione la frequenza che la cannabis ha avuto nell'insieme delle denunce per violazione della legge sugli stupefacenti (Tav. 2.2.4). Abbiamo pertanto calcolato che circa 6.200 soggetti nel 2007 e 7000 nel 2008 siano stati detenuti nelle carceri italiane per violazioni della normativa sulla cannabis, con un risultante costo di detenzione di circa 451 milioni di euro per il 2007 e di circa 386 milioni di euro per l'anno 2008. Questa spesa sarebbe comunque stata risparmiata qualunque fosse stato lo scenario che si fosse instaurato in seguito alla legalizzazione della cannabis³⁴. Occorre tuttavia precisare che, in analogia con il contrabbando delle sigarette, questa riduzione di spesa sarebbe stata in parte compensata dai costi di carcerazione dei contrabbandieri di cannabis.

10. - Il costo sanitario del consumo di cannabis

I principali costi sanitari consistono nei danni, sia fisici che mentali, indotti dal consumo di cannabis, sia quelli a carico degli stessi consumatori, sia quelli sopportati dal resto della società. Queste esternalità non consistono solo nella partecipazione della società alle spese mediche dovute all'abuso di cannabis, ma anche nei danni a terzi indotti dal consumo (es. incidenti di guida). I disturbi fisici che possono essere causati dall'abuso sistematico e prolungato della cannabis rientrano in quelli tipicamente causati dall'inalazione di prodotti combustibili, cioè malattie respiratorie e cardio-vascolari, l'abuso può inoltre accompagnarsi a disturbi psicotici quali depressione e schizofrenia³⁵. Occorre tuttavia ricordare che la carriera tipica del consumatore di cannabis inizia in età giovanile e termina generalmente

³³ CENTRO STUDI DI RISTRETTI ORIZZONTI (2009).

³⁴ Si noti che una simile riduzione della spesa carceraria si sarebbe potuta ottenere anche con una depenalizzazione del possesso di cannabis.

³⁵ Per una rassegna sui danni fisici e mentali indotti dall'uso di cannabis: WITTON J. (2008) e HALL W. e DAGENHARDT L. (2009).

prima che questi effetti nocivi possano cronicizzarsi. Infatti, sebbene con i suoi oltre cinque milioni di consumatori la cannabis sia largamente la sostanza illecita più consumata in Italia³⁶, il suo uso non ha indotto conseguenze sanitarie comparabili alle altre droghe. Tra i circa 500-600 decessi annuali causati in Italia dal consumo di droghe non si registrano casi imputabili ad “overdose” da cannabis, i trattamenti sanitari per abuso di cannabis sono una piccola parte del totale³⁷, e spesso la richiesta di trattamento non è spontanea, ma indotta dalla segnalazione ai servizi sanitari dei colpevoli di reati associati al consumo di cannabis da parte del sistema giudiziario³⁸. Inoltre, il consumo di cannabis non induce la diffusione di malattie infettive tra i consumatori o la popolazione in generale, mentre lo scambio di siringhe infette, spesso connesso all’uso di droghe “pesanti”, è causa della diffusione di malattie quali Aids, epatite, ecc. In conclusione, nello scenario in cui la legalizzazione della cannabis avesse indotto un aumento della diffusione dei consumi, i costi sanitari sarebbero proporzionalmente aumentati. Nel caso invece in cui la legalizzazione non avesse causato un aumento dei consumi, la spesa sanitaria si sarebbe potuta ridurre, si sarebbero infatti risparmiati i costi dei trattamenti imposti dalle autorità ai soggetti colpevoli di violazione della normativa proibizionista.

11. - Consumo di droga e produttività

Una delle ragioni sostenute a favore della proibizione del consumo di cannabis consiste nell’ipotesi che il consumo di cannabis riduca la produttività dei consumatori. Tuttavia, non solo l’evidenza empirica è inconcludente, ma anche la proibizione in sé può indurre degli effetti indesiderati che sortiscono un simile effetto. Varie ipotesi sostengono che l’uso di droga in generale possa implicare una perdita di produttività del consumatore. Si sostiene che il drogato abbia minori probabi-

³⁶ Nell’anno 2008 sono stati stimati i seguenti tassi di diffusione del consumo annuale di droghe tra la popolazione italiana adulta (età 15-64): anfetamine 0,6%, ecstasy 0,7%, oppiacei 0,7% cocaina 2,2%, cannabis 14,6% (UNODOC, *WDR 2010*, par. 4.2).

³⁷ Nell’anno 2008 sono stati registrati 167.011 trattamenti per abuso di droga, di questi il 73,3% riguardava gli oppiacei, il 16,7% la cocaina e solo il 9,1% la cannabis (UNODOC, *WDR 2010*, tab. 4.2.2.4)

³⁸ Infatti l’OEDT (2007) nella sua “*Relazione Annuale 2007*” afferma che: «I risultati degli studi dimostrano che la maggioranza delle persone indirizzate al trattamento dal sistema giudiziario sono consumatori principali di cannabis» (p. 44).

lità di essere impiegato, sia indotto all'assenteismo e sia meno produttivo. L'evidenza empirica sull'associazione tra consumo di droga e produttività è tuttavia inconcludente³⁹, sia che si assuma il reddito individuale come *proxy* della produttività, o si usino altri indicatori quali la probabilità di trovare un impiego, il tasso di assenteismo o il tempo speso in trattamento. Il proibizionismo in sé può, invece, avere delle implicazioni negative sulla produttività dei consumatori di cannabis. In primo luogo, l'adozione di una normativa proibizionista implica l'applicazione di sanzioni nei confronti dei suoi violatori, limitazioni della libertà personali (dalle sanzioni amministrative fino al carcere) che riducono le capacità contributive di questi soggetti. In secondo luogo la condanna morale associate all'uso di sostanze illecite può marginalizzare i consumatori, danneggiandone le capacità produttive. Questa è una delle ragioni che hanno condotto alla depenalizzazione del consumo di cannabis nell'Australia Occidentale. Si è infatti sostenuto che la normativa proibizionista, mentre aveva effetti modesti sulla riduzione della diffusione del consumo, riduceva significativamente le probabilità di trovare un impiego per coloro che erano stati sanzionati per aver violato tale normativa. In conclusione, nel caso in cui la legalizzazione non avesse causato un aumento dei consumi, la rimozione delle sanzioni (penali o amministrative) a carico dei soggetti che avevano violato la normativa proibizionista avrebbe indotto un aumento della produttività di questi lavoratori e, quindi, un aumento delle entrate fiscali. Altrimenti, sebbene il legame tra produttività ed uso di cannabis non sia empiricamente appurato, è possibile ipotizzare che la legalizzazione della cannabis avrebbe potuto indurre una perdita di produttività (e quindi erariale) proporzionale all'aumento della diffusione dei consumi causati dalla legalizzazione stessa.

12. - Conclusioni

Questo studio ha cercato di valutare le implicazioni fiscali di una estensione anche al mercato della cannabis della normativa che regola in Italia il commercio delle sigarette. In dettaglio, si è ipotizzato che nel biennio 2007-2008:

- a) la produzione, manifattura, distribuzione all'ingrosso ed al dettaglio della cannabis fosse concessa ad operatori autorizzati;
- b) la funzione d'offerta della cannabis fosse infinitamente elastica ad un prezzo pari a quello registrato nel mercato illegale;

³⁹ Per una rassegna sugli studi sulla produttività dei consumatori di droga: PACULA R.L. *et AL.* (2009, par. 4.2.1, pp. 19-20).

c) il 75,5 % del ricavo delle vendite della cannabis fosse andato all'erario (come accisa ed Iva) ed il restante a compensare produttori e distributori.

Le poste qui stimate sono consistite solamente nella spesa carceraria sostenuta per la detenzione di condannati per violazione della normativa vigente e nelle imposte (non riscosse) sulle vendite e redditi da cannabis. Numerose altre implicazioni fiscali sono omesse dalla nostra analisi. In primo luogo non abbiamo valutato le conseguenze indesiderate del proibizionismo, tra cui il costo fiscale del crimine indotto dallo sviluppo di un mercato illegale della cannabis. In secondo luogo non abbiamo valutato le conseguenze di una eventuale diffusione del consumo di cannabis, causata dalla sua legalizzazione, sulla produttività dei consumatori e sui costi sanitari. Per questo motivo i nostri risultati rappresentato solo una parziale valutazione delle possibili implicazioni fiscali dell'equiparazione normativa della cannabis alle sigarette. Questo lavoro si è limitato a valutare l'importo di due poste contabili, un'uscita ed una (mancata) entrata, relativamente ad un biennio, il 2007-2008, per il quale disponiamo di dati sufficientemente omogenei. La prima posta, cioè la spesa carceraria, è abbastanza facilmente stimabile (il numero ed il costo di detenzione dei carcerati è desumibile dalla contabilità pubblica) ed interpretabile (è ovvio che la legalizzazione della cannabis eliminerebbe questa ragione della loro incarcerazione). Nell'ambito delle spese direttamente sostenute per l'applicazione della normativa proibizionista essa rappresenta, più di altre, una spesa che può essere risparmiata legalizzando la cannabis. La stima delle (mancate) imposte sulle vendite di cannabis richiede di ipotizzare quali potrebbero essere le conseguenze della legalizzazione della cannabis sul suo consumo. Sulla base dell'analisi teorica e dell'evidenza empirica, abbiamo individuato due possibili scenari. Nel primo, l'equiparazione tra sigarette e cannabis è finalizzata alla soddisfazione di quei consumi definibili di "sussistenza", ossia quella domanda di cannabis che si è comunque rivolta al mercato illegale italiano. Nel secondo scenario, la legalizzazione assume una finalità pienamente commerciale. Nel primo scenario, le conseguenze sul consumo di cannabis sono presumibilmente nulle, nel secondo è possibile che, almeno nel breve termine, il consumo di cannabis aumenti di circa un terzo. A seconda degli scenari ipotizzati, abbiamo stimato che nel biennio 2007-2008 l'erario avrebbe potuto riscuotere una somma compresa tra circa 7 e 10 miliardi di euro dalle imposte sulle vendite di cannabis, ed una somma compresa tra circa 0,5 e 3 miliardi di euro da imposte sul reddito. A questi importi si può aggiungere un risparmio di spesa carceraria stimato in circa 840 milioni di euro nel biennio 2007-2008. Altre entrate sarebbero potute derivare dall'indotto e, in uno scenario commerciale tipo "Amsterdam", dal cosiddetto

“turismo dello spinello”. Sempre dal punto di vista delle eventuali entrate, una implicazione fiscale di particolare rilievo sarebbe potuta derivare dalla riduzione del crimine, sia quello legato al mercato illegale della cannabis, sia della criminalità in generale. Nello scenario in cui la legalizzazione avesse indotto una maggiore diffusione del consumo di cannabis, questi effetti positivi sarebbero comunque stati compensati da un aumento dei costi sanitari e da una riduzione della produttività dei consumatori. Una comparazione delle principali implicazioni fiscali del cambio di regime proposto è sviluppata nella Tavola 12.1.

In conclusione, questo lavoro fornisce una parziale valutazione di alcune implicazioni fiscali che potrebbe aver avuto una equiparazione normativa della cannabis alle sigarette in Italia nel biennio 2007-2008. Come accennato nell'introduzione, la scelta dell'operatore pubblico sulla desiderabilità sociale della legalizzazione della cannabis non può ridursi alla mera opportunità fiscale di tale provvedimento. L'argomento fiscale entra nella funzione obiettivo dell'operatore pubblico insieme ad altri fattori extra-fiscali, il cui peso dipende dal sistema di valori diffuso nella società. Ad esempio, la scelta tra i due scenari qui proposti, quello di “sussistenza” e quello “commerciale”, dipende dalla metrica adottata dall'operatore pubblico. Un Paese con i conti pubblici in ordine, fortemente impegnato in una campagna di prevenzione dei danni derivanti dal fumo, darebbe più importanza alle negative implicazioni sanitarie della maggiore diffusione dei consumi di cannabis piuttosto che alle maggiori entrate fiscali, e pertanto preferirebbe lo scenario di “sussistenza” rispetto a quello “commerciale”. Viceversa, un Paese con gravi problemi di finanza pubblica potrebbe preferire l'instaurazione dello scenario “commerciale”, che implicherebbe un maggiore aumento delle entrate fiscali a scapito di un incremento nella diffusione dei consumi di cannabis.

APPENDICI

A.1 Olanda

Nel 1976 l'Olanda, sulla base delle indicazioni fornite dalle Commissioni Baan e Huilsman, adottò formalmente una politica di non persecuzione del possesso/consumo di cannabis. Questa espressione di *gedoogcultuur* ("cultura della tolleranza") era finalizzata alla separazione dei mercati delle droghe: cioè aveva l'intento di far uscire i consumatori di droghe definite "leggere" (tra cui la cannabis) dal mercato nero. Inizialmente la vendita di cannabis era concessa solo in sedi prive di finalità commerciali (associazioni giovanili, ecc.), successivamente, negli anni Ottanta, fu concessa l'apertura di *coffee-shops* tramite un sistema di licenze municipali. Questa decisione portò ad una rapida diffusione di questi punti di vendita, che raggiunsero il loro numero massimo di 1.179 unità nel 1997. Le preoccupazioni per questa evoluzione commerciale del fenomeno hanno indotto una politica restrittiva, che ha ridotto a circa 700 i *coffee-shops* attualmente aperti in Olanda, in cui sono impiegati circa 3.400 lavoratori (Bieleman *et al.*, 2009). Nel lungo termine la depenalizzazione del consumo di cannabis in Olanda non ha indotto un aumento permanente della diffusione del consumo di cannabis tra gli olandesi: «When we look to data for all adults, however, it is much harder to see any evidence that the Dutch system produces an increase in the intensity or duration of using careers» (MacCoun, 2010, p. 8). Questa indipendenza tra livello del consumo e *status* legale è confermata dall'analisi comparata di Reinermann *et al.* (2004), i quali mostrano che la distribuzione del consumo individuale è sostanzialmente uguale sia per la popolazione dei consumatori di Amsterdam (dove il consumo è tollerato) sia per quella di San Francisco (dove è vietato). In particolare, il consumatore olandese pare avere una "carriera" precoce ma anche più breve rispetto ai consumatori di cannabis americani (MacCoun, 2010a, Graf. 7). Infine, comparando il prezzo della cannabis in Olanda con quelli di altri paesi, si può affermare che la depenalizzazione ha avuto un impatto trascurabile se non nullo sul prezzo al dettaglio della cannabis (MacCoun, 2010a). L'analisi temporale dell'andamento della diffusione del consumo di cannabis in Olanda svolta da MacCoun e Reuter (1997) mostra che la depenalizzazione del 1976 non ha influenzato la diffusione del consumo di cannabis in Olanda nei successivi sei/sette anni. Tuttavia, quando, negli anni Ottanta è cresciuta la diffusione dei *coffee-shops*, la frequenza di consumo mensile tra la popolazione olandese adulta è invece aumentata dall'8,5% del 1984 all'11,5% del 1992, ma è poi diminuita fino al 5,4% del 2005: «We hy-

pothesize that the dramatic mid-1980s escalation in Dutch cannabis use is the consequence of the gradual progression from a passive depenalization regime to the broader *de facto* legalization which allowed for greater access and increasing levels of promotion, at least until 1995 when the policy was revised» (MacCoun, 2010a, pag. 17). Secondo questa tesi, negli anni Novanta in Olanda, ed in particolare ad Amsterdam, abbiamo assistito ad un fenomeno di “commercializzazione” della cannabis. La depenalizzazione della cannabis, che originariamente era finalizzata alla separazione dei mercati delle droghe, è servita invece per lo sviluppo di una rete commerciale a fini di profitto privato. La crescita del numero dei *coffee-shops* e l’avvio di una loro aggressiva attività promozionale sarebbe quindi responsabile della diffusione del consumo di cannabis registrata tra la popolazione olandese negli anni Novanta. A conferma di questa tesi, quando, nella decade successiva, sono state imposte limitazioni alla proliferazione/localizzazione/promozione dei punti di vendita di cannabis, la diffusione del consumo tra la popolazione olandese è tornata ai suoi livelli pre-commercializzazione.

A.2 Australia ed Alaska

Quattro stati della confederazione australiana hanno depenalizzato la coltivazione domestica di cannabis: Australia del Sud (1987), il Territorio della Capitale (1992), il Territorio del Nord (1996) e l’Australia Occidentale (2004). In questi casi la vendita di cannabis rimane vietata mentre l’autoproduzione è appunto finalizzata alla prevenzione dello sviluppo di un mercato a fini commerciali. L’analisi comparata svolta da Donnelly *et al.* (2000) indica che, nel periodo dal 1985 al 1995, l’andamento della diffusione del consumo di cannabis nell’Australia del Sud sia stato in linea con quello degli altri Stati australiani che non avevano dotato simili politiche di depenalizzazione della coltivazione. Risultati simili, cioè effetti nulli della depenalizzazione sulla diffusione del consumo, sono stati registrati nel caso del Territorio della Capitale (McGeorge e Aitken, 1997) e dell’Australia Occidentale (Fetherson e Lenton, 2007). Williams (2004), analizzando questi dati, conclude: «No evidence is found that either participation or frequency of use is sensitive to the criminal status of marijuana». In due occasioni in anche in Alaska la coltivazione domestica di cannabis è stata decriminalizzata. L’analisi del secondo episodio (2003) mostra che, malgrado la depenalizzazione, in Alaska la diffusione del consumo di cannabis con frequenza mensile si è ridotto da circa il 10% della popolazione nel 2003 a circa l’8% nel 2007, mentre nel resto degli Usa è rimasto costante al 6% (MacCoun, 2010b, Graf. 2).

APPENDICE STATISTICA

TAV. 2.2.1

OPERAZIONI DI POLIZIA CONDOTTE IN ITALIA

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Totale Antidroga (A)	22.177	21.954	21.651	20.703	18.493	18.744	19.872	20.768	22.111	22.470
di cui: Hashish (H)	7.539	7.746	7.633	6.746	6.196	6.390	6.513	6.378	6.839	7.065
di cui: Marijuana (M)	3.492	3.376	3.564	3.330	2.311	1.642	1.852	2.517	2.208	1.942
Totale Cannabis (C= H+M)	11.031	11.122	11.197	10.076	8.507	8.032	8.365	8.895	9.047	9.007
Cannabis su totale (C/A)	0,50	0,51	0,52	0,49	0,46	0,43	0,42	0,43	0,41	0,40

Fonte: DCSA, *Relazione Annuale*, anni vari.

TAV. 2.2.2

CANNABIS SEQUESTRATA IN ITALIA

	2004	2005	2006	2007	2008	Totale	Media
Hashish (kg.)	16.423	23.199	19.728	20.064	34.107	113.521	22.704
Marijuana (kg.)	3.491	2.485	4.985	4.549	2.380	17.890	3.578
Piante (nr.)	311.124	140.956	95.368	1.524.248	148.152	2.219.848	443.970

Fonti: OEDT (2010), *National Report 2009: Italy Part B: Selected Issues - 11.2 Seizures - tab. 11.8*, e DCSA, *Relazione Annuale*, anni vari.

TAV. 2.2.3

SEGNALAZIONI ALL'A.G. PER VIOLAZIONI DELLA NORMATIVA SULLA CANNABIS

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Totale Antidroga (A)	34.396	34.367	34.133	33.184	29.584	31.483	31.605	33.056	33.451	35.097
di cui: Hashish (H)	10.279	10.673	10.631	9.725	8.775	8.845	8.300	8.528	8.999	9.327
di cui: Marijuana (M)	4.539	3.872	4.199	3.648	2.954	1.783	1.793	2.746	2.422	2.266
Totale Cannabis (C= H+M)	14.818	14.545	14.830	13.373	11.729	10.628	10.093	11.274	11.421	11.593
Cannabis su totale (C/A)	0,43	0,42	0,43	0,40	0,40	0,34	0,32	0,34	0,34	0,33

Fonte: DCSA, *Relazione Annuale*, anni vari.

TAV. 2.2.4

DETENUTI PER VIOLAZIONE DELLA NORMATIVA SULLA CANNABIS

	2004	2005	2007	2008
Det. ex art. 73*	21.730	23.156	18.222	21.037
Det. per cannabis	7.336	7.393	6.221	6.949

* Fonte: SCANDURRA A. (2008, tab. 04).

TAV. 3.1

DIFFUSIONE DEL CONSUMO ANNUALE DI CANNABIS
TRA LA POPOLAZIONE ITALIANA

Anno di rilevazione	Classe di età	Percentuale (stima Unodoc)	Percentuale (stima Ipsad)	Pop. Ital. 15-64 (migliaia)	Consumatori (migliaia)
2008	15-64	14,6	n.d.	39.906	5.826
2007	15-64	n.d.	14,0	39.017	5.462
2005	15-64	11,2	11,9	38.827	4.485

Fonti: UNODOC, WDR 2005, tab. 8.1.3; WDR 2006 tab. 6.1.3; WDR 2007/9, tab. 3.5.1.3. WDR 2010, tab. 4.2.1.3; OEDT (2007), Country overview - Italy - Key Statistics; OEDT (2009), National Report 2009: Italy Part B: Selected Issues - 11.1.2 Cannabis use; ISTAT, *Demografia in cifre* (<http://demo.istat.it>).

TAV. 3.2

PREZZI TIPICI AL DETTAGLIO REGISTRATI NEL MERCATO ILLEGALE DELLA
CANNABIS IN ITALIA (€/Gr.)

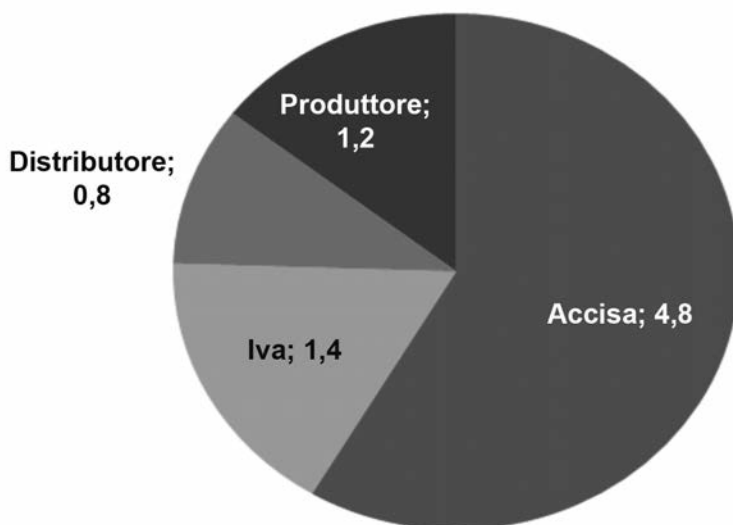
Anno	Marijuana	Hashish	Cannabis
2007	6,6	8,4	8,2
2008	7,2	8,8	8,6

Fonte: OEDT, *National Report 2009: Italy Part B: Selected Issues - 11. Cannabis Market and Production*.

Nota: Prezzo tipico della cannabis calcolato come media ponderata tra prezzo tipico dell'hashish e della marijuana, con peso uguale alla percentuale che hashish o marijuana hanno avuto sul totale dei sequestri di cannabis (piante escluse).

GRAF. 6.1.1

RIPARTIZIONE DEGLI INCASSI SULLA VENDITA DI CANNABIS



TAV. 12.1

PRINCIPALI IMPLICAZIONI FISCALI DEGLI SCENARI IPOTIZZATI

Conseguenze fiscali su:	$Q_L = Q_B$	$Q_L = 4/3 Q_B$
Imposte sulle vendite*	≈ 7 miliardi di euro	≈ 10 miliardi**
Imposte sul reddito*	≈ 530 milioni di euro	≈ 3 miliardi**
Riduzione della criminalità	Aumento delle entrate	Aumento delle entrate
Applicazione della normativa di cui carceri	Riduzione della spesa ≈ 836 milioni di euro	Riduzione della spesa ≈ 836 milioni di euro
Costi sanitari	Riduzione della spesa	Aumento della spesa
Produttività dei consumatori	Aumento delle entrate	Riduzione delle entrate

Legenda: Q_B = stima della quantità di cannabis consumata illegalmente; Q_L = quantità di cannabis che sarebbe stata consumata legalmente; Scenario di "sussistenza": $Q_L = Q_B$; Scenario "commerciale": $Q_L = 4/3 Q_B$.

Note: *Escluse le entrate da imposte sull'indotto; **Escluso il "turismo dello spinello".

BIBLIOGRAFIA

- BECKER G.S. - GROSSMAN M. - MURPHY K.M., «The Market for Illegal Goods: The Case of Drugs», *Journal of Political Economy*, vol. 114, no. 1, 2006, pages 38-60.
- BIELEMAN B.B. - NIJKAMP A.R. - DE BIE E., «Coffee-Shops in Nederland 2007», *Groningen Intraval*, 2009.
- BOUCHARD M., «Towards a Realistic Method to Estimate the Cannabis Production in Industrialized Countries», *Contemporary Drugs Problems*, vol. 35, July 2008, pages 291-300.
- CAULKINS J.P., «Estimated Cost of Production for Legalized Cannabis», Rand Drug Policy Research Center, *WR-764-C*, (www.rand.org), 2010.
- CENTRO STUDI RISTRETTI ORIZZONTI, <http://www.ristretti.it/areastudio/statistiche/index.htm>, 2009.
- DONNELLY N. - HALL W. - CHRISTIE P., «The Effects of the Cannabis Expiation Notice Scheme on Levels and Patterns of Cannabis Use in South Australia: Evidence from the National Drug Strategy Household Surveys 1985-1995», *Drug and Alcohol Review*, no. 19(3), 2000, pages 265-269.
- DCSA (Direzione Centrale Servizi Antidroga), *Relazione Annuale*, Ministero degli Interni, Roma, anni 2006-2010.
- FETHERSTON J. - LENTON S., *A Pre-Post Comparison of the Impact of the Western Australia Cannabis Infringement Note on Public Attitudes, Knowledge and Use*, National Drug Research Institute, Curtin University of Technology, Perth, 2007.
- GROVER A., *Report of the Special Rapporteur on the Right of Everyone to the Enjoyment of the Highest Attainable Standard of Physical and Mental Health*, United Nations, A/65/225, August 6 2010, New York.
- HALL W. - DAGENHARDT L., «Adverse Health Effects of Non-Medical Cannabis Use», *Lancet*, no. 374, 2009, pages 1383-1391.
- IPSA (Italian Population Survey on Alcohol and other Drugs), Consiglio Nazionale delle Ricerche, Istituto di Fisiologia Clinica, Sezione di Epidemiologia e Ricerca sui Servizi Sanitari, (http://www.epid.ifc.cnr.it/index.php?option=com_content&view=article&id=63&Itemid=78).
- KILMER B. - PACULA R.L., «Estimating the Size of the Global Drug Market: A Demand Side Approach - report 2», Rand Drug Policy Research Center - *Technical Report 711-EC*, (www.rand.org), 2009.
- KILMER B. - CAULKINS J.P. - PACULA R.L. - MACCOUN R.J. - REUTER P.H., «Altered State? How Marijuana Legalization in California Could Influence Marijuana Consumption and Public Budgets», Rand Drug Policy Research Center - *Occasional Paper*, (www.rand.org), 2010.
- MACCOUN R.J., «Drugs and the Law: A Psychological Analysis of Drug Prohibition», *Psychological Bulletin*, vol. 113, no. 3, 1993, pages 497-512.

- MACCOUN R.J., «What Can We Learn from the Dutch Cannabis Coffee-Shops Experience?», Rand Drug Policy Research Center - WR-768-RC, (www.rand.org), 2010a.
- .-, «Estimating the Non-Price Effects of Legalization on Cannabis consumption», Rand Drug Policy Research Center - WR-767-RC, (www.rand.org), 2010b.
- MACCOUN R.J. - REUTER P.H., «Interpreting Dutch Cannabis Policy: Reasoning by Analogy in the Legalization Debate», *Science*, vol. 278, Ottobre 1997.
- MCGEORGE J. - AITKEN C.K. «Effects of Cannabis Decriminalization in the Australian Capital Territory on University Students' Patterns of Use», *Journal of Drug Issues*, no. 27, 1997, pages 785-793.
- MENEGHINI G., «Droga, narrazioni, stereotipi. Retorica dell'identità subordinata», *I Fogli di Oriss*, nn. 1/2, 1994, pp. 79-98.
- MIRON J.A., «Drugs Prohibition», in NEWMAN P. (ed.), *The New Palgrave of Law and Economics*, MacMillan, London, 1997, pages 648-652.
- OEDT (Osservatorio Europeo delle Droghe e delle Tossicodipendenze), *Relazione annuale*, Unione Europea, Lisbona, 2007.
- .-, (Osservatorio Europeo delle Droghe e delle Tossicodipendenze), *Bollettino statistico*, Unione Europea, Lisbona, 2007.
- .-, (Osservatorio Europeo delle Droghe e delle Tossicodipendenze), *National Report 2009: Italy*, National Focal Point, Roma, 2010.
- PACULA R.L. «Examining the Impact of Marijuana Legalization on Marijuana Consumption: Insights from the Economics Literature», Rand Drug Policy Research Center - WR-770-RC, (www.rand.org), 2010.
- PACULA R.L. - GROSSMAN M. - CHALOUKPA F.J. - O'MALLEY P.M. - JOHNSTON L.D. - FARRELLY M.C., «Drug and Youth», Cambridge (MA), *NBER Working Paper*, no. 7703, 2000.
- PACULA R.L. - HOORENS S. - KILMER B. - REUTER P.H. - BURGDORF J.R. - HUNT P. «Issues in Estimating the Economic Cost of Drug Abuse in Consuming Nations», Rand Europe - *Technical Report # 3*, prepared for the European Commission, (www.rand.org), 2009.
- PACULA R.L. - KILMER B. - GROSSMAN M. - CHALOUKPA F.J., «Risks and Prices: The Role of User Sanctions in Marijuana Markets», Cambridge (MA), *NBER Working Paper*, no. 13415, 2007.
- REINERMANN C. - COHEN P. - KAAL H. «The Limited Relevance of Drug Policy: Cannabis in Amsterdam and in San Francisco», *American Journal of Public Health*, no. 94(5), 2004, pages 836-842.
- REUTER P.H., *Disorganized Crime*, MIT Press, Cambridge, 1983.
- .-, «Marijuana Legalization: What Can Be Learned from Other Countries», Rand Drug Policy Research Center, *Working Paper # WR-771-RC*, (www.rand.org), July 2010.
- SCANDURRA A., «Tre anni di applicazione della legge Fini-Giovanardi: la parola ai dati», in CORLEONE F. - GONNELLA P. - ZUFFA G. (A CURA DI), *Libro bianco sulla Fini - Giovanardi*, *I Quaderni di Fuoriluogo*, (<http://www.fuoriluogo.it>), 2008.

- SCHAFFER C. - PAOLI L., *Drogen und Strafverfolgung*, Freiburg, Max Planck Institut für Ausländisches Internationales Strafrecht, 2006.
- SLACK A. - O'DEA D. - SHEERIN I. - NANA G.- WU J. - NORMAN D., «New Zealand Drug Harm Index: Report to the New Zealand Police», Wellington, Business and Economics Research Ltd., (<http://www.berl.co.nz/754a1.page>), 2008.
- UNODOC (United Nations, Office on Drugs and Crime), *Bulletin on Narcotics: Illicit Drug Markets*, Vienna, 2006.
- .-, (United Nations, Office on Drugs and Crime), *World Drug Report*, anni vari, Vienna.
- VAN LAAR M.G.C. - VAN GAGELDONK A. - VAN OUYEN-HOUBEN M. - CROES E. - MEIJER R. - KETELAARS T., *The Netherland Drug Situation 2009: Report to the EMCDDA*, Reritox National Focal Point, Timbos Instituut, Utrecht, 2010.
- WILLIAMS J., «The Effect of Price and Policy on Marijuana Use: What Can Be Learned from the Australian Experience?», *Health Economics*, no. 13, 2004, pages 123-137.
- WITTON J., «Cannabis Use and Physical and Mental Health», in *A cannabis Reader: Global Issues and Experiences*, Monographs series 8, vol. 2, cap. 6, EMCDDA, Lisbon, 2008, pages 115-140.

Disuguaglianze, beni relazionali ed elezioni nelle periferie di Roma

Federico Tomassi*

Ministero dello Sviluppo Economico, Roma

In the city of Rome economic and territorial polarizations increased between centre and peripheral belts, paralleling suburban sprawl. Electoral data 2005-2011 highlight sharp political polarization, too. Left-wing candidates prevail where social centrality exists, i.e. in dense zones with widespread social relationships and many public or collective places. On the contrary, right-wing parties prevail in far-off sprawled zones, with less opportunities to meet each other, where relational goods are less diffused. Since such goods, according to civil economics, are key factor in the individual well-being, economic and territorial differences lead to different political choices.

[JEL Classification: H41; R14; Z13].

Keywords: commons; relational goods; social capital; urban sprawl; elections; Rome.

* <federico.tomassi@uniroma1.it>, Unità di valutazione degli investimenti pubblici (UVAL-DPS). L'autore ringrazia per l'ispirazione e i consigli Paolo Leon, Laura Pennacchi e Walter Tocci, insieme con i partecipanti ai seminari della Fondazione Basso sui beni comuni e al Laboratorio Roma del CRS, nonché i *referee* anonimi per gli utili commenti. È inoltre debitore nei confronti di Rossana Rosati dell'Ufficio statistico di Roma Capitale e di Paolo Iannini di Provinciattiva per la loro disponibilità nel fornire alcuni dei dati necessari al lavoro. L'autore rimane l'unico responsabile di eventuali errori e imprecisioni, e inoltre quanto scritto non impegna l'Amministrazione di appartenenza o gli Enti citati.

1. - Introduzione: economia, urbanistica e politica

Nell'Italia di oggi è diffusa la percezione di “malessere sociale” come risultato di tre processi paralleli: l'insicurezza generata dalla volatilità dei redditi, la segmentazione sociale dovuta alla polarizzazione dei redditi stessi, il freno alla mobilità sociale causato dalla persistenza intergenerazionale delle disuguaglianze¹. Il “malessere” colpisce inevitabilmente anche Roma – in termini economici e di reddito, socio-culturali e di opportunità e, infine, politici ed elettorali – nonostante che la Capitale sia stata celebrata a lungo come un esempio economico di successo.

In effetti durante le giunte di centrosinistra dal 1993 al 2008 (prima con Francesco Rutelli e poi con Walter Veltroni), il cosiddetto “modello romano”² ha registrato risultati positivi per quanto riguarda la crescita del Pil, la ricchezza pro capite e le presenze turistiche. Si è tuttavia allargata la disuguaglianza nei redditi e nelle condizioni economiche, tra chi ha un reddito adeguato e può godersi le molte opportunità di consumo e divertimento che fornisce la città, e chi subisce la crisi di molti settori industriali, la sotto-occupazione e la precarietà, lo scarto tra il proprio titolo di studio e la bassa qualità dei lavori. Di sicuro, la crescita economica che pure c'è stata non si è tuttavia ripartita in maniera sufficientemente omogenea sul territorio, creando diffuse situazioni di disagio.

Negli stessi anni è andato progressivamente aggravandosi un altro tipo di polarizzazione, questa volta in ambito urbanistico. Con l'espandersi dello *sprawl* urbano, ossia la dispersione insediativa causata dalla crescita disordinata delle periferie urbane (analogamente alle altre grandi metropoli ma anche a una regione policentrica come il Veneto), si crea infatti una nuova disuguaglianza tra chi ha la disponibilità di beni comuni, collettivi o relazionali intorno a sé e chi non ce l'ha, ossia tra i quartieri dove è frequente incontrarsi e interagire con gli altri abitanti e le zone dove al contrario le famiglie vivono maggiormente isolate dal contesto urbano.

Si intendono qui come beni comuni, nell'accezione tradizionale dei *commons*, o collettivi secondo la letteratura più recente³, le reti umane che hanno carattere virtuale o artificiale – come i saperi, la fiducia, le tecnologie – ma che si sviluppano a livello locale⁴. Sono tali i contatti e le relazioni urbane create e riprodotte su un

¹ FRANZINI M. (2007), pp. 41-42).

² VELTRONI W. (2006).

³ OSTROM E. (1990).

⁴ DONOLO C. (2005).

territorio specifico dall'interazione tra ecosistema sociale ed ecosistema naturale. Secondo il nuovo filone di ricerca dell'economia civile⁵, più che gli scambi di cose, sono proprio gli incontri tra le persone a determinare non solo il benessere degli individui, ma anche lo sviluppo locale e la competitività territoriale, mediante sia il consumo di beni relazionali che rappresentano l'output dell'incontro sia i processi di apprendimento collettivo che ne derivano. Nella stessa ottica, Zamagni (2005a, pp. 71-72) sottolinea che «È ormai ampiamente ammesso che lo star bene (*well-being*) delle persone dipende non solamente dal soddisfacimento dei bisogni materiali, ma anche da quello dei bisogni relazionali. [...Essi] richiedono beni relazionali, beni cioè la cui utilità per il soggetto che li consuma dipende, oltre che dalle loro caratteristiche intrinseche e oggettive, dalle modalità di fruizione con altri soggetti. Relazionale è il bene che può essere prodotto e fruito soltanto assieme da coloro i quali ne sono, appunto, gli stessi produttori e fruitori mediante le relazioni che li coinvolgono».

Ma le reti sociali non sono diffuse equamente nei diversi quartieri, e di conseguenza non lo sono le opportunità di incontro tra persone che permettono la produzione e il contestuale consumo di beni relazionali. In effetti, le reti sociali sembrano presenti con maggiore intensità nei quartieri più vecchi, solitamente densi e intensivi, dove le relazioni interpersonali sono forti e le possibilità di partecipazione più numerose. Al contrario le periferie estreme e più recenti sono lontane dai grandi eventi e dai grandi progetti, poco fornite di servizi di base (tali da costringere a un utilizzo forzato dell'automobile), vincolate a una scarsa presenza di spazi collettivi che riducono le occasioni di contatto e privano di identità collettive e storicizzate. Le due polarizzazioni – quella economica e quella urbanistica – incidono sui risultati delle elezioni, il cui dato medio romano nasconde grandi differenze nei vari quartieri, determinando così una terza dicotomia, di tipo elettorale.

Questo saggio intende contribuire al dibattito sull'economia civile analizzando le relazioni tra la polarizzazione elettorale e quella economico-urbanistica grazie a un *dataset* appositamente realizzato. L'obiettivo è quello di individuare gli elementi di fondo che contribuiscono a spiegare le differenze dei quartieri romani rispetto alla media cittadina, differenze che sembrano permanere in ogni elezione, indipendentemente dal suo esito. Lo scopo non è infatti quello di analizzare il valore assoluto dei voti, e quindi i motivi perché il centrosinistra o il centrodestra vincono o perdono la singola elezione, che dipendono molto dal candidato scelto

⁵ BRUNI L. - ZAMAGNI S. (2004); CALAFATI A.G. (2002); GUI B. (2002); ZAMAGNI S. (2005a; 2005b).

dallo schieramento e dal contesto politico nazionale, e che variano continuamente nelle diverse elezioni.

Il punto di partenza è che il centrosinistra mostra risultati sistematicamente peggiori nei quartieri più esterni, dove è il centrodestra a prevalere, mentre ottiene più voti in quartieri centrali e semicentrali. Quella che in passato era definita la “cintura rossa” nella periferia oggi non esiste più⁶, tanto che Siti (2009) afferma: «Non riesco a immaginare un borgataro riformista». Ma secondo Tocci (2008, p. 95) è già dalla metà degli anni '80 che «la mitica cintura rossa si spezza in due», con la periferia storica che va a costituire la base sociale delle giunte di centrosinistra e con i quartieri più lontani intorno al Grande Raccordo Anulare (Gra) in cui comincia a prevalere il centrodestra, ulteriormente rafforzato poi negli anni '90 dai conflitti legati all'immigrazione.

Gli elementi di originalità del presente lavoro risiedono sia nell'interdisciplinarietà della trattazione che nell'oggetto stesso dell'analisi. In primo luogo, vi è il tentativo di mettere in relazione metodologie e concetti provenienti da ambiti teorici differenti: dal lato dell'economia l'impatto delle disuguaglianze sociali e di reddito sul benessere degli individui, dal lato della sociologia la categoria dei beni comuni o relazionali applicata alle dinamiche urbane, dal lato della politologia e della statistica le analisi elettorali. Ma in secondo luogo è il dualismo elettorale tra centro e periferie a essere con scarsa frequenza oggetto di attenzione: se abbondano ovviamente le analisi sul comportamento degli elettori nelle varie votazioni e sui fattori che portano alla vittoria uno schieramento o un candidato, raramente il focus è posto sugli elementi di fondo del voto che permangono costanti in elezioni il cui esito finale invece diverge. Sarà infatti evidente come, nonostante in valore assoluto le differenze tra un'elezione e l'altra possano essere molto marcate, i quartieri dove il centrosinistra e il centrodestra ottengono i migliori e i peggiori risultati sono quasi sempre gli stessi, a denotare una differenza rispetto alla media cittadina abbastanza stabile nel corso del tempo. Sicuramente inconsueta è infine la scala adottata per l'analisi, ossia le zone urbanistiche romane, che assicurano una sufficiente numerosità per ottenere risultati statistici robusti, ma anche un'adeguata omogeneità al loro interno per distinguere efficacemente tra il centro e le diverse fasce periferiche. A tale scopo è stato realizzato un *dataset* a livello di zona urbanistica che combina dati di diversa fonte: i risultati elettorali per sezione, l'anagrafe comunale, le caratteristiche socio-economiche rilevate con

⁶ FERRAROTTI F. - MACIOTTI M.I. (2009).

il censimento e un paio di indagini specifiche sia sul reddito che sulla dotazione di servizi e spazi pubblici.

Il saggio inizia con la descrizione delle differenze tra il centro di Roma e le fasce periferiche sia in termini di polarizzazione economica e urbanistica (paragrafo 2), sia con riferimento ai dati elettorali dal 1997 al 2011 delle zone urbanistiche (paragrafo 3). Su questa base i principali fattori socio-economici, demografici e urbanistici vengono discussi come motivo delle differenze elettorali tra i vari quartieri (paragrafo 4), e messi in relazione con la scarsa disponibilità di beni comuni e relazionali nelle periferie, e con la conseguente limitata interazione sociale e politica che avviene nei quartieri più lontani (paragrafo 5). Segue la presentazione e la discussione dei risultati prima di un'analisi esplorativa per individuare eventuali fattori latenti, e poi di alcune regressioni dei dati elettorali a partire dal 2005 rispetto ai fattori di contesto come variabili esplicative (paragrafo 6). Infine vengono proposte alcune osservazioni sintetiche sui motivi per i quali la carenza e lo scarso uso dei beni comuni e relazionali nelle periferie sembra danneggiare più il centrosinistra che il centrodestra (paragrafo 7).

2. - Le periferie romane

Negli anni del cosiddetto “modello romano” la crescita non si è ripartita in maniera sufficientemente omogenea sul territorio, anzi in alcuni casi ha aggravato gli squilibri esistenti nella ripartizione sociale della ricchezza e delle opportunità⁷. Il Lazio ha uno dei più alti livelli di concentrazione del reddito in Italia (l'indice di Gini era nel 2008 pari a 0,324), inferiore solo a Campania e Sicilia, a mostrare una disuguaglianza distributiva elevata e insolita per una regione con un reddito medio elevato, tanto che l'indicatore Eurostat di deprivazione arriva a 16,8 contro la media nazionale di 15,2 e la quota di abitanti che non riesce ad affrontare una spesa imprevista di 750 euro è del 38,8% contro il 33,3% medio italiano⁸. Negli ultimi anni la crescita è peraltro rallentata, con un aumento del reddito medio familiare nella provincia di Roma solo del 7,5% tra il 2005 e il 2009 a fronte di un incremento medio nazionale del 9,4%, e più in particolare una caduta del reddito romano dello 0,8% a seguito della crisi economica tra 2008 e 2009⁹.

⁷ CASTRONOVI A. (2008); DE MURO P. *et AL.* (2008).

⁸ ISTAT (2010).

⁹ EURES (2010, cap. 7).

Un maggiore dettaglio dimostra che i risultati degli anni di crescita non sono comunque uniformi nei 19 Municipi in cui è suddivisa amministrativamente Roma (Tav. 1). Le differenze nel reddito pro capite sono notevoli tra il massimo del I, II e XVII Municipio (tra 25.000 e 28.600 euro nel 2003) e il minimo dell'VIII e delle aree del V e X esterne al Gra (circa 14.000 euro). Anche l'indice di capitale economico – che comprende proprietà delle case, valori immobiliari e tassi di occupazione e disoccupazione – registra il massimo in I, II, III, XII e XVII Municipio e il minimo nell'VIII. Infine, l'indice di disagio socio-economico costruito dal Censis sulla base di un'indagine campionaria multidimensionale varia tra i valori migliori in II e III (0,30-0,35) e i peggiori in VII, VIII e XV (tra 0,65 e 0,74).

TAV. 1

DIFFERENZE SOCIO-ECONOMICHE NEI MUNICIPI DI ROMA

Municipi	Reddito pro capite	Euro 2003	Capitale economico	Indice 2009	Disagio socio-economico	Indice 2002
Migliori 5	II (Pinciano e Parioli)	28.641	I	1,73	II	0,299
	II (Salario e Trieste)	27.615	II	1,35	III	0,353
	I (Centro storico)	25.067	XVII	1,17	IX	0,408
	II-XVII (Flaminio e Della Vittoria)	24.895	XII	0,93	XI	0,444
	XII (Eur e Torrino)	24.861	III	0,63	XII	0,456
Media	Roma	19.615	Roma	0,00	Roma	0,526
Peggiori 5	XV-XVI (Castel di Guido e Ponte Galeria)	14.913	IV	-0,74	VI	0,614
	VIII (Lunghezza, Acqua Vergine, San Vittorino)	14.293	V	-0,78	XIII	0,620
	V (Settecamini)	13.985	XV	-0,91	VII	0,652
	VIII-X (Torrenova, T. Angela, T. Gaia, Morena)	13.950	XX	-0,93	XV	0,671
	VIII (Borghesiana)	13.949	VIII	-2,40	VIII	0,736

Fonte: ISTITUTO TAGLIACARNE (2005) per il reddito *pro capite*; BONCIANI M. (2009, pp. 108-117) per il capitale economico e CENSIS (2002) per il disagio socio-economico.

Negli stessi anni Roma ha vissuto anche una continua valorizzazione immobiliare dei patrimoni di grandi imprese ed enti previdenziali, e dei suoli pregiati dell'Agro romano, che non ha comportato solo un aumento delle rendite private¹⁰. A fronte di una popolazione residente a Roma stabile da più di un decennio

¹⁰ TOCCI W. (2009).

intorno ai 2,7 milioni di abitanti¹¹, l'impennata nel costo delle case e degli affitti ha provocato soprattutto il progressivo spostamento degli abitanti verso la periferia esterna al Gra e i comuni della Provincia, per un totale di circa 200mila abitanti che si sono trasferiti fuori dal Gra tra il 1991 e il 2007¹². Alla base di tale dinamica vi sono diversi processi che operano contemporaneamente: l'espulsione delle classi medie dalla città consolidata per l'impossibilità di sostenere i prezzi di mercato; la "vetrinizzazione" del centro a vantaggio delle attività commerciali, turistiche e culturali; la dismissione del patrimonio pubblico e degli enti previdenziali; la domanda crescente da parte di soggetti con limitata capacità di spesa quali immigrati, single, giovani coppie, separati e divorziati¹³. I trasferimenti di residenza all'interno della città tra diverse zone urbanistiche mostrano che il saldo positivo più elevato – quantomeno nel periodo tra il 2003 e il 2007 – è caratteristico di tutte le aree esterne al Gra, mentre continuano a perdere abitanti il centro storico e i quartieri a sud e a est¹⁴. Poiché i posti di lavoro rimangono per lo più localizzati nelle aree centrali, le traiettorie dei trasferimenti di residenza e degli spostamenti pendolari risultano avere direzioni opposte: i romani traslocano verso le periferie ma poi tornano ogni giorno verso il centro per lavorare.

L'espansione edilizia fa sì che le differenze socio-economiche non riguardino più il tradizionale dualismo dell'urbanistica romana tra centro e periferia¹⁵, visto che nel primo – definito nella maniera più estensiva possibile, non solo storico-geografica – rimangono 400mila persone mentre l'85% dei residenti vive altrove. Bisogna ormai ragionare di "periferie" al plurale, distinguendo almeno tre fasi che hanno lasciato impronte molto diverse nella struttura della città: la periferia storica, anulare e regionale¹⁶. La prima espansione nella periferia storica segue l'enorme incremento di popolazione a partire dagli anni '20 fino alla metà degli anni '70 ed è guidata dalle borgate, nate isolate dal resto della città e poi man mano raggiunte

¹¹ Secondo i dati Istat, applicando agli anni precedenti i risultati del confronto tra Censimento e Anagrafe completato nel 2007 (stime dell'autore), la popolazione residente media annua nel Comune ha oscillato dai 2.709mila nel 2000 al minimo di 2.686mila nel 2003 per poi risalire a 2.753mila nel 2010; il totale della Provincia è invece continuamente cresciuto da 3.849mila nel 2001 a 4.174mila nel 2010. Il lavoro di ricostruzione compiuto da SONNINO E. *et AL.* (2010, cap. II) stima infatti una popolazione "virtuale" di 2.754mila abitanti a fine 2008. Secondo i dati anagrafici del Comune, nel periodo considerato i residenti sono invece circa 100-120mila in più, per un ammontare pari a 2.882mila a fine 2010.

¹² LUCCIARINI S. - VIOLANTE A. (2006); CAUDO G. *et AL.* (2009); CAUDO G. (2011).

¹³ LUCCIARINI S. - VIOLANTE A. (2007, par. 6).

¹⁴ CRISCI M. (2010).

¹⁵ VIOLANTE A. (2008, pp. 37-59).

¹⁶ TOCCI W. (2008, pp. 74-109).

dai nuovi quartieri, in un intreccio perverso di abusivismo irresponsabile e pianificazione mancata. La seconda espansione nella periferia anulare, molto più rapida, impegna tra il 1975 e il 1995 una fascia immensa che va dalla circonvallazione alle aree a cavallo del Gra, raddoppiando la superficie urbanizzata pur in presenza di una popolazione stabile, con una forma insediativa prevalentemente granulare. Negli anni 90 riparte un nuovo ciclo edilizio che arriva a superare i confini comunali già molto estesi e ad interessare la prima e seconda cintura dei comuni della Provincia: una periferia regionale ancora più sfilacciata, con nuovi quartieri circondati da terreni agricoli e isolati fisicamente dal resto della città, ossia per usare le parole di Berdini (2008, pp. 165-187), una «sconfinata periferia metropolitana». Tocci (2008, p. 73) rappresenta queste nuove periferie come «un tessuto slabbrato costituito da tante piccole monadi urbane collocate in modo caotico nello spazio geografico, senza alcun rispetto delle preesistenze culturali e naturali, con densità locali a volte anche alte, ma con una bassa densità a scala vasta».

Ai fini operativi del presente lavoro, intorno al nucleo di quartieri centrali in senso sia storico che socio-economico (i Municipi I, II, III e XVII più i quartieri Eur, Fleming e Ponte Milvio), sono state quindi considerate tre successive fasce: la periferia storica di edilizia intensiva e vecchie borgate ufficiali o abusive (l'area consolidata più o meno interna alla circonvallazione); la periferia anulare di costruzioni estensive con ampia dotazione di verde (il restante territorio interno al Gra, più i grandi quartieri dell'espansione verso il mare a Ostia, Acilia, Mezzocamino e Spinaceto); la periferia regionale con i rimanenti insediamenti isolati fuori dal Gra fino ai confini del Comune (Tav. 11 in Appendice).

Queste tre fasce mostrano una riduzione continua della densità di popolazione nel tessuto urbano, che scende da 93,3 abitanti per ettaro nella periferia storica a 28,8 nella periferia anulare e a Ostia, fino a 5,4 fuori dal Gra (Tav. 2). Se tale situazione appariva già evidente a metà degli anni '90¹⁷, la dinamica tra il 1997 e il 2007 mostra una ulteriore forte crescita della popolazione nella periferia regionale (+20,9%), un moderato aumento nella periferia anulare (+3,5%) e un significativo calo nella periferia storica (-7,1%). In maniera analoga, il rapporto tra volumi costruiti e superficie urbanizzata, che misura quanto sono estensivi i nuclei abitati, decresce da 4,5 m³/m² entro l'anello ferroviario a 2,3 entro il Gra, fino a 1-1,1 fuori dal Gra e nei Comuni di prima cintura, ma è addirittura pari a 0,7 nelle ex borgate abusive della periferia romana¹⁸.

¹⁷ MALTESE C. - PEREGO F. (1996, pp. 140-142).

¹⁸ AIC - UNIONE BORGATE (2010, p. 37).

TAV. 2

POPOLAZIONE E DENSITÀ NEL COMUNE DI ROMA, PER AREA URBANA

Area urbana	Superficie (ettari)	Residenti al 31.12.1997	Residenti al 31.12.2007	Variazione % 1997-2007	Densità al 31.12.2007
Centro	5.195	413.092	405.095	-1,9	78,0
Periferia storica	13.735	1.379.323	1.280.992	-7,1	93,3
Periferia anulare	21.849	607.866	629.158	3,5	28,8
Periferia extra-Gra	87.630	392.106	474.038	20,9	5,4
Totale	128.409	2.811.547	2.838.047	0,9	22,1

Nota: per poter effettuare il confronto tra 1997 e 2007, la tavola è costruita sulla base delle suddivisioni toponomastiche (Tav. 11 in Appendice) e non mediante le zone urbanistiche utilizzate nel paragrafo 3.

Fonte: elaborazione dell'autore su dati di Roma Capitale - Ufficio anagrafico.

3. - Le elezioni dal 1997 al 2011

Il dualismo economico ed urbanistico si riflette sulle scelte elettorali: in tutte le elezioni tra il 1997 e il 2011 vige una sorta di legge gravitazionale, in cui l'attrazione del Campidoglio ha un effetto opposto sulle due principali coalizioni man mano che ci si allontana verso le periferie. I voti ottenuti dal centrosinistra sono direttamente proporzionali alla vicinanza con il centro, con il massimo nella periferia storica, un valore intermedio nella periferia anulare, e il minimo nella periferia regionale. Viceversa, i voti per il centrodestra sono inversamente proporzionali alla vicinanza dal centro, con il massimo nella periferia regionale, un valore intermedio nella periferia anulare, e il minimo nella periferia storica. Nel centro invece entrambi gli schieramenti raggiungono livelli simili alla media romana, con alcune eccezioni. Ciò non vale solo per i candidati alle cariche amministrative monocratiche (Sindaco, Presidente della Provincia e della Regione), ma anche alle elezioni politiche e alle europee, sia considerando la somma delle liste delle due coalizioni per la parte proporzionale, sia esaminando i soli partiti principali (gli attuali PD e PdL), e persino per l'affluenza ai referendum del 2011, che hanno avuto una forte valenza politica più che civica, tanto da poter individuare le caratteristiche tipiche del voto al centrosinistra.

Considerazioni di questo tipo non sono nuove, tanto che subito dopo le elezioni è stato evidenziato dai mezzi di comunicazione¹⁹ come il maggiore successo di Alemanno e dei candidati presidenti municipali sia avvenuto nei Municipi più estesi in periferia. L'analisi dei singoli seggi, in particolare per le elezioni della Ca-

¹⁹ A titolo di esempio, si rimanda all'articolo di D'AVANZO G., «Cinecittà, così la borgata rossa ha voltato le spalle alla sinistra», in *La Repubblica*, 30 aprile 2008.

mera dei Deputati nel 2006²⁰, mostra la tendenza del centrosinistra a ottenere il maggiore consenso lungo una diagonale che va da sud-ovest (Ostia) a nord-est (Tiburtina), con diramazioni verso nord nei Municipi IV e XIX; al contrario, il centrodestra prevale in un'ampia corona intorno alla città, soprattutto fuori dal Gra, con una presenza marcata a nord-ovest nei Municipi II e XX.

Nel presente lavoro l'unità di analisi è rappresentata dalle suddivisioni dei Municipi nelle 143 zone urbanistiche (sul totale di 155) dove sono presenti seggi elettorali, aree sufficientemente piccole e omogenee, dal punto di vista sia storico-geografico che socio-economico (Graf. 8 in Appendice), per le quali sono disponibili numerosi dati statistici²¹. Poiché i candidati del centrosinistra hanno prevalso nel territorio comunale in tutte le competizioni elettorali fino al 2008, si è scelto di considerare come variabile di analisi il differenziale della percentuale dei voti per i candidati del centrosinistra nelle singole zone urbanistiche rispetto alla media romana²². Le elaborazioni sono perfettamente simmetriche per quanto riguarda il voto al centrodestra.

Alle ultime elezioni del 2010, quando alla Regione Lazio è stata eletta Renata Polverini del centrodestra contro Emma Bonino del centrosinistra, il raggruppamento delle zone urbanistiche di Roma in cinque classi di dimensione omogenea mostra chiaramente una riconoscibile trama sul territorio, con il Gra a rappresentare un vero e proprio spartiacque tra il centrosinistra e il centrodestra (Graf. 1). Le zone con il maggiore consenso per la Bonino sono al centro, a est e sud entro il Gra, più il XIII Municipio (il litorale di Ostia e Acilia) che pur essendo lontano mantiene comunque caratteristiche da centro abitato a sé stante. Le zone

²⁰ TRUGLIA F.G. (2006).

²¹ I Municipi comprendono aree sia centrali che periferiche (si pensi al XVI, che va da Monteverde Vecchio a Malagrotta), e non rappresentano l'unità di analisi più significativa. Viceversa, i 2.600 seggi permettono di disaggregare l'analisi al massimo livello possibile, ma mancano le caratteristiche sociali, economiche e demografiche della popolazione che vi risiede, poiché tali dati sono facilmente disponibili solo per suddivisioni territoriali più grandi. A un livello intermedio vi sono le 115 suddivisioni toponomastiche storiche in rioni, quartieri, suburbi e zone dell'Agro romano, che tuttavia presentano il difetto sia di una ampia eterogeneità in termini di superficie, tra i piccoli rioni del centro storico e le vaste zone esterne al Gra, sia di non essere riconducibili ai Municipi, poiché spesso ritagliate a cavallo di più di uno di essi, e peraltro in maniera poco razionale dal punto di vista urbanistico.

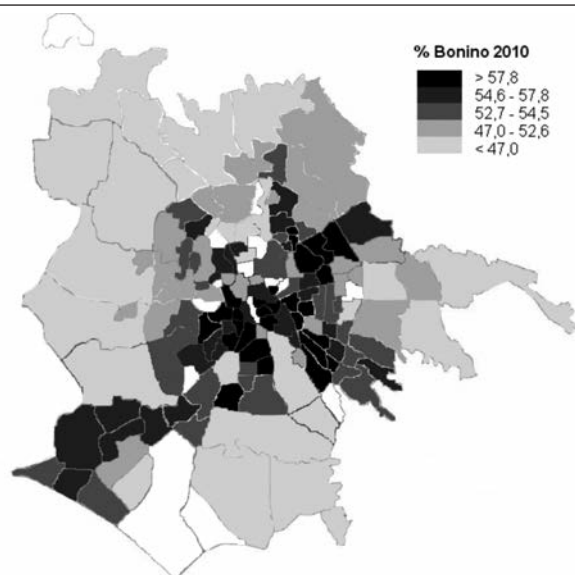
²² Elaborazione dell'autore sulla base dei dati elettorali (riportati nell'archivio *on-line* www.elezioni.comune.roma.it), del viario elettorale (dove sono presenti le strade afferenti a ciascun seggio) e dello stradario (dove si ricava la zona urbanistica di ciascuna strada). Le percentuali riportate differiscono lievemente dai risultati ufficiali sia per errori e lacune nell'archivio *on-line*, sia perché sono esclusi i seggi ospedalieri.

con il minore consenso per il centrosinistra sono invece quelle fuori dal Gra, più il XX Municipio (Cassia e Flaminia) e le ovvie eccezioni dei quartieri centrali tradizionalmente borghesi (Parioli, Eur e Appia Antica), nonché, ma in misura minore, il XVIII e XIX a nord-ovest. Risultati simili si ottengono analizzando anche le precedenti elezioni, che per semplicità non vengono riportate.

A scopo di sintesi, i risultati elettorali delle zone urbanistiche sono stati riaggregati nelle quattro aree urbane corrispondenti al centro e alle tre fasce periferiche individuate nel paragrafo precedente, ampliando la presentazione dei dati a tutte le elezioni a partire dal 2006 e mostrando le zone migliori e peggiori per ogni area (Tav. 3). Nel 2010 la candidata sconfitta del centrosinistra, Emma Bonino, ha vinto in città ovunque tranne che fuori dal Gra: rispetto al massimo del 57% nella periferia storica, e tralasciando il centro, è scesa al 53,8% nella periferia anulare e al 48,2% nella periferia regionale, con uno scarto tra periferia storica e regionale pari a 8,8 punti percentuali. È significativo anche il dettaglio dei quartieri: il massimo pari al 68-69% è stato raggiunto nei rioni popolari ma centrali di San Lorenzo (III Municipio) e Trastevere (I) e, in misura minore con circa il 64%, Testaccio (sempre I) e Tiburtino Sud (V, primo residuo della “cintura rossa”); al contrario, il minimo la Bonino lo ha subito a Prima Porta e Cesano (XX Municipio) con poco meno del 38%.

GRAF. 1

VOTI DEL CANDIDATO DI CENTROSINISTRA NEL 2010, PER ZONA URBANISTICA (%)



Fonte: elaborazione dell'autore su dati di Roma Capitale - Servizi elettorali.

TAV. 3

VOTI DEI CANDIDATI DI CENTROSINISTRA E AFFLUENZA REFERENDARIA, PER AREA URBANA (%)

Area urbana	Referendum affluenza 2011	Regione Presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune Sindaco 2006	Regione Presidente 2005
Centro	55,9	53,1	46,4	50,3	58,0	51,1
Periferia storica (A)	63,0	57,0	49,4	54,1	63,6	57,8
Periferia anulare	61,1	53,8	45,3	50,0	61,7	56,0
Periferia extra-Gra (B)	57,2	48,2	38,8	43,8	57,7	52,5
Totale Roma	60,5	54,2	46,3	50,9	61,4	55,5
<i>Scarto (A – B)</i>	5,8	8,8	10,7	10,2	5,9	5,3
Massimo	Tiburtino Sud 70,7	S. Lorenzo 68,9	S. Lorenzo 62,5	S. Lorenzo 66,4	Testaccio 74,7	Testaccio e S. Lorenzo 67,2
Minimo	Parioli 44,8	Prima Porta 37,6	Boccea 29,1	Prima Porta 34,3	Cesano 45,9	Appia Antica Nord 37,6
<i>Candidati (il vincitore è in grassetto):</i>						
Centrosinistra	-	Bonino	Rutelli	Zingaretti	Veltroni	Marrazzo
Centrodestra	-	Polverini	Alemanno	Antoniozzi	Alemanno	Storace

Nota: aggregazione dei seggi elettorali sulla base delle zone urbanistiche (cfr. Tav. 11 in Appendice); massimo e minimo si riferiscono alle zone con almeno 1.000 iscritti alle liste elettorali.

Fonte: elaborazione dell'autore su dati di Roma Capitale - Servizi elettorali.

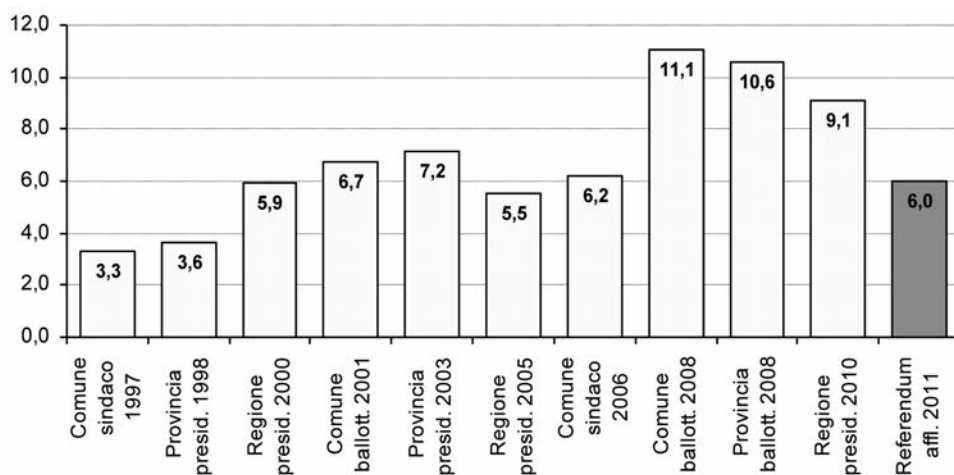
Nel 2008 il candidato sindaco del centrosinistra Francesco Rutelli venne nettamente sconfitto da Gianni Alemanno ovunque, avendo ottenuto il 49,4% nella periferia storica, il 45,3% nella periferia anulare e il 38,8% nella periferia regionale, con uno scarto tra la prima e la terza area di 10,7 punti percentuali. Pur vincendo le elezioni, lo stesso andamento lo aveva registrato il nuovo presidente della Provincia Nicola Zingaretti, candidato dal centrosinistra, il cui risultato migliore è nella periferia storica con il 54,1% dei voti, che scendono al 50% nella periferia anulare, fino al minimo del 43,8% nella periferia più esterna, per uno scarto di 10,2 punti. In entrambe le votazioni il massimo era stato raggiunto a San Lorenzo, Trastevere e Testaccio (I e III Municipio), dove Zingaretti ha superato il 66% e Rutelli il 62%; al contrario, in nessun quartiere esterno al Gra Zingaretti ha ottenuto più del 50% e Rutelli più del 45%, con il minimo a Prima Porta (XX Municipio) e Boccea (XVIII) rispettivamente del 34,3% e del 29,1%.

Le stesse differenze esistevano anche nelle elezioni di due anni prima per il Comune, sia pure tenendo presente che il Sindaco uscente e riconfermato del

centrosinistra, Walter Veltroni, aveva vinto ovunque contro Gianni Alemanno con ampio margine. Nonostante uno scarto tra periferia storica e regionale di soli 5,9 punti percentuali, l'andamento decrescente allontanandosi dal centro si conferma in pieno: il centrosinistra prese il 63,6% nella periferia storica, il 61,7% nella periferia anulare e "solo" il 57,7% nella periferia regionale. Considerazioni simili valgono infine per l'elezione di Piero Marrazzo alla Presidenza della Regione nel 2005, una candidatura poco politicizzata che aveva registrato uno scarto ridotto ma comunque significativo tra periferia storica e regionale (5,3 punti percentuali) e vincendo ovunque (ma con il minimo del 51% proprio al centro): nella periferia storica 57,8%, nella periferia anulare 56% e nella periferia regionale 52,5%. Lo scarto tra le periferie non è però questione recente: è almeno dal 2000 che si registra una differenza significativa in ogni voto per i candidati di centrosinistra, sebbene nelle tre elezioni dal 2008 al 2010 appaia particolarmente rilevante (Graf. 2).

GRAF. 2

SCARTO NEI VOTI DEI CANDIDATI DI CENTROSINISTRA TRA PERIFERIA STORICA E REGIONALE (punti percentuali)



Nota: al fine di rappresentare una serie storica sufficientemente lunga, i seggi elettorali sono aggregati sulla base delle suddivisioni toponomastiche (cfr. Tav. 11 in Appendice), e quindi i dati differiscono rispetto alla Tavola 3 e al testo.

Fonte: elaborazione dell'autore su dati di Roma Capitale - Servizi elettorali.

Potrebbe sembrare improprio aggiungere anche l'affluenza agli ultimi referendum del 2011 nella serie storica del voto per il centrosinistra, ma tale inserimento appare giustificato dalla forte valenza politica attribuita a tale consultazione, molto al di là dell'effettivo significato civico della partecipazione e delle materie oggetto dei quesiti. Si noti che viene presa in considerazione l'affluenza e non la percentuale di "sì", in quanto il gioco politico si basava essenzialmente sull'andare o non andare a votare, da cui deriva una limitata significatività del risultato effettivo. In effetti, la partecipazione ai referendum è notevolmente correlata con il voto ai candidati del centrosinistra, avendo raggiunto il massimo nella periferia storica con il 63%, rispetto al 61,1% della periferia anulare e al 57,2% della periferia regionale; sebbene si confermi uno scarto tra la prima e la terza area pari a 5,8 punti percentuali, la particolarità è rappresentata tuttavia dal fatto che il minimo si registra non fuori dal Gra, ma proprio al centro, con il 55,9%. Tra i quartieri con la maggiore affluenza vi è un insieme eterogeneo guidato da Tiburtino Sud (V Municipio) con oltre il 70% e da Grottaferetta (XI) con il 68%; i valori minori si registrano invece con meno del 45% a Parioli (II), circa il 47% al Centro Storico (I) e circa il 48% a Farnesina ed Acquatraversa (XX).

Nonostante che in valore assoluto le differenze tra un'elezione e l'altra possano essere molto marcate, i quartieri dove il centrosinistra e il centrodestra ottengono i migliori e i peggiori risultati sono quasi sempre gli stessi, a denotare una differenza rispetto alla media cittadina abbastanza stabile nel corso del tempo. La costanza degli elementi di fondo del voto nelle zone urbanistiche si rispecchia ovviamente anche in termini di coefficienti di correlazione, tanto che i dati elettorali sono fortemente collineari, con l'eccezione dell'affluenza ai referendum (Tav. 4). In particolare emergono un paio di gruppi: sono molto correlate da un lato le elezioni più recenti dal 2008 al 2010 (coefficienti compresi tra 0,94 e 0,98), e dall'altro lato le elezioni 2005 e 2006 (0,95), mentre incrociando i due gruppi risulta una correlazione comunque forte sebbene minore (tra 0,80 e 0,87). Ciò vuol dire che solo l'elezione di Piero Marrazzo alla Regione nel 2005 e soprattutto il secondo mandato da sindaco di Walter Veltroni nel 2006, quando le sue caratteristiche personali e la gestione del primo mandato gli hanno evidentemente attribuito un'apparenza *bipartisan*, hanno creato una discontinuità nel comportamento elettorale tradizionale dei romani. Tale discontinuità sarà evidente anche nella diversa significatività delle variabili esplicative nelle regressioni. Discorso a parte per i referendum, la cui affluenza mostra coefficienti di correlazione comunque alti, intorno a 0,71 con le elezioni dal 2005 al 2008 fino al massimo di 0,78 con le regionali del 2010.

CORRELAZIONE TRA I VOTI DEI CANDIDATI DI CENTROSINISTRA

Elezione	Referendum affluenza 2011	Regione Presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune Sindaco 2006	Regione Presidente 2005
Referendum 2011	-	0,78**	0,70**	0,72**	0,71**	0,72**
Regione 2010	0,78**	-	0,95**	0,94**	0,87**	0,87**
Comune 2008	0,70**	0,95**	-	0,98**	0,83**	0,80**
Provincia 2008	0,72**	0,94**	0,98**	-	0,88**	0,84**
Comune 2006	0,71**	0,87**	0,83**	0,88**	-	0,95**
Regione 2005	0,72**	0,87**	0,80**	0,84**	0,95**	-

Nota: coefficienti di correlazione di Pearson; ** = significativo all'1%.

Fonte: elaborazione dell'autore su dati di Roma Capitale - Servizi elettorali.

4. - Fattori socio-economici, demografici e urbanistici

Nel paragrafo precedente i dati elettorali sono stati letti solamente in relazione all'appartenenza alle diverse aree urbane, ossia sintetizzando rispetto alla distanza dal Campidoglio. Ma ovviamente la "perifericità" è solo un sintomo di fenomeni più complessi, che riguardano le dinamiche storiche e demografiche, le condizioni socio-economiche e l'assetto urbanistico. Ad esempio Anderlini (2009), per spiegare i recenti successi della Lega fuori dalle grandi città in una regione tradizionalmente "rossa" come l'Emilia-Romagna, mette in contrapposizione la cultura materiale del lavoro manuale diffusa nelle periferie (siano esse povere o ricche, con bassi livelli di attività oppure con una florida economia manifatturiera) con gli elementi intellettuali post-moderni tipici delle centralità urbane.

Serve quindi un'analisi più approfondita sulle zone urbanistiche, per le quali le fonti disponibili²³ sono il censimento Istat 2001 (nucleo familiare, titolo di studio, condizione lavorativa, settore di attività, stato delle abitazioni), gli aggiornamenti annuali dell'anagrafe comunale (residenti, fasce d'età, cittadinanza, stato civile) e la rilevazione di Provinciattiva sulla dotazione di servizi (Tav. 12 in Appendice). Alcune correlazioni di tali variabili con i dati elettorali appaiono particolarmente ele-

²³ I risultati del censimento 2001 sono stati gentilmente forniti dall'Ufficio Statistico di Roma Capitale, poiché l'ISTAT (2006), a cui si rimanda comunque per la metodologia e la descrizione dei dati, ha pubblicato solo le suddivisioni toponomastiche. I dati anagrafici sono periodicamente resi disponibili su internet dall'Ufficio Statistico stesso. La rilevazione sulla dotazione di servizi è stata gentilmente fornita da Provinciattiva (PROVINCIA DI ROMA, 2010, pp. 22-25), a cui si rimanda anche in questo caso per metodologia e descrizione dei dati.

vate, a mostrare un'associazione stretta e insolita nelle analisi socio-economiche (Tav. 5).

TAV. 5

CORRELAZIONE TRA I VOTI DEI CANDIDATI DI CENTROSINISTRA
E IL CONTESTO URBANISTICO, SOCIALE ED ECONOMICO

Variabile	Referendum affluenza 2011	Regione Presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune Sindaco 2006	Regione Presidente 2005
Distanza dal centro	-0,33**	-0,46**	-0,51**	-0,47**	-0,25**	-0,20*
Densità di popolaz.	0,36**	0,51**	0,60**	0,58**	0,40**	0,36**
Variatz. % residenti	-0,37**	-0,45**	-0,50**	-0,49**	-0,34**	-0,32**
Stranieri non Ocse	-0,47**	-0,25**	-0,27**	-0,35**	-0,35**	-0,37**
Stranieri Europa est	-0,36**	-0,36**	-0,48**	-0,51**	-0,36**	-0,36**
Età <15 anni	-0,20*	-0,41**	-0,57**	-0,56**	-0,24**	-0,19*
Età >75 anni	0,19*	0,40**	0,54**	0,48**	0,16	0,08
Indice di vecchiaia	0,20*	0,43**	0,59**	0,55**	0,20*	0,14
Single	0,08	0,41**	0,53**	0,47**	0,20*	0,10
Imprenditori e prof.	-0,27**	-0,09	0,02	-0,07	-0,35**	-0,49**
Addetti edilizia	-0,41**	-0,49**	-0,57**	-0,53**	-0,26**	-0,15
Addetti scuola	0,25**	0,44**	0,55**	0,47**	0,17*	0,06
Grande distribuzione	-0,47**	-0,37**	-0,29**	-0,23**	-0,26**	-0,21*
Piazze per ettaro	0,07	0,37**	0,47**	0,43**	0,22**	0,16

Nota: coefficienti di correlazione di Pearson; ** = significativo all'1%, * = significativo al 5%.

Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

La densità di popolazione incide notevolmente su tutte le ultime elezioni, con coefficienti di correlazione positivi che nel caso del ballottaggio per il Comune nel 2008 arrivano a 0,6. Altre variabili significative risultano essere dal punto di vista demografico la variazione percentuale dei residenti (negativa), la presenza di stranieri (negativa, sia essa misurata in termini di provenienza da paesi non Ocse oppure dall'Europa dell'est) e la quota di giovani (negativa), di anziani e di single (positiva); dal punto di vista occupazionale l'incidenza di imprenditori e professionisti (negativa) e dei settori dell'edilizia (negativa) e della scuola (positiva); dal punto di vista della dotazione di servizi, la disponibilità di grandi superfici di vendita e la densità di piazze per ettaro (entrambe negative). La sostanziale costanza delle scelte elettorali dei romani, sia pure con la discontinuità del 2005 e 2006, come mostrato nel paragrafo precedente, rende sufficientemente robuste le considerazioni che seguono al variare delle elezioni. Di conseguenza, per semplicità, sarà presa in esame soprattutto la differenza tra le zone urbanistiche

nel voto al ballottaggio per il Comune nel 2008, che rappresenta l'elezione dove lo scarto tra periferia storica e regionale è maggiore.

La forte correlazione tra le variabili di contesto e i dati elettorali ha un primo effetto sulla differenza nel voto tra zone urbanistiche più o meno caratterizzate dai fattori socio-economici, demografici e urbanistici in esame. In termini puramente descrittivi, emergono grandi diversità nella media del voto tra il quarto di zone con i valori più alti e il quarto con i valori più bassi per alcune variabili, evidenziando così i fattori più significativi per spiegare la trama territoriale delle scelte elettorali (Tav. 6).

TAV. 6

VOTI DEL CANDIDATO DI CENTROSINISTRA ALLE COMUNALI DEL 2008
SECONDO IL CONTESTO URBANISTICO, SOCIALE ED ECONOMICO (%)

Variabile	Zone urbanistiche con valori superiori al primo quartile	Zone urbanistiche con valori inferiori al terzo quartile	Differenza
<i>Prevalenza per il centrosinistra</i>			
Densità di popolazione	49,1	35,6	13,5
Piazze per ettaro	49,4	36,7	12,7
Indice di vecchiaia	50,3	39,2	11,1
Addetti pubblici	49,7	39,0	10,7
Età 60-75 anni	49,0	39,9	9,1
Età >75 anni	49,2	41,2	8,1
Single	49,0	42,2	6,8
Addetti alla scuola	48,0	41,3	6,7
<i>Prevalenza per il centrodestra</i>			
Compon. nucleo familiare	39,0	50,2	-11,2
Variazione % dei residenti	39,6	49,3	-9,7
Stranieri dall'est Europa	39,3	48,8	-9,5
Età <15 anni	40,7	50,1	-9,4
Addetti all'edilizia	40,3	49,0	-8,7
Età 30-45 anni	39,7	48,0	-8,2
Tasso di attività	42,5	49,4	-6,8
Coniugati	42,4	48,6	-6,2

Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

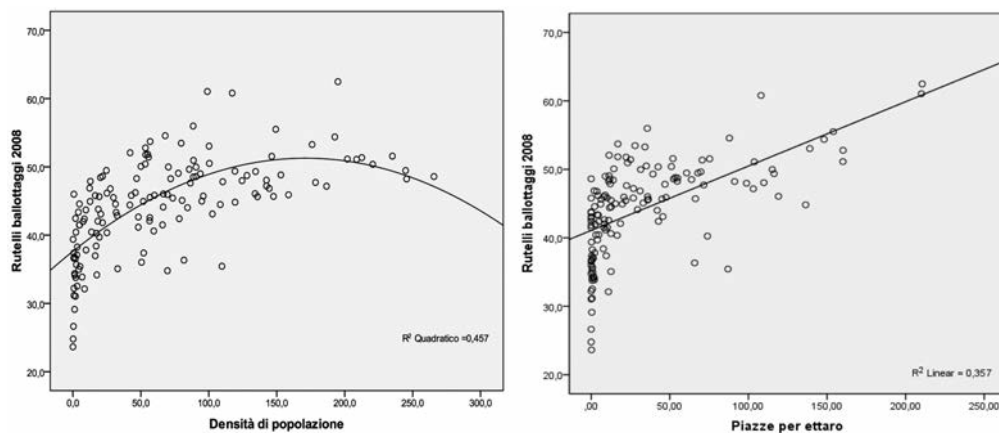
Lo scarto più grande riguarda due indicatori urbanistici: la densità di popolazione e la densità di piazze per ettaro. Nelle zone urbanistiche con la maggiore densità di popolazione il centrosinistra ha ottenuto il 49,1% mentre in quelle con i valori minori solo il 35,6%, ossia 13,5 punti percentuali in meno. Analogamente, nelle zone urbanistiche con la maggiore densità di piazze il suo consenso ha raggiunto il 49,4%

mentre in quelle con i valori minori si è fermato al 36,7%, pari a 12,7 punti percentuali in meno. Scarti superiori ai 10 punti percentuali, in positivo o in negativo, risultano in relazione all'indice di vecchiaia (50,3% rispetto a 39,2%), alla quota di dipendenti pubblici (49,7% contro 39%) e al numero di componenti il nucleo familiare (39% rispetto a 50,2%). Ma vi sono scarti rilevanti compresi tra 6 e 10 punti percentuali anche considerando (con impatto positivo) la quota di residenti tra 60 e 75 anni oppure oltre i 75 anni, di single e di dipendenti scolastici, nonché (con impatto negativo) la variazione percentuale dei residenti, la presenza di stranieri provenienti dall'Europa dell'est, la quota di residenti con meno di 15 anni o tra 30 e 45 anni, gli addetti all'edilizia, il tasso di attività e la percentuale di coniugati.

La densità di popolazione appare particolarmente importante nello spiegare le differenze, ma in una forma non lineare rispetto al voto al centrosinistra: quest'ultimo, infatti, aumenta al crescere della densità, ma solo fino a una certa soglia di "affollamento" urbano, oltre la quale la correlazione diventa negativa. Tale variabile sembra quindi avere una relazione quadratica con i dati elettorali, che è particolarmente evidente nel caso del voto per il Comune nel 2008 (Graf. 3). Ciò può essere interpretato in relazione agli svantaggi crescenti del vivere in quartieri ad alta densità, che fino a una certa soglia – individuabile approssimativamente in circa 200 abitanti per ettaro – sono più che compensati dai benefici in termini di relazioni interpersonali e opportunità di partecipazione collettiva, ma oltre tale soglia i problemi prevalgono sui vantaggi a causa dell'eccessiva concentrazione di persone, attività e automobili, con l'inevitabile corollario di traffico, rumore e confusione. Altrettanto rilevante appare la disponibilità di piazze per ettaro, di cui risulta una forte correlazione lineare rispetto a tutte le elezioni, e in particolare anche in questo caso con le comunali 2008, escludendo il valore *outlier* del Centro Storico e considerando che alcune zone urbanistiche non hanno alcuna "piazza" nella propria toponomastica.

GRAF. 3

VOTI DEL CANDIDATO DI CENTROSINISTRA ALLE COMUNALI DEL 2008
SECONDO LA DENSITÀ DI POPOLAZIONE E LA DISPONIBILITÀ DI PIAZZE



Fonte: elaborazione dell'autore sui dati della Tavola 12 in appendice.

Nel paragrafo successivo sarà data un'interpretazione dei legami tra le variabili analizzate (e in particolare la densità e le piazze) e i dati elettorali, ma già sulla base di queste prime analisi descrittive, tra le variabili maggiormente capaci di spiegare le differenze di voto a Roma, appare interessante la mancanza dell'istruzione e del reddito dei residenti, che al contrario nelle analisi elettorali nazionali risultano decisive (Graf. 4).

La capacità per chi possiede titoli di studio più elevati di cercare informazioni, attivare forme di partecipazione politica e crearsi una propria opinione, è legata in Italia al voto per il centrosinistra, e il contrario vale per i gruppi sociali che partecipano e si informano meno²⁴. A Roma, invece, altre indagini hanno già mostrato come il titolo di studio, vista la diffusione dei livelli di istruzione elevati, non risulti un indicatore particolarmente incisivo per identificare i comportamenti e localizzare i gruppi sociali; in particolare, si osserva una scarsa segregazione residenziale sia in termini socio-economici per le classi elevate (per quanto storicamente presenti al centro e nelle appendici a nord tra Monte Mario e Salaria e a sud intorno all'Eur), sia su basi etniche²⁵. È probabile quindi che l'istruzione elevata a Roma non sia una caratteristica tale da individuare una categoria specifica di persone, ma denoti invece almeno due tipologie contrapposte. Da un lato è ti-

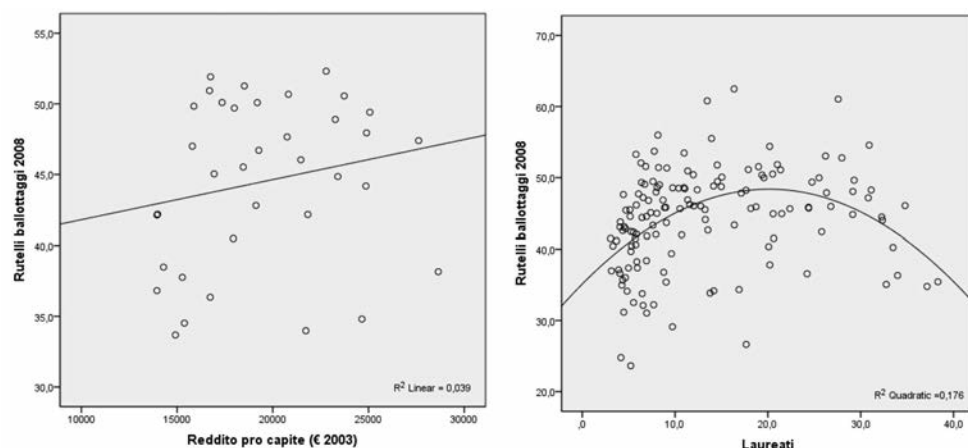
²⁴ FRUNCILLO D. (2010, pp. 177-204; pp. 233-234).

²⁵ VIOLANTE A. (2008, pp. 114-116).

pico un ceto medio-borghese di piccoli imprenditori e liberi professionisti con tendenze conservatrici e moderate (il cosiddetto “generone”), storicamente schierato a destra, e dall’altro lato caratterizza un ceto medio-impiegatizio soprattutto nel settore della pubblica amministrazione e della scuola, che rappresenta la tipologia urbana sui cui il centrosinistra sembra basare sempre più il suo consenso, non solo a Roma²⁶. Per entrambe le tipologie è comprensibile che le variabili relative allo status occupazionale e al settore di attività incidano maggiormente rispetto al titolo di studio.

GRAF. 4

VOTI DEL CANDIDATO DI CENTROSINISTRA ALLE COMUNALI DEL 2008
SECONDO IL REDDITO *PRO CAPITE* E L’INCIDENZA DEI LAUREATI



Fonte: elaborazione dell'autore sui dati della Tavola 12 in aAppendice.

Neanche il reddito sembra spiegare i risultati elettorali, ma in questo caso almeno in parte la ragione risiede in una lacuna statistica, ossia nella mancanza di tali dati a un livello di disaggregazione sufficiente per la presente analisi. L'unica indagine dettagliata sembra essere quella condotta dall'Istituto Tagliacarne (2005) su gruppi di suddivisioni toponomastiche romane, ma il reddito pro capite delle 36 aree analizzate presenta indici di correlazione non significativi con i dati elettorali degli ultimi anni, peraltro sia positivi che negativi. Ciò dipende probabilmente da un consenso per il centrosinistra che risulta abbastanza trasversale in termini socio-economici, essendo forte sia nei quartieri centrali con reddito elevato (es. Centro storico e Gianicolense) sia in quelli della prima periferia con red-

²⁶ ANDERLINI F. (2011, pp. 72-73).

dito medio-basso (es. Tiburtino, Prenestino, Tuscolano), e analogamente vale per il centrodestra tra i quartieri tradizionalmente borghesi (es. Parioli e Fleming) e le periferie relativamente più povere (es. Borghesiana, Ponte Galeria e Prima Porta). Il reddito pro capite è comunque positivamente correlato dal punto di vista urbanistico con la disponibilità di piazze per abitante e per ettaro, e dal punto di vista socio-economico con le fasce d'età oltre i 50 anni, un basso numero di componenti del nucleo familiare, i livelli di istruzione più elevati, il settore della finanza e dei servizi alle imprese, la quota di imprenditori e professionisti e (piuttosto sorprendentemente) con la quota di studenti.

In linea con quanto rilevato a livello nazionale da Fruncillo (2010, p. 206), anche a Roma “non sembra venga incrinato lo schema fondamentale che assegna al centrosinistra la capacità di ottenere un surplus di voti tra coloro che hanno caratteristiche di maggiore centralità sociale e soprattutto consegna al centrodestra un vantaggio più evidente nei ceti sociali periferici ai quali in passato i partiti di sinistra si rivolgevano prioritariamente”²⁷. Solo che questa dicotomia tra centralità e perifericità non va intesa nel senso di un alto o basso status socio-economico, quanto piuttosto in termini urbanistici, ossia in relazione all'effettiva dislocazione territoriale dei diversi quartieri. Densità, distanza, fasce d'età, nazionalità, condizione professionale e settore di attività dicono in fondo la stessa cosa: il consenso per il centrosinistra è maggiore nei quartieri un tempo popolari perché in periferia (ma nella prima fascia, quella oggi “storica”), dove la densità è alta (ma fino a un certo punto), la distanza dal centro ridotta, i residenti anziani, le famiglie poco numerose, gli stranieri rari, gli imprenditori e professionisti pochi e al contrario molti i dipendenti pubblici e scolastici.

5. - La disponibilità di beni comuni e relazionali

In prima approssimazione, la regolarità tra voto e centralità/perifericità sembrerebbe legata all'idea intuitiva che nelle aree più centrali della città si registri la maggiore qualità della vita e man mano che ci si allontana nelle periferie la sod-

²⁷ Sebbene ciò non rientri nell'oggetto del presente lavoro, si osservi come queste dinamiche creino al centrosinistra non solo un problema politico, avendo ormai le roccaforti in quartieri borghesi e intellettuali (es. Monteverde e Trastevere) e poco spazio di manovra nei nuovi quartieri popolari. Infatti, anche volendo scegliere di mantenere questa nuova fisionomia, vi è un problema elettorale: la periferia storica da sola non basta comunque a vincere le elezioni, perché i residenti continuano a trasferirsi da lì verso quartieri più lontani, tanto che i voti validi della prima fascia periferica sul totale romano sono diminuiti dal 51,4% nel 1996 al 47,4% nel 2010.

disfazione declini, in relazione sia alle maggiori distanze da percorrere che a una minore efficacia dei servizi pubblici. Tuttavia, le poche indagini campionarie che hanno tentato di rilevare le valutazioni soggettive in merito alla qualità della vita dei residenti non suffragano tale convincimento.

L'Agenzia per il controllo e la qualità dei servizi pubblici locali del Comune di Roma (2011, pp. 11-15) effettua annualmente rilevazioni di *citizen satisfaction* volte ad approfondire, secondo la percezione soggettiva dei cittadini/utenti romani, il legame esistente tra il sistema dei servizi pubblici locali e la qualità della vita. Dal punto di vista territoriale è addirittura la residenza nella periferia esterna al Gra a mostrare il maggiore impatto positivo sulla qualità della vita, con un punteggio pari a 6,44 (in una scala da 0 a 10) rispetto al voto medio di 6-6,1 nelle altre aree urbane. Anche la soddisfazione media per i servizi pubblici oggetto dell'indagine è risultata pari a 6,53-6,55 sia nel centro (all'interno dell'anello ferroviario) che nella periferia esterna al Gra, rispetto a valori intorno a 6,4 nella fascia compresa tra l'anello e il Gra. Tale risultato è confermato dall'indagine campionaria del Comune di Roma (2007, cap. 5), secondo cui sono proprio i Municipi più periferici a mostrare la maggiore soddisfazione di vivere sia a Roma in generale (IV, VIII e X raggiungono oltre il 95% di molto o abbastanza soddisfatti) sia nella specifica zona di residenza (XX, XVI, XVIII e XII superano il 90%, insieme al XVII che è centrale).

Danno esiti analoghi le rilevazioni ai testimoni privilegiati e a un campione di residenti realizzate dall'AIC e Unione Borgate (2010, pp. 151-187) e dalla Provincia di Roma (2010, pp. 31-36), rispettivamente nelle borgate ex abusive e nei quartieri periferici della prima cintura dell'hinterland. Sebbene i quartieri indagati siano eterogenei, i risultati mostrano che una parte consistente degli intervistati esprime una buona soddisfazione per la vita che conduce, grazie alla proprietà dell'alloggio, al basso costo di beni e servizi, alla mancanza di preoccupazioni economiche, all'elevata integrazione degli stranieri, alla percezione di sicurezza diffusa (eccetto il rischio dei furti in casa). In generale i residenti delle periferie, comprese le storiche borgate ex abusive, sono abbastanza contenti di dove vivono e difficilmente vorrebbero trasferirsi altrove.

Sebbene ciò possa apparire controintuitivo, poiché siamo inclini a individuare nella perifericità una fonte di insoddisfazione *a priori*, evidentemente la scelta economica di vivere fuori dalla città consolidata presenta anche aspetti positivi, ad esempio in termini di maggiori spazi e minore inquinamento, che migliorano la percezione della propria qualità della vita nonostante la maggiore lontananza dai principali servizi. Peraltro, in numerosi nuovi insediamenti periferici, la di-

stanza dal centro viene compensata con ampiezza e dotazioni dell'abitazione e degli spazi condominiali tali da far crescere il prezzo fino ad equipararlo con quello di case più centrali, creando quartieri per famiglie ad alto reddito accanto ad aree più popolari²⁸. Quindi non può essere la qualità della vita il fattore decisivo per spiegare le dinamiche elettorali.

Altre indagini hanno suggerito la possibilità che, almeno a Roma, a incidere in maniera significativa sul benessere dei cittadini siano non tanto i fattori "macro", quanto piuttosto il livello "meso", vale a dire le opportunità di interrelazioni sociali nel contesto in cui si vive, e quello "micro", ossia le risorse psicosociali individuali²⁹.

Un ragionamento alternativo potrebbe quindi fondarsi sul fatto che la periferia storica, proprio perché densa e intensiva, è composta da quartieri dove le relazioni interpersonali sono forti e le opportunità di partecipazione collettiva sono numerose. Al contrario ciò che caratterizza in maniera nettamente percepibile l'estrema periferia di Roma, sia quella abitata dai ceti medio-alti (basti pensare alla *gated community* dell'Olgiatea) sia quella più popolare, è piuttosto lo *sprawl* urbano, ossia la dispersione insediativa causata dalla crescita rapida e disordinata³⁰, nonostante la popolazione stabile, e le sue conseguenze sociali meno visibili ma altrettanto forti in termini di isolamento rispetto al resto della società. Se lo *sprawl*, insieme alla mancata pianificazione urbanistica (nelle periferie ex abusive) o alla progettazione lasciata nelle mani dei costruttori privati (nei nuovi insediamenti residenziali), non determina una bassa qualità della vita, aumenta però i costi di gestione dei servizi pubblici³¹ e, soprattutto, limita fortemente le potenzialità della città.

I residenti nelle periferie estreme sono lontani dai grandi eventi e dai grandi progetti, deboli utenti dei servizi pubblici di alta qualità al centro, utilizzatori forzati dell'automobile, poco forniti sia di servizi di base che di spazi pubblici e collettivi. Sono minori e più deboli le relazioni interpersonali e i legami sociali perché si hanno poche occasioni di contatto, e le persone sono più isolate e prive di identità collettive e storicizzate. Appaiono limitate le opportunità di incontro tra persone, e di conseguenza è scarsamente probabile la produzione e il contestuale

²⁸ VIOLANTE A. (2008, p. 134).

²⁹ BONCIANI M. (2009).

³⁰ GIBELLI M.C. - SALZANO E. (2006).

³¹ CAMAGNI R. *et al.* (2002); EEA (2006).

consumo di beni relazionali, che nell'ottica dell'economia civile rappresentano il fattore determinante sia del benessere individuale che dello sviluppo locale³².

In particolare, si utilizza molto più frequentemente l'automobile, visto che negli ultimi anni pochi insediamenti residenziali, commerciali o terziari sono stati realizzati o progettati in prossimità delle stazioni su ferro³³. Prendendo come riferimento gli spostamenti per servizi di gestione familiare (acquisti, uffici pubblici, ambulatori, ecc.), che in genere necessitano di percorrenze limitate nel quartiere dove si vive, dentro l'anello ferroviario prevale l'uso dei mezzi pubblici (46,9% rispetto al 33,4% di auto e moto, ma con un notevole 19,7% di spostamenti a piedi o in bici), mentre dentro la fascia verde vi è una sostanziale parità (41% contro 43,6%) e oltre dominano nettamente i mezzi privati (61,1% dentro il Gra e 69,5% fuori dal Gra contro rispettivamente 29,9% e 22,9% per il trasporto pubblico³⁴, e meno del 10% per piedi e bici)³⁵.

Perde significato lo spazio intermedio tra le abitazioni e le destinazioni, siano esse lavorative, ricreative o commerciali: se per Reale (2010) «non è solo un fatto fisico, ma anche sociale: la città dello *sprawl* è la città degli individui, il luogo del privato», AIC e Unione Borgate (2010, p. 151) sottolineano che gli individui «non si incontrano, non hanno mantenuto una socialità, né forse mai sviluppato un'identità collettiva. [...] l'individualismo si trasforma in isolamento, in chiusura nella propria casa e nella cura del proprio giardino, nel senso sia simbolico che letterale, in disinteresse per le questioni di ordine collettivo e civile».

Sono pochi i luoghi di socializzazione, e in particolare pochissime le piazze: analogamente al voto per il centrosinistra, la disponibilità di spazi pubblici è inversamente proporzionale alla distanza dal Campidoglio (Graf. 5). Il numero di piazze ogni 1.000 abitanti scende infatti progressivamente da 1,36 nei quartieri centrali a 0,42 nella periferia storica, fino a 0,32 nella periferia anulare e 0,21 fuori dal Gra; ancora più nettamente, la densità di piazze per ettaro diminuisce da 78,1 nei quartieri centrali a 41,9 nella periferia storica, fino a 10,8 nella periferia anulare e 1,3 fuori dal Gra³⁶.

³² BRUNI L. - ZAMAGNI S. (2004); CALAFATI A.G. (2002); GUI B. (2002); ZAMAGNI S. (2005a; 2005b).

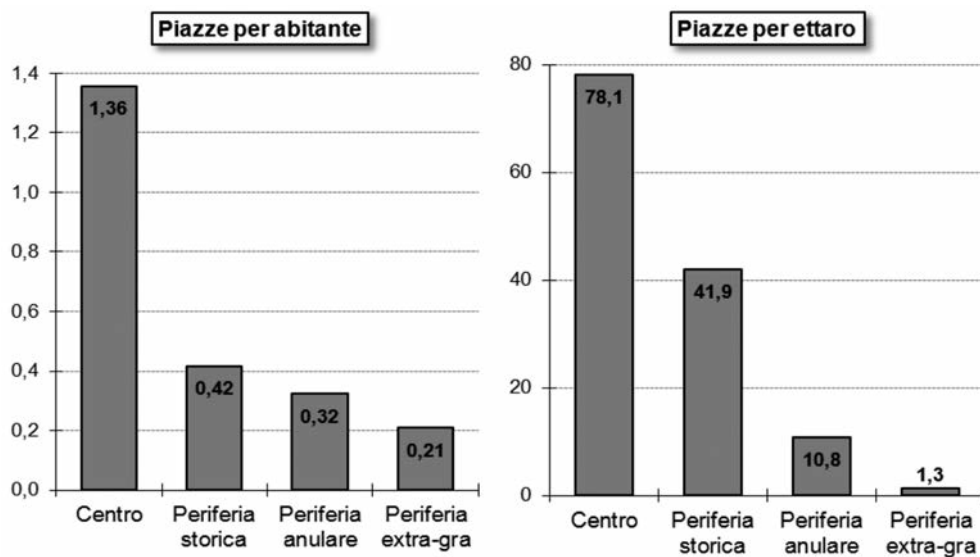
³³ MANIGRASSO M. (2011, p. 97).

³⁴ Peraltro, l'uso del trasporto pubblico dipende fortemente dalla non disponibilità dell'automobile, poiché tra i residenti fuori dal Gra ben il 40% di chi usa i mezzi pubblici lo fa solo per questo motivo.

³⁵ AGENZIA PER IL CONTROLLO E LA QUALITÀ DEI SERVIZI PUBBLICI LOCALI DEL COMUNE DI ROMA (2009, cap. 2).

³⁶ AIC - UNIONE BORGATE (2010, pp. 42-44); PROVINCIA DI ROMA (2010, pp. 52-55).

DISPONIBILITÀ DI PIAZZE PER AREA URBANA (2009)



Fonte: elaborazione dell'autore su dati di PROVINCIAITIVA.

Secondo Salzano (2010), sono proprio gli spazi pubblici a esprimere l'espressione più originale e persino l'essenza delle città europee: «i luoghi nei quali stare insieme, commerciare, celebrare insieme i riti religiosi, svolgere attività comuni e utilizzare servizi comuni. [...] la cerniera tra il dentro e il fuori, il luogo dove la città – tramite l'incontro con il “diverso” – si apre al mondo, lo conosce e ne diviene parte: [...] il luogo della socializzazione di tutti i cittadini». Anzi, il disordine e l'eterogeneità tipici degli spazi pubblici, solitamente più caotici rispetto a quelli privati, possono rappresentare non tanto un fattore di fastidio e insicurezza, quanto piuttosto un elemento che favorisce la creatività e il cambiamento. Nella città contemporanea, invece, gli spazi pubblici – compresi quelli verdi – sono ridotti a “vuoti tra le case”, non sono oggetto né di cura né di un progetto comune, e mancano connessioni fruibili a piedi, contribuendo a rendere i quartieri poco permeabili e a impedirne l'attraversamento, a tutto vantaggio della circolazione automobilistica, che beneficia di rotonde, svincoli e vasti parcheggi su strada. I centri commerciali diventano il fulcro della vita dei quartieri anche a fini ricreativi, dove si passa il tempo perché non c'è altro da fare o perché si trovano servizi non disponibili altrove³⁷.

³⁷ AIC - UNIONE BORGATE (2010, pp. 168-169); ANNUNZIATA S. - COSSU M. (2010); CELLAMARE C. (2011); COTTINO P. (2003); CREMASCHI M. (2010).

La riflessione teorica degli urbanisti sul “welfare urbano” e sul “diritto alla città” mostra che vi è la possibilità sia di fruire dei beni comuni costituiti dall’organizzazione urbana del territorio, sia di partecipare alle decisioni sulle trasformazioni, solo se esiste una pianificazione urbanistica governata dell’“habitat” umano³⁸. Quando al contrario il territorio è una mera aggregazione di frammenti, per Salzano (2010) il welfare urbano è indebolito dal fatto di «sostituire agli spazi pubblici i “non luoghi”, caratterizzati dalla ricerca dei requisiti opposti a quelli che rendono pubblica una piazza (lo spazio pubblico per antonomasia): la recinzione mentre la piazza è aperta, la sicurezza mentre la piazza è avventura, l’omologazione mentre la piazza è differenza e identità, la natura delle persone che la abitano, clienti anziché di cittadini, la distanza dalla vita quotidiana anziché la sua prossimità».

Ma Burkhalter e Castells (2009, pp. 20-21, nostra traduzione) specificano che: «lo spazio pubblico non è necessariamente una piazza, un parco, un monumento, o qualsiasi altra configurazione spaziale che deriva dalla piazza. Può essere tutto ciò, ma anche un incrocio, una strada, un campo sportivo, l’ingresso di una stazione ferroviaria, il sagrato di una chiesa, una sequenza di ristoranti, bar e cinema che facilitano incontri spontanei. Uno spazio pubblico è dove le persone si incontrano senza scopi particolari oppure per un’ampia varietà di motivi». Nella stessa ottica, fin dagli anni ‘60, Insolera (2011, pp. 213-214) aveva compreso che se nei quartieri intensivi la qualità dell’abitare è scarsa, al contempo però l’ampia disponibilità di negozi e servizi – favorita proprio dall’elevata densità – rende soddisfacente il “comfort cittadino” e le possibilità di partecipazione alla “cultura della città”. Cellamare (2011) spiega come invece nelle nuove periferie venga meno la funzione sociale del commercio, a causa di «una radicale riduzione della frequentazione di queste strade, l’innescò di un fenomeno di percezione dell’insicurezza, in un circolo vizioso che porta ad un complessivo impoverimento della vita di quartiere».

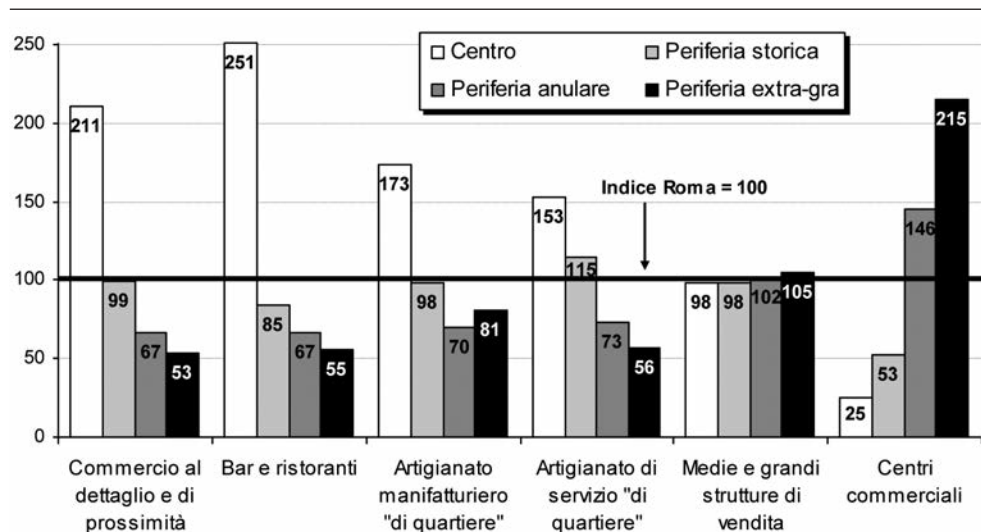
Rapportata a 100 la dotazione media romana, i quartieri più lontani dal centro mostrano una scarsa disponibilità pro capite di quasi tutti gli esercizi privati che creano occasioni di incontro (Graf. 6): commercio al dettaglio e di prossimità (nella periferia esterna al Gra il 53% della media), bar e ristoranti (55%), artigianato manifatturiero (81%) e di servizio (56%). Al contrario, mentre le medie e grandi superfici di vendita sono equamente distribuite in tutte le fasce urbane, l’unico aspetto per il quale la periferia storica è meno dotata sono ovviamente i centri commerciali, la cui disponibilità è oltre 4 volte maggiore nella periferia esterna al Gra (215% rispetto alla media romana). Scendendo al livello delle zone

³⁸ SALZANO E. (2011).

urbanistiche, mentre il numero di centri commerciali è troppo basso – e anzi nullo nella maggior parte di esse – perché possa avere un risultato significativo nelle analisi statistiche, la disponibilità della grande distribuzione è invece sufficientemente eterogenea da poter essere utilmente impiegata come *proxy* della carenza di beni comuni.

GRAF. 6

DISPONIBILITÀ *PRO CAPITE* DI ESERCIZI PRIVATI PER AREA URBANA
(ROMA = 100; 2009)



Fonte: elaborazione dell'autore su dati di PROVINCIAITIVA.

La mancanza di spazi pubblici e quindi di occasioni anche casuali di incontro e scambio (il che avviene solo marginalmente nei centri commerciali) riduce la propensione a impegnarsi in attività collettive, rende improbabile dedicare tempo ad attività sociali e politiche, limita lo stimolo a partecipare alla vita delle associazioni presenti sul territorio. In termini di capitale sociale, in generale in Italia è proprio la presenza di legami *linking* (dentro organizzazioni e imprese) e *bridging* (rapporti anche di vicinato tra individui eterogenei) che sembra favorire la partecipazione sociale, la coscienza civica e lo sviluppo umano³⁹, a differenza dei legami *bonding* che caratterizzano i gruppi omogenei ed esclusivi (anche con riferimento a interessi personali o familiari di natura economica, che in alcuni territori pos-

³⁹ SABATINI F. (2008, 2009a).

sono sfociare nel fenomeno del voto di scambio). La presenza sia di spazi pubblici (misurata in positivo dal numero di piazze per ettaro) che di attività private con una certa valenza sociale (misurata in negativo dalla grande distribuzione per abitante) può appunto rappresentare un indicatore della disponibilità di capitale sociale *bridging*, adottando una definizione di capitale sociale come fattore di benessere e qualità dello sviluppo⁴⁰. Una misura sia pure imperfetta di capitale sociale *bonding* può invece essere rinvenuta nella percentuale di coniugati sui residenti, in termini quindi di esistenza di meri legami familiari.

È proprio attraverso la scarsa qualità urbana e la carenza di beni comuni e relazionali che lo *sprawl* delle periferie può arrivare ad avere conseguenze sociali e culturali, che a loro volta incidono sulle scelte elettorali, legando così le dinamiche urbanistiche con quelle politiche. La logica sembra simile a quanto avviene a livello nazionale, laddove il centrodestra si impone tra i gruppi sociali, anche ricchi e istruiti, che però hanno carenze nelle reti di partecipazione alla vita collettiva⁴¹ e di conseguenza un basso livello di relazioni, scambi e fiducia interpersonale.

6. - Analisi statistiche

Gli elementi alla base delle differenze elettorali risultano bene evidenti mediante un'analisi in componenti principali (ACP) per mostrare l'associazione tra le variabili di contesto socio-economico, demografico e urbanistico, nonché individuare eventuali fattori latenti⁴². Per spiegare almeno il 60% della variabilità complessiva è necessario estrarre tre assi (Graf. 7): la prima componente (35,6%) è interpretabile come "status socio-economico" legato al titolo di studio, alla professione autonoma e all'età adulta, la seconda (13,4%) come "anzianità" sia degli abitanti che delle abitazioni nei quartieri densi vicini al centro, e infine la terza non mostrata nel grafico (10,8%) come "carenza di legami familiari" perché single e divorziati o perché stranieri (Tav. 13 in Appendice).

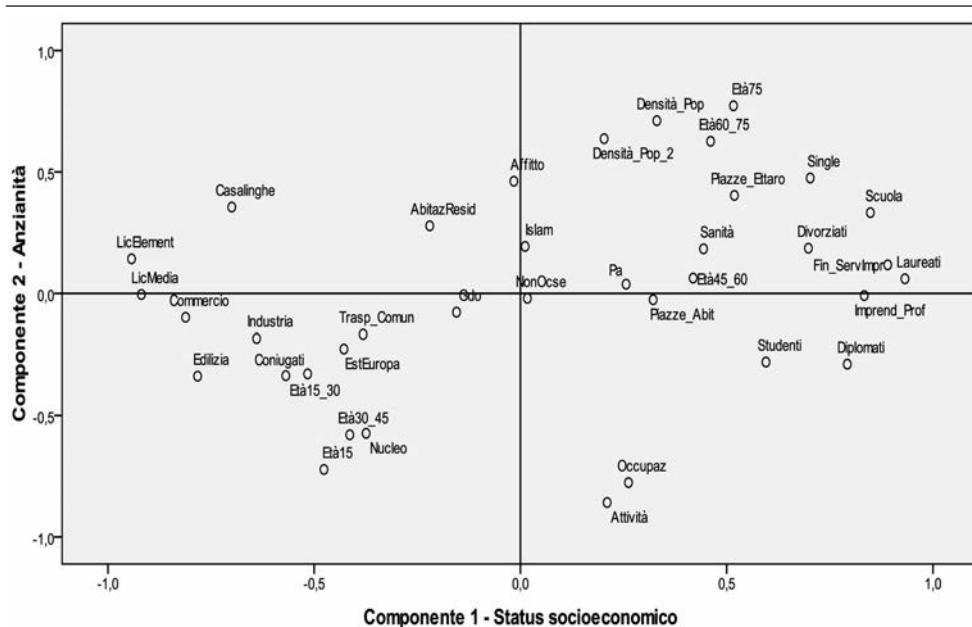
⁴⁰ SABATINI F. (2009b).

⁴¹ FRUNCILLO D. (2010, pp. 164-176; pp. 241-262).

⁴² In via preventiva è stata calcolata la correlazione tra le variabili oggetto di analisi, in modo da verificarne l'eventuale collinearità, da cui non emerge nessuna coppia con indice di Pearson maggiore di 0,80 eccetto le variabili elettorali come già messo in evidenza.

GRAF. 7

ANALISI IN COMPONENTI PRINCIPALI SUI PRIMI DUE ASSI DELLE VARIABILI SOCIO-ECONOMICHE, DEMOGRAFICHE E URBANISTICHE



Nota: analisi in componenti principali con rotazione Varimax e normalizzazione di Kaiser; varianza totale spiegata = 59,8%; autovalori, comunalità e pesi delle variabili sulle tre componenti sono riportati nella Tavola 13 in Appendice.

Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

Le zone che caratterizzano i tre assi presentano differenze ben marcate (Tav. 7). Sul semiasse positivo della prima componente troviamo le zone accomunate da attrattività e valori immobiliari superiori alla media romana (a sud Grottaferretta, Cecchignola, Torrino e Tre Fontane; a nord Acquatraversa e Medaglie d'Oro), mentre sul semiasse negativo alcuni quartieri recenti e molto periferici meno attrattivi (in particolare a est Lunghezza, dove ricade il nuovo quartiere di Ponte di Nona). La seconda componente è invece caratterizzata in positivo dai grandi insediamenti storici e densi nella periferia storica o anulare, costruiti non oltre gli anni '70 (Don Bosco, Tufello, Gordiani, Torrespaccata, Villaggio Olimpico, Eroi, Marconi), e in negativo dalle zone a bassa densità con numerose ville (soprattutto l'Appia Antica e la parte interna del XIII Municipio). Infine la terza componente ha sul semiasse positivo le aree centrali dove le famiglie standard sono poche per la presenza di residenti temporanei e stranieri (Centro Storico, Traste-

vere, Esquilino, XX Settembre, oltre a zone poco abitate) e sul semiasse negativo gli insediamenti densi della periferia anulare (a nord Serpentara; a est Tor Tre Teste, Tiburtino Sud e Osteria del Curato; a sud Cecchignola e Grotta-perfetta).

TAV. 7

PUNTEGGI DELLE ZONE URBANISTICHE SUI TRE ASSI DELL'ANALISI IN COMPONENTI PRINCIPALI

Zone	1 ^a componente ("status socio-economico")	2 ^a componente ("anzianità")	3 ^a componente ("carenza di legami familiari")
Prime 7	11G Grotta-perfetta 2,25	10A Don Bosco 2,35	1A Centro Storico 5,06
	20B Acquatraversa 2,20	4I Tufello 2,26	1B Trastevere 4,08
	12E Cecchignola 2,10	6D Gordiani 2,07	7G Centro dir. Centocelle 2,32
	12C Torrino 1,92	8A Torrespaccata 1,83	1E Esquilino 2,04
	11F Tre Fontane 1,92	17C Eroi 1,69	15E Magliana 2,04
	5I S. Alessandro 1,55	2A Villaggio Olimpico 1,68	4O Tor S. Giovanni 1,66
	19A Medaglie d'Oro 1,46	15A Marconi 1,62	1F XX Settembre 1,62
Ultime 7	8E Lunghezza -1,60	11Y Appia Antica Sud -1,85	4E Serpentara -1,33
	8H S. Vittorino -1,67	15E Magliana -1,95	12E Cecchignola -1,34
	9C Tor Fiscale -1,89	13A Malafede -2,00	11G Grotta-perfetta -1,54
	4O Tor S. Giovanni -1,91	13I Infernetto -2,18	7E Tor Tre Teste -1,55
	12N S. Palomba -2,02	13X Castel Porziano -2,24	5D Tiburtino Sud -1,71
	8D Acqua Vergine -2,35	10I Barcaccia -2,33	10F Osteria del Curato -1,77
	15E Magliana -2,48	5I Sant'Alessandro -3,11	12M Castel Romano -2,46

Nota: cfr. Grafico 7.

Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

La regressione delle variabili elettorali con i tre assi fattoriali risulta poco interessante in generale, poiché il coefficiente di determinazione R^2 non raggiunge mai valori rilevanti eccetto le due elezioni del 2008 (comunali 0,45 e provinciali 0,41). Tuttavia, l'interesse di questa analisi risiede nella sintetica individuazione dei fattori alla base delle differenze elettorali tra zone, e quindi del consenso più o meno maggiore raccolto dai due schieramenti. Si conferma il ragionamento teorico dei paragrafi precedenti: a vantaggio del centrosinistra vi sono l'età dei residenti e degli edifici, legati alla densità e alla centralità dei quartieri; a vantaggio

del centrodestra vi sono invece lo status socio-economico elevato e il lavoro autonomo, soprattutto nelle zone a nord e sud della città dove si concentrano maggiormente i redditi superiori; con effetti più ambigui la quota di stranieri, single e non residenti, che appaiono particolarmente numerosi al centro, dove entrambi gli schieramenti ottengono risultati in linea con la media romana.

Tali considerazioni sono confermate dalla regressione dei risultati elettorali con le variabili di contesto, mediante il semplice metodo dei minimi quadrati ordinari sui valori standardizzati. Risulta una rilevante capacità di spiegare le differenze nel voto tra zone urbanistiche, particolarmente inattesa quando si analizzano dinamiche elettorali e socio-economiche, con l' R^2 compreso tra 74,1 per le regionali 2010 e 80,3 per le regionali 2005 (Tav. 8).

Sia pure considerando alcuni elementi di difformità individuati nel paragrafo 3 fra le diverse elezioni, in generale risultano alcune variabili ricorrenti e particolarmente significative. Con un impatto positivo sul voto al centrosinistra: la densità di popolazione, gli addetti al settore trasporti e comunicazioni, gli addetti al settore scolastico e la quota di abitazioni occupate da residenti⁴³. Con un effetto negativo: la densità di popolazione al quadrato (coerentemente con la correlazione quadratica rilevata nel paragrafo 4), gli imprenditori e professionisti, i coniugati nonché (eccetto nel 2006 e 2008) le casalinghe. La particolarità delle comunali 2006 e dei referendum 2011 consiste nella non significatività né della densità né del suo quadrato, unici casi tra le sei elezioni analizzate, mostrando al contrario l'effetto positivo (nel 2006) della fascia d'età tra 30 e 45 anni, degli affitti e della licenza media, nonché (nel 2011) della fascia d'età tra 45 e 60 anni, della licenza elementare e degli studenti. Per i referendum 2011 va anche rimarcato il forte effetto negativo degli addetti all'industria, che nelle regionali 2005 erano invece positivamente correlati col centrosinistra, sia pure più debolmente.

Più incerto l'impatto della presenza degli stranieri, che sembra dipendere dalla definizione che si dà: nelle due elezioni del 2008 e nel 2011 è significativa la provenienza dai paesi dell'Europa dell'est (in negativo), mentre nel 2006 e nel 2005 sono significative sia la religione islamica (in positivo) che la provenienza da paesi non Ocse (in negativo). Tale evidenza trova riscontro nella differente localizzazione degli stranieri in base alla nazionalità, sia pure non configurandosi una vera e propria segregazione territoriale sull'esempio londinese o parigino. Sono infatti gli islamici a risiedere maggiormente nei quartieri dove il centrosinistra è forte, sia

⁴³ Abitazioni occupate da persone che hanno dimora abituale nelle stesse, con esclusione quindi degli immobili vuoti, ad uso turistico od occasionale, ecc.

quelli più centrali, in particolare Centro Storico e Trastevere (dove rappresentano oltre il 10% dei residenti) e in misura minore l'Esquilino, sia quelli della periferia storica a sud-est (Quadraro, Tor Fiscale, Tor Pignattara). Al contrario, gli est europei risiedono in gran parte nei quartieri dove prevale il centrodestra, soprattutto nella periferia anulare o fuori dal Gra (Tor San Giovanni, Magliana, Decima, Ponte Galeria), e con una certa frequenza persino nei quartieri di ville dove evidentemente esercitano lavori domestici (Cesano, Appia Antica nord e sud).

TAV. 8

REGRESSIONE DEI VOTI DEL CENTROSINISTRA IN FUNZIONE DELLE
VARIABILI DI CONTESTO, CON LA DENSITÀ DI POPOLAZIONE

Variabili esplicative	Referendum affluenza 2011	Regione Presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune Sindaco 2006	Regione Presidente 2005
<i>Urbanistica</i>						
Densità di popolazione		0,60 (3,92)**	0,53 (3,46)**	0,55 (3,45)**		0,50 (3,35)**
Densità di popolazione al quadrato		-0,40 (-2,83)**	-0,35 (-2,60)**	-0,38 (-2,71)**		-0,33 (-2,54)*
<i>Occupazione</i>						
Imprenditori e professionisti	-0,63 (-7,57)**	-0,63 (-8,24)**	-0,55 (-8,10)**	-0,65 (-9,18)**	-0,45 (-3,90)**	-0,36 (-3,47)**
Addetti a trasporti e comunicazione	0,26 (5,33)**	0,30 (5,65)**	0,22 (4,35)**	0,14 (2,67)**	0,16 (2,82)**	0,19 (3,71)**
Addetti alla scuola	0,54 (7,93)**	0,56 (6,40)**	0,58 (7,56)**	0,53 (6,60)**	0,65 (6,69)**	0,61 (6,40)**
Addetti al settore pubblico		-0,11 (-2,37)*				
Addetti all'industria	-0,45 (-8,46)**					0,12 (2,20)*
Studenti	0,13 (2,34)*					
Casalinghe	-0,26 (-4,17)**	-0,19 (-3,70)**				-0,17 (-3,60)**
<i>Fasce d'età e stato civile</i>						
Età tra 30 e 45 anni					0,30 (4,93)**	0,22 (4,01)**
Età tra 45 e 60 anni	-0,16 (-3,36)**					
Coniugati		-0,43 (-6,83)**	-0,40 (-6,38)**	-0,36 (-5,47)**	-0,30 (-3,60)**	-0,36 (-5,61)**

./.

Variabili esplicative	Referendum affluenza 2011	Regione Presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune Sindaco 2006	Regione Presidente 2005
<i>Nazionalità</i>						
Stranieri dall'est Europa	-0,17 (-3,32)**		-0,24 (-4,49)**	-0,32 (-5,80)**		
Stranieri da paesi islamici					0,42 (4,62)**	0,30 (3,92)**
Stranieri da paesi non Ocse					-0,46 (-3,38)**	-0,35 (-3,87)**
<i>Istruzione</i>						
Licenza media					0,23 (2,07)**	0,52 (5,00)**
Licenza elementare	0,25 (2,48)*					
<i>Abitazioni</i>						
Abitazioni occupate da residenti	0,27 (5,68)**	0,29 (5,45)**	0,16 (3,15)**	0,12 (2,27)*	0,28 (4,63)**	0,33 (6,16)**
Abitazioni in affitto					0,13 (2,30)*	
<i>R² corretto</i>	78,7	74,1	77,1	74,8	74,4	80,3
<i>Numero di osservazioni</i>	143	143	143	143	143	141

Nota: coefficienti della regressione lineare OLS sui valori standardizzati; variabile dipendente = affluenza alle urne o voto per il candidato di centrosinistra; statistica *t* tra parentesi; ** = significativo all'1%, * = significativo al 5%.

Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

La capacità di spiegare le differenze elettorali migliora inserendo al posto della densità di popolazione i fattori che, secondo la logica discussa nel paragrafo 5, possono rappresentare *proxy* della disponibilità e del consumo di beni comuni o relazionali. Non sembra credibile infatti che sia la densità in sé a influire così nettamente sulle scelte elettorali, ma piuttosto altri fattori di interrelazione sociale che sono legati alla densità stessa e che vanno meglio esplicitati. Vengono utilizzate dunque l'incidenza *pro capite* delle medie e grandi superfici di vendita, ossia della grande distribuzione organizzata (GDO), e la densità di piazze per ettaro; nel primo caso è atteso un impatto negativo, dovuto alla sostituzione dei piccoli esercizi di quartiere raggiungibili facilmente a piedi e capaci di creare frequenti occasioni di incontro e scambio, e nel secondo caso positivo grazie alla maggiore disponibilità di spazi pubblici.

Adottando sempre il metodo dei minimi quadrati ordinari, l' R^2 raggiunge valori compresi tra 78,9 per le comunali 2006 e 86,4 per i referendum 2011, confermando comunque gran parte delle evidenze già discusse (Tav. 9). Con un impatto positivo sul voto al centrosinistra troviamo la densità di piazze per ettaro, la quota di abitazioni occupate da residenti, gli addetti nei settori trasporti e comunicazioni,

scuola e industria, i single (compresi gli anziani vedovi); con un effetto negativo le medie e grandi superfici di vendita, gli imprenditori e professionisti, le casalinghe. Le due nuove variabili inserite sono sempre significative, eccetto le piazze in due occasioni, e hanno i segni attesi. Appaiono invece poco robuste perché meno ricorrenti le considerazioni relative ad altre variabili: la quota di addetti nel settore edilizia (impatto negativo dal 2008 al 2011) e finanza o servizi alle imprese (negativo nel 2005 e 2006), le fasce d'età (negativo per i minori di 15 anni nel 2008 e 2011, positivo tra 30 e 45 anni in quattro elezioni), la quota di divorziati (negativo dal 2008 al 2011), la presenza di stranieri provenienti dall'Europa dell'est o paesi non Ocse (negativo rispettivamente in due e tre elezioni), la licenza elementare (positivo nel 2008 e 2010) e la percentuale di case in affitto (positivo dal 2005 al 2008). Si noti come l'inserimento di variabili che rappresentano *proxy* del capitale sociale *bridging* renda irrilevante la variabile che avevamo individuato come misura dei legami *bonding*, ossia la percentuale di coniugati.

TAV. 9

REGRESSIONE DEI VOTI DEL CENTROSINISTRA IN FUNZIONE DELLE
VARIABILI DI CONTESTO, CON GDO E PIAZZE COME *PROXY* DEI BENI
COMUNI E RELAZIONALI

Variabili esplicative	Referendum affluenza 2011	Regione presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune sindaco 2006	Regione presidente 2005
<i>Beni comuni e relazionali</i>						
Medie e grandi superfici di vendita (GDO)	-0,48 (-12,96)**	-0,43 (-8,88)**	-0,34 (-7,15)**	-0,33 (-6,79)**	-0,33 (-6,12)**	-0,40 (-8,07)**
Piazze per ettaro		0,21 (4,00)**	0,18 (3,38)**	0,17 (2,98)**		0,20 (3,84)**
<i>Occupazione</i>						
Imprenditori e professionisti	-0,46 (-8,10)**	-0,42 (-5,43)**	-0,42 (-5,76)**	-0,48 (-7,77)**	-0,35 (-3,59)**	-0,33 (-2,96)**
Addetti a trasporti e comunicaz.	0,19 (4,63)**	0,27 (5,91)**	0,21 (4,76)**	0,17 (3,66)**	0,12 (2,37)*	0,20 (4,35)**
Addetti alla scuola	0,33 (4,20)**	0,43 (5,23)**	0,45 (5,71)**	0,44 (5,12)**	0,53 (5,65)**	0,64 (7,09)**

./.

Variabili esplicative	Referendum affluenza 2011	Regione presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune sindaco 2006	Regione presidente 2005
Addetti all'industria		0,29 (4,55)**	0,28 (4,65)**	0,31 (5,53)**	0,20 (3,19)**	0,42 (6,35)**
Addetti all'edilizia	-0,17 (-2,84)**	-0,18 (-2,20)*	-0,16 (-2,02)*			
Addetti a finanza e servizi alle imprese					-0,26 (-2,60)*	-0,23 (-2,20)*
Casalinghe	-0,24 (-5,97)**	-0,38 (-6,93)**	-0,27 (-5,14)**	-0,19 (-4,34)**		-0,28 (-5,76)**
<i>Fasce d'età e stato civile</i>						
Età <15 anni	-0,23 (-4,02)*		-0,15 (-2,53)*	-0,40 (-5,77)**		
Età tra 30 e 45 anni	0,24 (4,44)**			0,27 (4,22)**	0,60 (6,13)**	0,17 (3,37)**
Età tra 45 e 60 anni					0,28 (4,63)**	
Età tra 60 e 75 anni					0,18 (2,24)*	
Divorziati	-0,25 (-4,99)**	0,15 (2,76)**	0,14 (2,60)*			
Single	0,18 (2,45)*	0,38 (4,59)**	0,35 (4,43)**	0,39 (5,00)**	0,42 (5,45)**	0,43 (5,93)**
<i>Nazionalità</i>						
Stranieri dall'est Europa			-0,12 (-2,55)*			-0,15 (-2,85)**
Stranieri da paesi islamici					0,44 (5,73)**	
Stranieri da paesi non Ocse	-0,18 (-4,38)**			-0,18 (-4,01)**	-0,59 (-6,57)**	
<i>Istruzione</i>						
Laurea						-0,48 (-2,76)**
Licenza elementare		0,41 (4,49)**	0,33 (3,70)**			
<i>Abitazioni</i>						
Abitazioni occupate da residenti	0,28 (5,78)**	0,40 (8,64)**	0,28 (5,59)**	0,28 (5,49)**	0,40 (6,54)**	0,46 (7,93)**
Abitazioni in affitto				0,11 (2,65)**	0,14 (3,23)**	0,14 (3,38)**
<i>R² corretto</i>	86,4	82,9	84,8	82,7	78,9	84,7
<i>Numero di osservazioni</i>	143	143	143	143	143	141

Nota: coefficienti della regressione lineare OLS sui valori standardizzati; variabile dipendente = affluenza alle urne o voto per il candidato di centrosinistra; statistica *t* tra parentesi; ** = significativo all'1%, * = significativo al 5%.
Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

Per chiarire il ruolo del reddito *pro capite* come possibile variabile esplicativa, è stata infine condotta un'altra regressione utilizzando non più le zone urbanistiche ma le 36 aree individuate nell'indagine campionaria dell'Istituto Tagliacarne (2005). Per nessuna elezione il reddito appare significativo, come in effetti era lecito attendersi vista l'assenza di correlazione individuata già nel paragrafo 4 (Tav. 10). Pur tenendo conto che le unità di analisi sono in questo caso differenti, meno numerose (36 contro 143) e anche meno omogenee, per le altre variabili indipendenti viene confermato in gran parte quanto detto in precedenza, se escludiamo le elezioni comunali 2006 per le quali la capacità esplicativa appare limitata. Con un impatto positivo emerge soprattutto la densità di popolazione, mentre hanno un persistente effetto negativo la densità al quadrato, la quota di casalinghe, gli addetti a finanza e servizi alle imprese, la quota di stranieri provenienti dall'Europa dell'est. Considerazioni analoghe valgono per le regressioni con la disponibilità di piazze, che per semplicità non vengono riportate.

TAV. 10

REGRESSIONE DEI VOTI DEL CENTROSINISTRA IN FUNZIONE DELLE
VARIABILI DI CONTESTO, CON IL REDDITO *PRO CAPITE*

Variabili esplicative	Referendum affluenza 2011	Regione presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune sindaco 2006	Regione presidente 2005
<i>Urbanistica</i>						
Densità di popolazione			1,12 (4,95)**	1,22 (5,14)**		1,16 (4,78)**
Densità di popolaz. al quadrato			-0,69 (-3,60)**	-0,77 (-3,54)**		-0,70 (-3,30)**
<i>Occupazione</i>						
Tasso di attività		-0,17 (-2,10)*				
Tasso di lavoro dipendente				0,25 (2,96)**	0,88 (8,34)**	0,27 (3,34)**
Imprenditori e professionisti	-0,95 (-17,65)**					
Addetti al commercio		-0,45 (-3,54)**	-0,64 (-4,52)**			
Addetti a trasporti e comunicaz.		0,19 (3,01)**				
Addetti all'industria						0,42 (2,58)*

./.

Variabili esplicative	Referendum affluenza 2011	Regione presidente 2010	Comune ballottaggio 2008	Provincia ballottaggio 2008	Comune sindaco 2006	Regione presidente 2005
<i>Addetti a finanza e servizi alle imprese</i>						
			-0,47 (-2,45)*	-0,44 (-3,21)**		-0,73 (-4,67)**
Casalinghe	-0,34 (-6,34)**	-0,49 (-7,05)**	-0,27 (-4,66)**	-0,22 (-2,94)**		-0,29 (-3,92)**
<i>Fasce d'età e stato civile</i>						
<i>Età <15 anni</i>						
			-0,20 (-2,46)*			
<i>Età >65 anni</i>						
	0,33 (4,50)**					
<i>Componenti del nucleo familiare</i>						
		-0,72 (-6,68)**			-0,51 (-4,81)**	
<i>Single con figli</i>						
	-0,25 (-4,04)**					
<i>Nazionalità</i>						
<i>Stranieri dall'est Europa</i>						
	-0,71 (-9,21)**	-0,67 (-8,91)**	-0,48 (-4,51)**	-0,61 (-4,54)**		-0,80 (-6,18)**
<i>Istruzione</i>						
<i>Laurea</i>						
			-0,69 (-4,27)**			
<i>Licenza media</i>						
		1,48 (9,50)**				
<i>R² corretto</i>	93,8	93,2	92,7	88,5	66,9	89,8
<i>Numero di osservazioni</i>	36	36	36	36	36	36

Nota: coefficienti della regressione lineare OLS sui valori standardizzati; variabile dipendente = affluenza alle urne o voto per il candidato di centrosinistra; statistica *t* tra parentesi; ** = significativo all'1%, * = significativo al 5%.

Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

7. - Osservazioni conclusive

C'è quindi una correlazione significativa tra la densità di popolazione, la disponibilità di beni comuni o relazionali e le scelte elettorali. Con tutte le cautele che comporta l'isolamento di alcuni fattori nell'analisi del comportamento politico, si può comunque dire che l'elettore di centrosinistra abita per lo più in quartieri densi, dove cioè sono più ravvicinate e ricche le relazioni sociali, e dove sono maggiormente accessibili gli spazi pubblici e collettivi; al contrario, gli insediamenti sparsi fuori dal Gra privilegiano i comportamenti più isolati e individuali dell'elettorato di centrodestra. La città contemporanea è soprattutto del secondo

tipo e questo non poteva non far sentire la sua influenza nel lungo periodo. Si è trattato di un processo lento e costante che solo sul finire del quindicennio di amministrazioni di centrosinistra ha manifestato tutti i suoi effetti fino a tradursi nella vittoria del centrodestra: un processo che è andato a scontrarsi con l'immagine scintillante del "modello romano", che pur corrispondendo a una crescita reale dell'economia cittadina, trascurava l'accentuarsi del dualismo territoriale, sociale e culturale.

Il presunto imborghesimento del centrosinistra non sembra una ragione sufficiente a spiegare tale dinamica. Come evidente dalla scarsa significatività del reddito *pro capite*, infatti, il suo consenso politico non è limitato alle fasce sociali più ricche, ma comprende sia i quartieri benestanti e centrali sia le zone popolari della prima periferia, che mantengono un reddito medio-basso rispetto alla media romana, e considerazioni analoghe valgono per il centrodestra. Sicuramente i due schieramenti hanno cambiato fisionomia elettorale rispetto ai decenni passati, con il centrosinistra che ha guadagnato consensi nei quartieri con elevato benessere perdendo al contempo voti nelle zone più difficili e marginali, e con il centrodestra che ha allargato la propria azione politica fuori dai quartieri tradizionalmente borghesi. Ma i dati non mostrano una vera e propria inversione nella capacità attrattiva, quanto piuttosto una certa omogeneizzazione in relazione alle caratteristiche socio-economiche, e di conseguenza le differenze attuali appaiono maggiormente legate a fattori urbanistici e relazionali che non all'istruzione o al reddito.

Il centrosinistra sembra prevalere laddove sono disponibili e condivisi i beni comuni, e di conseguenza vi sono numerose opportunità di produrre e consumare beni relazionali, e al contrario il centrodestra ottiene un maggiore consenso nelle periferie che hanno una limitata disponibilità di spazi collettivi. Ma perché la carenza e lo scarso uso dei beni collettivi nelle periferie danneggia più il centrosinistra che il centrodestra? Per qualunque schieramento si voti, in effetti, si rimane sempre lontani dal Campidoglio, che riesce a controllare e intervenire meglio nelle aree più centrali dove è maggiore l'attenzione mediatica, rispetto alle periferie amministrate da Municipi senza poteri effettivi e con poche risorse. Tuttavia, almeno due ragioni inducono a ritenere che sia soprattutto il centrosinistra a subire gli effetti negativi dello *sprawl* urbano.

Il primo elemento è che ovviamente il centrodestra non era identificato, almeno fino alla vittoria di Alemanno nel 2008, con la gestione dei servizi pubblici nel quindicennio precedente, quando l'Amministrazione comunale e gran parte dei Municipi sono stati ininterrottamente guidati dal centrosinistra che – dopo la fine dell'esperienza del pentapartito con Franco Carraro e grazie al nuovo sistema di

elezione diretta del Sindaco – aveva portato al Campidoglio prima Francesco Rutelli (dal 1993 al 2001) e poi Walter Veltroni (dal 2001 al 2008). Ma il secondo elemento è più strutturale, e riguarda le forme di attività politica sul territorio, molto più in sintonia con la tradizione dei partiti organizzati e radicati da cui nasce l'odierno centrosinistra che alle modalità mediatiche e "leggere" dell'attuale centrodestra. I candidati del centrosinistra e le strutture territoriali che li appoggiano appaiono maggiormente legati a forme di campagna elettorale organizzata nei quartieri, che però è tanto più difficile attuare quanto più i residenti non passeggiano per le strade perché salgono sull'automobile appena usciti da casa, non frequentano piazze e spazi pubblici perché non ce ne sono, non entrano nei negozi di vicinato perché si dirigono piuttosto verso i grandi centri commerciali.

Di conseguenza, per usare le parole di Cerasa (2009, pp. 10-12), il centrosinistra è sembrato abbandonare «il contatto con la città fuori dalle mura [...] perdendo progressivamente il controllo e il consenso delle zone più periferiche», al contrario dell'interesse del centrodestra verso i problemi reali o percepiti delle periferie, a cominciare dalla sicurezza e dal degrado. Gli elettori hanno avvertito da parte del centrosinistra una maggiore attenzione verso i quartieri più centrali in termini di progetti, opere ed eventi: «una gestione asserragliata al centro, concentrata eccessivamente sulla creazione di un pacchetto di eventi spettacolari [...] a cui solo una piccolissima parte della città aveva davvero la possibilità di accedere».

APPENDICE STATISTICA

TAV. 11

SUDDIVISIONI TOPONOMASTICHE E ZONE URBANISTICHE

Area urbana	Classificazione secondo le suddivisioni toponomastiche	Classificazione secondo le zone urbanistiche dei Municipi
Centro	Rioni: tutti Quartieri: 1-5, 6 e 14 parte, 15, 17, 18, 32	1 tutto, 2 tutto, 3 tutto, 11 x, 12 a, 17 tutto, 20 abdx
Periferia storica	Quartieri: 6 parte, 7-13, 14 parte, 16, 19-25, 27-31 Zone: 1, 7 parte	4 abcghi, 5 abcdgh, 6 tutto, 7 abg, 8 a, 9 tutto, 10 abcd, 11 abcdefg, 12 b, 15 abc, 16 adx, 18 ad, 19 abf
Periferia anulare	Quartieri: 26, 33-35 Suburbi: tutti Zone: 2, 4, 7 parte, 8, 12, 15, 18, 21, 22, 24, 27, 28, 31-33, 35, 40, 44, 50, 53, 56	4 defl, 5 ef, 7 cdefh, 8 b, 10 ef, 12 cdefgx, 13 bcfg, 15 def, 16 bc, 18 bc, 19 cde, 20 cef
Periferia <i>extra</i> -Gra	Zone: 3, 5, 6, 9-11, 13, 14, 16, 17, 19, 20, 23, 25, 26, 29, 30, 34, 41, 45-49, 51, 52, 54, 55, 57, 58	4 mno, 5 il, 8 cdefgh, 10 ghilx, 11 y, 12 hilmn, 13 adehix, 15 g, 16 ef, 18 ef, 19 gh, 20 ghilmno

Fonte: elaborazione dell'autore sulla base di TOCCI W. (2008, pp. 74-82).

ZONE URBANISTICHE DI ROMA



Fonte: elaborazione dell'autore su dati di Roma Capitale - Ufficio statistico.

TAV. 12

VARIABILI ELETTORALI E DI CONTESTO UTILIZZATE NELLE
ANALISI STATISTICHE

Variabile	Unità di misura	Fonte*	Anno	Media	Deviazione standard
<i>Variabili elettorali:</i>					
<i>voto per il centrosinistra</i>					
Regionali 2005	% voti	ELET	2005	54,1	6,2
Comunali 2006	% voti	ELET	2006	59,9	6,4
Comunali ballottaggio 2008	% voti	ELET	2008	44,3	7,0
Provinciali ballottaggio 2008	% voti	ELET	2008	48,9	6,8
Regionali 2010	% voti	ELET	2010	52,3	7,0
Affluenza al referendum 2011	% elettori	ELET	2011	59,0	6,1
Variazione affluenza 2005-10	Δ	ELET	2005-10	-12,3	6,7
Variazione % iscritti 2005-10	Δ	ELET	2005-10	3,0	15,3
<i>Variabili di contesto socio-economico, demografico e urbanistico</i>					
Abitazioni in affitto	% abitazioni	CENS	2001	26,2	12,6
Abitazioni occupate da residenti	% abitazioni	CENS	2001	87,5	5,5
Addetti a finanza e servizi alle imprese	% occupati	CENS	2001	15,5	4,6
Addetti a trasporti e comunicazioni	% occupati	CENS	2001	7,4	2,1
Addetti al commercio	% occupati	CENS	2001	18,5	4,2
Addetti all'edilizia	% occupati	CENS	2001	5,9	2,4
Addetti all'industria	% occupati	CENS	2001	11,7	3,9
Addetti alla sanità	% occupati	CENS	2001	8,1	2,0
Addetti alla scuola	% occupati	CENS	2001	7,8	2,7
Addetti pubblici	% occupati	CENS	2001	14,6	4,8
Casalinghe	% residenti	CENS	2001	16,6	2,6
Componenti del nucleo familiare	Numero	CENS	2001	2,8	0,7
Coniugati	% pop.	ANAG	2007	47,3	3,7
Densità di popolazione	Ab./ettaro	ANAG	2007	66,4	64,1
Densità di popolazione al quadrato	Ab./ettaro	ANAG	2007	8.491,1	13.804,2
Diplomati	% residenti	CENS	2001	34,3	5,5
Divorziati	% residenti	ANAG	2007	2,6	0,7
Età <15	% residenti	ANAG	2007	13,5	2,8
Età 15-30	% residenti	ANAG	2007	14,4	2,2
Età 30-45	% residenti	ANAG	2007	25,8	4,0
Età 45-60	% residenti	ANAG	2007	20,8	1,9
Età 60-75	% residenti	ANAG	2007	16,4	3,2
Età >75	% residenti	ANAG	2007	9,1	3,9
Imprenditori e professionisti	% residenti	CENS	2001	9,2	5,4
Laureati	% residenti	CENS	2001	13,5	9,1

./.

Variabile	Unità di misura	Fonte*	Anno	Media	Deviazione standard
Licenza elementare o nessun titolo	% residenti	CENS	2001	25,9	7,6
Licenza media	% residenti	CENS	2001	26,3	6,2
Medie e grandi strutture di vendita (GDO)	Numero/ab.	PROV	2009	1,1	5,3
Piazze per abitante	Numero/ab.	PROV	2009	0,6	0,9
Piazze per ettaro	Numero/ett.	PROV	2009	36,8	56,6
Reddito pro capite	€	TAGLIA	2003	19.552	4.083
Single	% residenti	ANAG	2007	26,1	8,3
Stranieri est-europei	% residenti	ANAG	2007	3,6	2,8
Stranieri islamici	% residenti	ANAG	2007	1,9	1,7
Stranieri non Ocse	% residenti	ANAG	2007	8,5	4,8
Studenti	% residenti	CENS	2001	8,7	1,7
Tasso di attività	% residenti	CENS	2001	57,9	4,6
Tasso di occupazione	% residenti	CENS	2001	45,2	4,4
Variazione % dei residenti	Δ	ANAG	2010/2007	13,1	19,2

(*) Fonte: ANAG = Ufficio anagrafico di Roma Capitale, CENS = Censimento Istat, ELET = Servizi elettorali di Roma Capitale, PROV = Provinciattiva, TAGLIA = Istituto Tagliacarne (per 36 aree); cfr. nota 23.

TAV. 13

ANALISI IN COMPONENTI PRINCIPALI SU TRE ASSI
DELLE VARIABILI DI CONTESTO

Variabile	Comunalità	Pesi sulla matrice ruotata delle componenti		
		1° - Status socioeconomico	2° - Anzianità	3° - Carenza di legami famil.
Abitazioni in affitto	0,21	-0,14	0,44	0,08
Abitazioni occupate da residenti	0,59	-0,02	0,36	-0,68
Addetti a finanza e servizi alle imprese	0,85	0,85	0,30	0,18
Addetti a trasporti e comunicazioni	0,32	-0,17	-0,15	-0,52
Addetti al commercio	0,68	-0,75	-0,25	-0,24
Addetti all'edilizia	0,76	-0,70	-0,50	-0,17
Addetti all'industria	0,44	-0,54	-0,29	-0,27
Addetti alla sanità	0,23	0,35	0,25	0,22
Addetti alla scuola	0,84	0,65	0,45	0,47
Addetti pubblici	0,11	0,29	0,12	-0,07
Casalinghe	0,64	-0,78	0,19	-0,08
Componenti del nucleo familiare	0,59	-0,33	-0,69	0,11
Coniugati	0,79	-0,16	-0,28	-0,83
Densità di popolazione	0,62	0,17	0,76	0,12
Densità di popolazione al quadrato	0,45	0,05	0,66	0,11

./.

Variabile	Comunalità	Pesi sulla matrice ruotata delle componenti		
		1° - <i>Status</i> socioeconomico	2° - Anzianità	3° - Carenza di legami famil.
Diplomati	0,80	0,89	-0,08	0,00
Divorziati	0,59	0,50	0,25	0,53
Età <15	0,77	-0,31	-0,80	-0,15
Età 15-30	0,44	-0,42	-0,41	-0,30
Età 30-45	0,62	-0,36	-0,69	0,07
Età 45-60	0,26	0,46	0,19	-0,08
Età 60-75	0,70	0,38	0,74	-0,04
Età >75	0,86	0,29	0,84	0,28
Imprenditori e professionisti	0,77	0,65	0,08	0,58
Laureati	0,89	0,78	0,19	0,49
Licenza elementare o nessun titolo	0,94	-0,95	-0,06	-0,19
Licenza media	0,86	-0,78	-0,14	-0,48
Medie e grandi strutture di vendita (GDO)	0,05	-0,08	-0,08	-0,19
Piazze per abitante	0,22	0,18	-0,04	0,43
Piazze per ettaro	0,59	0,23	0,40	0,61
Single	0,83	0,40	0,51	0,64
Stranieri est-europei	0,58	-0,54	-0,41	0,34
Stranieri islamici	0,58	-0,30	0,05	0,70
Stranieri non Ocse	0,71	-0,28	-0,18	0,77
Studenti	0,46	0,66	-0,14	0,06
Tasso di attività	0,78	0,39	-0,79	-0,01
Tasso di occupazione	0,72	0,50	-0,66	-0,19
<i>Autovalori</i>	-	13,2	5,0	4,0
<i>Varianza spiegata (%)</i>	-	35,6	13,4	10,8

Nota: cfr. Grafico 7.

Fonte: elaborazione dell'autore sui dati della Tavola 12 in Appendice.

BIBLIOGRAFIA

- AGENZIA PER IL CONTROLLO E LA QUALITÀ DEI SERVIZI PUBBLICI LOCALI DEL COMUNE DI ROMA, *Indagine sulla mobilità e il trasporto pubblico a Roma*, Roma, 2009.
- .-, *Quarta indagine sulla qualità della vita e i servizi pubblici locali nella città di Roma*, Roma, 2011.
- ANDERLINI F., «Il mito dell'espansione leghista», Bologna, *il Mulino*, n. 5, 2009.
- .-, «PD e dintorni visti dall'Emilia», *Tam Tam Democratico*, n. 1, 2011.
- ANNUNZIATA S. - COSSU M., *Roma oltre il Piano: forme di urbanità per la città contemporanea*, atti della XIII conferenza della Società Italiana degli Urbanisti "Città e crisi globale: clima, sviluppo e convivenza", Roma, 2010.
- ASSOCIAZIONE ITALIANA CASE (AIC) - UNIONE BORGATE, *Periferie di mezzo. Condizione sociale, economica e territoriale nei quartieri ex abusivi di Roma*, Roma, 2010.
- BERDINI P., *La città in vendita. Centri storici e mercato senza regole*, Roma, Donzelli, 2008.
- BONCIANI M., «La qualità della vita a Roma: fattori ambientali, socio-economici e psicosociali», *Sociologia e ricerca sociale*, n. 89, 2009.
- BRUNI L. - ZAMAGNI S., *Economia civile. Efficienza, equità, felicità pubblica*, Bologna, il Mulino, 2004.
- BURKHALTER L. - CASTELLS M., *Beyond the Crisis: Towards a New Urban Paradigm*, atti della IV conferenza dell'International Forum on Urbanism "The New Urban Question: Urbanism Beyond Neo-Liberalism", Amsterdam, 2009.
- CALAFATI A.G., «Apprendimento collettivo e sviluppo locale», in CAMAGNI R.P. - CAPPELLO R. (a cura di), *Apprendimento collettivo e competitività territoriale*, Milano, Franco Angeli, 2002.
- CAMAGNI R. - GIBELLI M.C. - RIGAMONTI P., *I costi collettivi della città dispersa*, Firenze, Alinea, 2002.
- CASTRONOVÌ A., «Introduzione. Lavoro, disagio sociale e questione urbana», in CANNATA B. - CARLONI S. - CASTRONOVÌ A. (a cura di), *Le periferie nella città metropolitana*, Roma, Ediesse, 2008.
- CAUDO G., «L'abitare e la città al tempo della crisi», in BIANCHI D. - ZANCHINI E. (a cura di), *Ambiente Italia 2011. Il consumo di suolo in Italia. Rapporto annuale di Legambiente*, Milano, Edizioni Ambiente, 2011.
- CAUDO G. - FIORETTI C. - SEBASTIANELLI S., *La questione casa. Competenze, dinamiche e territorio*, contributo per il rapporto CREL, Roma, 2009.
- CELLAMARE C., *Ways of Inhabiting the "Hypermarket City". Bufalotta and the Porta di Roma Shopping Mall*, mimeo, 2011.
- CENSIS, *La povertà nel Lazio*, rapporto per la Regione Lazio, 2002.
- CERASA C., *La presa di Roma*, Milano, Rizzoli, 2009.

- COMUNE DI ROMA, *Vivere Roma. Un'indagine sulla qualità della vita percepita dai cittadini romani*, Roma, 2007.
- COTTINO P., *La città impreveduta. Il dissenso nell'uso dello spazio urbano*, Milano, Eleuthera, 2003.
- CREMASCHI M., *Rendita fondiaria e sviluppo urbano nella riqualificazione urbana: per un'ipotesi interpretativa*, atti del XXIV Convegno della Società Italiana di Scienza Politica, Venezia, 2010.
- CRISCI M., «Roma, una città a spicchi. Sistemi urbani della mobilità pendolare e domiciliare», in SONNINO E. - BERTINO S. - CASACCHIA O. - CRISCI M. - D'ORIO G. - ROSATI R., *Popolazione e previsioni democratiche nei municipi di Roma Capitale. Dinamiche attuali e previsioni fino al 2024. Città e Municipi*, Roma, Gangemi, 2011.
- DE MURO P. - MONNI S. - TRIDICO P., «Knowledge-Based Economy and Social Exclusion: Shadows and Lights in the Roman Socio-Economic Model», Università di Roma Tre - Dipartimento di Economia, *Working Paper*, no. 91, 2008.
- DONOLO C., «Reti as Commons», *Parolechiave*, no. 34, 2005.
- EURES, *Rapporto 2010 sullo stato delle Province del Lazio*, rapporto per UPI Lazio, 2010.
- EUROPEAN ENVIRONMENT AGENCY (EEA), *Urban Sprawl in Europe. The Ignored Challenge*; rapporto di ricerca con il Joint Research Centre della Commissione Europea, Copenhagen, 2006.
- FERRAROTTI F. - MACIOTI M.I., *Periferie da problema a risorsa*, Roma, Teti, 2009.
- FRANZINI M., «Disuguaglianze economiche e non solo: l'Italia del malessere sociale», *la Rivista delle Politiche Sociali*, n. 4, 2007.
- .-, *Ricchi e poveri. L'Italia e le disuguaglianze (in)accettabili*, Milano, Università Bocconi Editore, 2010.
- FRUNCILLO D., *Politica senza reti. L'Italia al voto nel 2006 e nel 2008*, Roma, Ediesse, 2010.
- GIBELLI M.C. - SALZANO E. (a cura di), *No Sprawl*, Firenze, Alinea, 2006.
- GUI B., «Più che scambi, incontri. La teoria economica alle prese con i fenomeni interpersonali», in SACCO P.L. - ZAMAGNI S. (a cura di), *Complessità relazionale e comportamento economico. Materiali per un nuovo paradigma di razionalità*, Bologna, il Mulino, 2002.
- INSOLERA I., *Roma moderna. Da Napoleone I al XXI secolo* (nuova edizione ampliata), Torino, Einaudi, 2011.
- ISTAT, *Popolazione residente e abitazioni nei grandi comuni italiani. Roma*, 14° Censimento generale della popolazione e delle abitazioni, Roma, 2006.
- .-, «Distribuzione del reddito e condizioni di vita in Italia. Anni 2008-2009», *Statistiche in Breve*, 29 dicembre 2010.
- ISTITUTO TAGLIACARNE, *Reddito e consumi delle famiglie nei comuni del Lazio e nei quartieri di Roma*, rapporto per la Regione Lazio, 2005.

- LUCCIARINI S. - VIOLANTE A., «Il modello mediterraneo di housing e welfare e il caso di Roma», *La Rivista delle Politiche Sociali*, n. 3, 2006.
- .-, -.-, «Trasformazioni socio-economiche e mutamento della città: il caso di Roma», *Argomenti*, n. 19, 2007.
- MALTESE C. - PEREGO F. (a cura di), *Le Porte di Roma. I nuovi centri dello scambio urbano*, Comune di Roma - VII Dipartimento, 1996.
- MANIGRASSO M., «Il consumo di suolo nei comuni di Roma e Fiumicino», in BIANCHI D. - ZANCHINI E. (a cura di), *Ambiente Italia 2011. Il consumo di suolo in Italia. Rapporto annuale di Legambiente*, Milano, Edizioni Ambiente, 2011.
- OSTROM E., *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press, 1990.
- PROVINCIA DI ROMA, *Capitale metropolitana, periferie comuni. Rapporto di studio sulla condizione sociale, economica e territoriale*, a cura di Provinciattiva, Roma, 2010.
- REALE L., «Roma. Densità, spazio pubblico, paesaggio», *Edilizia Popolare*, vol. 53, 2010, pp. 284-285.
- SABATINI F., «Social Capital and the Quality of Economic Development», *Kyklos*, vol. 61, no. 3, 2008.
- .-, «Social Capital as Social Networks: A New Framework for Measurement and an Empirical Analysis of Its Determinants and Consequences», *Journal of Socio-Economics*, vol. 38, no. 3, 2009a.
- .-, «Il capitale sociale nelle regioni italiane: un'analisi comparata», *Rivista di Politica Economica*, vol. 99, n. 2, 2009b.
- SALZANO E., *Crisi dello spazio urbano o fine (morte) delle città?*, atti del convegno «Le settimane della politica - Ma cos'è questa crisi», Torino, 2010.
- .-, «L'habitat dell'uomo bene comune», in CACCIARI P. (a cura di), *La società dei beni comuni. Una rassegna*, Roma, Ediesse, 2011.
- SITI W., «Non riesco a immaginare un borgataro riformista», in ILARDI M. - SCANDURRA E. (a cura di), *Ricominciamo dalle periferie. Perché la sinistra ha perso Roma*, Roma, ManifestoLibri, 2009.
- SONNINO E. - BERTINO S. - CASACCHIA O. - CRISCI M. - D'ORIO G. - ROSATI R., *Popolazione e previsioni demografiche nei municipi di Roma Capitale. Dinamiche attuali e prospettive fino al 2024. Città e Municipi*, Roma, Gangemi, 2011.
- TOCCI W., «La città del tram», in TOCCI W., INSOLERA I., MORANDI D., *Avanti c'è posto. Storie e progetti del trasporto pubblico a Roma*, Roma, Donzelli, 2008.
- .-, «L'insostenibile ascesa della rendita urbana», *Democrazia e Diritto*, n. 1, 2009.
- .-, «Disuguaglianze metropolitane», *Tam Tam Democratico*, n. 0, 2011.
- TRUGLIA F.G., «La configurazione spaziale del consenso elettorale nel Comune di Roma», *Sociologia e Ricerca Sociale*, n. 81, 2006.
- VELTRONI W., «Il "modello Roma"», *Italianieuropei*, n. 1, 2006.

VIOLANTE A., *La metropoli spezzata. Sviluppo urbano di una città mediterranea*, Milano, Franco Angeli, 2008.

ZAMAGNI S., «Libertà economica e mercato civile: il caso dell'impresa sociale», *Atlantide*, n. 3, 2005a.

-.-, «L'economia civile e i beni relazionali», in VIALE R. (a cura di), *Le nuove economie: dall'economia cognitiva a quella evolutiva oltre i fallimenti dell'economia neoclassica*, Milano, Il Sole 24 Ore, 2005b.

PAROLE CHIAVE

Industria postale e corriere espresso

Francesca Mazzarella

Nera Economic Consulting, Rome

Vincenzo Visco Comandini

“Tor Vergata” University of Rome
and

Nera Economic Consulting, Rome

Postal and delivery services are an important sector of the economy. While express courier services are booming, traditional postal services are gradually declining because of the e-substitution effect. It is vital for the sector to foster competitive playing level fields also by granting access to the network. However postal service regulation concerning access should not be confused or identified with telecommunication regulation. Differently from telecom, postal services are a highly labor intensive industry where the last mile delivery is not an essential facility .

[JEL Classification: L87; L43; L51].

Keywords: postal and delivery services; network access; price regulation.

* <francesca.mazzarella@nera.com>; ** <visco.comandini@economia.uniroma2.it>.

Gli autori desiderano ringraziare Lina Magrane e Carlo Scarpa per i commenti ricevuti.

1. - Introduzione

I servizi postali e di corriere espresso rappresentano oggi una parte significativa dell'economia dei paesi industrializzati: il fornitore postale nazionale è quasi dappertutto il primo datore di lavoro del Paese, il corriere espresso svolge un ruolo significativo sui processi di riorganizzazione delle imprese, contribuendo ad accrescerne la competitività sia nel settore manifatturiero che esporta sia nella distribuzione finale delle merci, in cui sempre più spesso sta prevalendo la logica del *just in time*, ossia un magazzino ridotto al minimo grazie ad un sistema di rifornimenti pressoché continuo nell'arco della giornata.

I due settori, i servizi postali e quelli di corriere espresso, presentano ormai da anni andamenti divergenti: in declino costante i primi, soprattutto a causa del fenomeno dell'*e-substitution*, ovvero la progressiva sostituzione elettronica della posta cartacea (di cui esempi sono le lettere singole sostituite da *email*, estratti conto e fatture inviati *on-line*); in crescita accentuata, nonostante la crisi economica, i secondi.

A subire la sostituzione elettronica sono infatti le lettere e lo scambio dei documenti, non i pacchi che anzi sono cresciuti in modo sostanziale grazie all'*e-commerce*, che costituisce un beneficio positivo sui sistemi economici nazionali. In quasi tutti i paesi, tuttavia, il mercato dei pacchi non è gestito come le lettere dagli operatori postali nazionali, ma in netta prevalenza dalle imprese di corriere espresso (incluse quelle, come vedremo al paragrafo 3, di proprietà degli stessi operatori). Ciò pone agli operatori postali, e quindi ai regolatori chiamati a garantire la competitività di un settore solo di recente liberalizzato, sfide complesse che impongono scelte selettive, capaci di mantenere elevata la concorrenza e l'equilibrio economico-finanziario degli operatori.

Sia perché condivide un quadro concettuale comune, sia perché spesso esercitata da chi ha svolto esperienza su quel settore, la regolazione dell'industria postale appare culturalmente e tecnicamente vicina a quella delle telecomunicazioni. I due settori, fino a due decenni fa, erano gestiti in tutti i paesi da un unico soggetto, il Ministero delle Poste e Telecomunicazioni che ne regolava i mercati sottoposti a monopolio pubblico. Le Autorità di regolazione e quelle *antitrust* hanno sviluppato competenze ed *expertise* facendo esperienza proprio nel processo di liberalizzazione delle telecomunicazioni, i cui mercati erano peraltro in forte crescita.

In quasi tutti i paesi europei, e dalla fine dello scorso anno anche in Italia, le competenze sul settore postale sono state affidate all'Autorità di regolazione delle telecomunicazioni. Naturale quindi che vi sia un travaso concettuale e di impostazione metodologica da queste ultime ai primi.

In questo lavoro si cercherà di evidenziare su quali aspetti TLC e poste hanno profili regolatori comuni e su quali invece divergono. Il problema non è banale, perché la regolazione di settore può presentare forti connotati di *path dependence*, in cui scelte iniziali finiscono per condizionare quelle successive, con scarse possibilità di recuperare gli eventuali errori commessi¹.

In particolare, uno degli errori dovuti ad un'erronea identificazione di fattispecie è quello della teoria del *ladder of investments*, sviluppata in relazione al mercato delle telecomunicazioni mobili e dei servizi broadband, che numerose evidenze empiriche hanno verificato non essere appropriata al settore postale. La Commissione Europea è caduta in quest'errore quando ancora nel 2005 sosteneva la sua applicabilità all'industria postale in via di liberalizzazione². La ragione è che nel settore postale i costi fissi non sono identificabili, come nella maggior parte delle industrie a rete, con gli investimenti fissi in impianti e macchinari, ma hanno a che fare con il dimensionamento della rete (a parità della quale vanno conseguite economie di densità) e, con particolare ancorché non unico riferimento al fornitore universale nazionale, anche con i relativi obblighi di qualità. A differenza delle tlc, nel settore postale non vi è una ripartizione di natura tecnologica fra costi fissi e costi variabili, essendo questi ultimi quelli preponderanti (costo del personale, *outsourcing*).

Ovvero per espandersi un concorrente non ha bisogno di realizzare rilevanti investimenti fissi irrecuperabili, bensì di conseguire economie di scala (far crescere la propria rete riducendo i costi unitari) grazie ad un oculato mix di utilizzo degli *input*, gran parte dei quali è variabile. Ulteriore prova di tale differenza è data dal target obiettivo di dimensionamento della rete che gli operatori concorrenti si danno: nel caso dei servizi radiomobili, la copertura dell'intero territorio nazionale in modo da evitare il roaming; nel settore postale, invece, un grado di copertura che in nessun paese ha mai superato³ il 70-75% degli indirizzi servibili o il 50% dei Cap.

Tlc e servizi postali presentano invece meccanismi simili nei sistemi di accesso alla rete, ma diversità di funzione e ruolo rivestito dalla regolazione su tale aspetto, come discusso nel paragrafo 5.

¹ Sul concetto di *path dependence* si veda LIEBOWITZ S. e MARGOLIS S. (1995).

² «Access can help facilitate market entry for upstream consolidators (...). New competitors who want to establish a delivery network can also use access for a transitional period to build up customer relationships and volumes, before being able to compete end to end with the incumbent. Using this model, consolidators and competitors, as well as major business customers can already carry out part of the value chain process before handing mail over to the incumbent» (COMMISSIONE EUROPEA, 2005, annex, p. 19).

³ In Svezia il mercato è stato liberalizzato nei primi anni Novanta e la concorrenza è stata fin dall'inizio significativa.

Un'altra, rilevante, differenza fra TLC e servizi postali è la dinamica del mercato, in continua crescita nei nuovi servizi nelle prime, in costante calo nei secondi. Nei servizi postali si impone una generale riconsiderazione strategica: considerato il prevedibile (e ormai ineludibile) calo dei volumi, per quanti anni gli operatori *incumbent* e concorrenti riusciranno a mantenersi finanziariamente sani? Il tema è affrontato in uno studio commentato nel paragrafo 6.

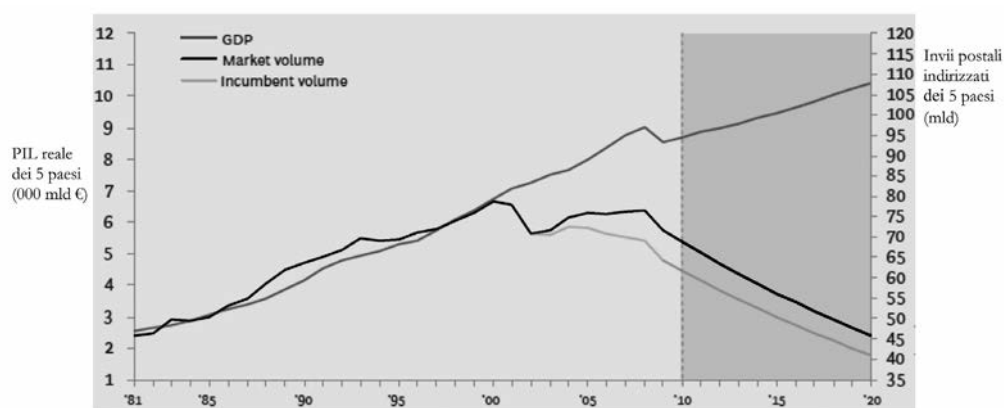
Dopo aver descritto le principali tendenze del mercato del corriere espresso (al paragrafo 8), il lavoro è concluso con l'analisi al paragrafo 9 del prodotto postale *direct mail* (la pubblicità per corrispondenza), visto all'interno della più generale competizione fra media per l'accesso alle risorse pubblicitarie, in cui solo in Italia la TV ha un ruolo di fortissima dominanza.

2. - L'e-substitution

In coincidenza con l'inizio del nuovo secolo il mercato postale delle lettere, che nella maggior parte dei paesi industrializzati era sempre cresciuto in relazione all'aumento del prodotto interno lordo, ha iniziato a mutare il proprio andamento. I volumi sono dapprima rimasti costanti, e da circa la metà del decennio hanno iniziato a seguire un andamento chiaramente decrescente (Graf. 1).

GRAF. 1

VOLUMI DI POSTA E PRODOTTO INTERNO LORDO
NEI 5 MAGGIORI PAESI EUROPEI

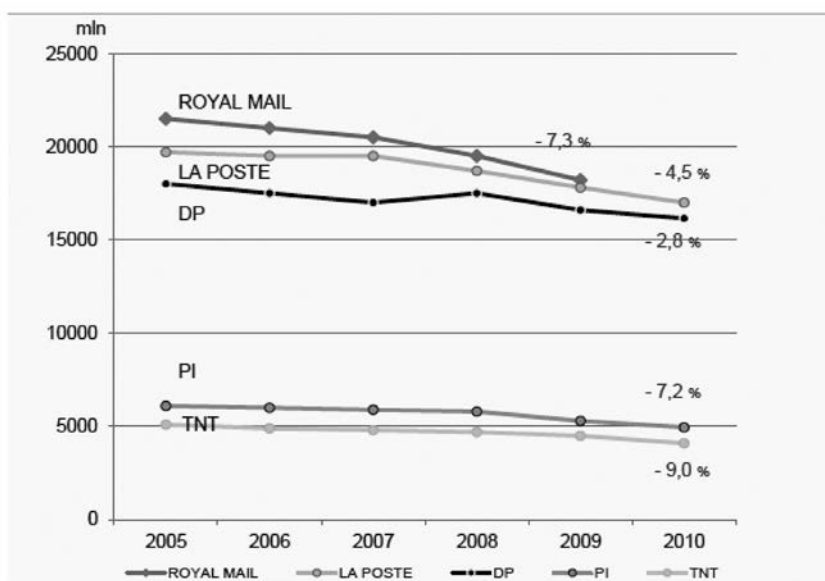


Fonte: BOSTON CONSULTING GROUP (2011) su dati UPU, IMF, *Economic Intelligence Unit*, ECORYS, bilanci degli operatori. Paesi: Francia, Germania, Italia, Regno Unito, Spagna.

Il decremento osservato dei volumi è relativamente lento ma costante, più accentuato nei paesi che hanno grandi volumi di posta e dove maggiore è la penetrazione dei servizi di Internet che accelerano la sostituzione. Il Grafico 2 indica la riduzione dei volumi dei maggiori operatori europei dal 2005 al 2010, peraltro dovuta anche alla bassa crescita economica e alla perdita di quote di mercato a favore dei concorrenti, in Germania, Olanda, Svezia e Italia (ma non negli altri Paesi) già presenti prima della direttiva europea di liberalizzazione completa del mercato, entrata in vigore nel 2011.

GRAF. 2

RIDUZIONE DEI VOLUMI DI POSTA INDIRIZZATA
DEI GRANDI OPERATORI EUROPEI



Fonte: bilanci degli operatori.

La sostituzione elettronica sembra avvenire con intensità diverse anche nei paesi di una medesima area geografica. Ad esempio le poste danesi, che due anni fa hanno unificato con quelle svedesi i loro servizi postali nella società PostNord, fra il 2000 e il 2011 hanno perso ben il 45% dei propri volumi, mentre nel mercato svedese la riduzione è stata di pochi punti percentuali.

Per l'Italia le stime più recenti (fonte: TNT) indicano una previsione di riduzione media annua (2010-2015) pari a circa:

- -1,3% per la posta *business* transazionale (*e-substitution* puro con estratti conto *on line*);

- -1,3% per posta prioritaria (*e-substitution* di *email, social network*);
- -1,8% per la posta registrata (tagli alla spesa pubblica spingono le amministrazioni a contenere il suo utilizzo, PEC);
- +1,4% per il *direct mail* (anche in parte per effetto spostamento volumi da editoria sempre meno sussidiata).

Come si può osservare, anche in Italia i volumi sono in calo in tutti i segmenti, con la sola eccezione del *direct mail* (invii pubblicitari per corrispondenza) che ha invece prospettive positive di crescita, come vedremo un'opportunità importante da sfruttare pur in presenza di vincoli normativi e concorrenziali rilevanti (discussi nel paragrafo 9).

Negli Stati Uniti l'*e-substitution* sta creando forti tensioni sul bilancio dell'operatore pubblico USPS nonostante questi goda di un esteso e mai seriamente discusso regime di monopolio nel recapito: fra il 2006 e il 2012 il calo dei volumi è stato pari al 25%, tale da spingere l'operatore a proporre di abolire la consegna della posta il sabato (provvedimento ad oggi fortemente contestato dal Congresso nonostante sia ormai la norma nella parte dei Paesi europei).

È quindi evidente che le prospettive di contrazione del mercato sono destinate a cambiare le strategie dei governi e delle imprese, che si trovano a fare i conti con redditività calante, a meno di non ridurre seriamente le dimensioni della rete e/o la qualità dei servizi prestati.

3. - I grandi operatori postali nazionali

La separazione, o meno, dei servizi postali dagli altri servizi (finanziari, assicurativi) è avvenuta con modalità diverse per i grandi operatori nazionali. Per esempio, Deutsche Post ha venduto a Deutsche Bank la propria società PostBank, mentre La Poste (operatore francese) l'ha invece mantenuta all'interno del proprio gruppo.

Nel 2005, l'Assemblea nazionale francese ha approvato la legge di regolazione delle attività postali. Questa legge ha da un lato recepito le prime due direttive europee del 97/67 e 02/39 sulla liberalizzazione della corrispondenza e dall'altro istituito la Banque Postale, un'istituzione finanziaria autonoma con licenza bancaria, prevista dal Contratto di programma siglato nel 2004. La Poste è molto attiva nel mercato del corriere espresso, attraverso Geopost e DPD che continua nella sua politica di partnership con operatori globali come FedEx, ma anche con operatori locali che le permettono di essere presente su mercati esteri come quello spagnolo, tunisino e dell'Europa Orientale.

Royal Mail è un'azienda autonoma controllata dallo Stato Britannico (attualmente è in discussione in Parlamento l'ennesimo progetto di privatizzazione parziale attraverso

un'IPO) che opera nel settore della corrispondenza e del corriere espresso. Negli ultimi tempi vi è stato un ampliamento dell'offerta di prodotti finanziari grazie alla *joint venture* con Bank of Ireland Group. Dal 2003 sono, inoltre, offerti prestiti personalizzati, polizze assicurative automobilistiche, prodotti di risparmio, carte di credito e mutui.

Come La Poste, anche l'operatore postale inglese utilizza un network internazionale prevalentemente europeo, General Logistics Systems (GLS), presente in 30 Paesi, che costituisce la gran parte dei ricavi del settore pacchi corriere espresso.

Il mercato postale svedese, totalmente liberalizzato da oltre dieci anni (1993), costituisce un'eccezione all'interno del mercato europeo. Nonostante la liberalizzazione, l'operatore postale Posten AB è riuscito, grazie un'attenta politica dei prezzi e, soprattutto, attraverso una strategia di contenimento dei costi ottenuta sfruttando ampie economie di scala (si pensi ai 700 invii *pro capite* annui in Svezia), a mantenere una quota di mercato di circa il 90%, lasciando al suo principale competitor, *CityMail* (oggi di proprietà delle poste norvegesi), una quota residuale. Attualmente, i concorrenti operano solo nelle aree urbane di Stoccolma, Malmö, Gotenborg e l'isola di Gotland.

La liberalizzazione non sembra aver portato benefici evidenti per il consumatore singolo: ha infatti generato una sostanziale diminuzione delle tariffe di *bulk mailing*, con forti risparmi per le aziende, mentre ha provocato un notevole aumento del prezzo dell'affrancatura singola utilizzata da consumatori e piccole imprese.

L'operatore postale tedesco Deutsche Post AG è uno dei pochi operatori privatizzati (attualmente circa il 70%) e quotato in borsa. Si è recentemente trasformato in un *global operator* che ha sviluppato una propria rete logistica internazionale dotata di infrastrutture terrestri e aeree (avendo acquistato la DHL) e opera nel settore dei trasporti e dei servizi accessori in favore di clienti stabiliti in tutto mondo.

Una scelta imprenditoriale simile è anche quella perseguita dal gruppo olandese TNT. L'operatore olandese è stato il primo a trasformarsi in operatore postale globale, sviluppando una strategia innovativa fondata su di un network internazionale attivo nei settori mail e corriere espresso. Nel 2011 è avvenuta la completa separazione proprietaria tra TNT PostNL (il fornitore del servizio universale postale olandese, con branch nazionali in Italia, Germania e Regno Unito) e TNT express (gruppo attivo nel mercato del corriere espresso).

Poste Italiane, il fornitore del servizio universale in Italia, presenta la caratteristica, unica nel mondo insieme alle poste giapponesi, di presentare un *deficit* strutturale nel settore postale e contemporaneamente realizzare ingenti utili nel settore finanziario e assicurativo. Si tratta di un sussidio incrociato legittimo, poiché non c'è alcun trasferimento vietato da settori in monopolio a quelli in concorrenza, ma certamente improprio, che presenta alcune criticità sia di ordine

proprietario che di rispetto delle norme *antitrust*. A livello di gruppo, Poste Italiane ha realizzato un profitto che nel 2011 è stato pari a 846 mln € di utile netto.

Poste Italiane, dalla sua trasformazione in S.p.a. avvenuta nel 1998, è riuscita a far crescere significativamente il risparmio postale (che finanzia la Cassa Depositi e Prestiti) con circa 6 milioni di correntisti, e diventare il secondo operatore nel ramo assicurazioni vita. Poste, infine, gestisce i servizi del primo operatore virtuale mobile in Italia, Postemobile, unico soggetto non infrastrutturato che è finora riuscito raggiungere il pareggio di bilancio.

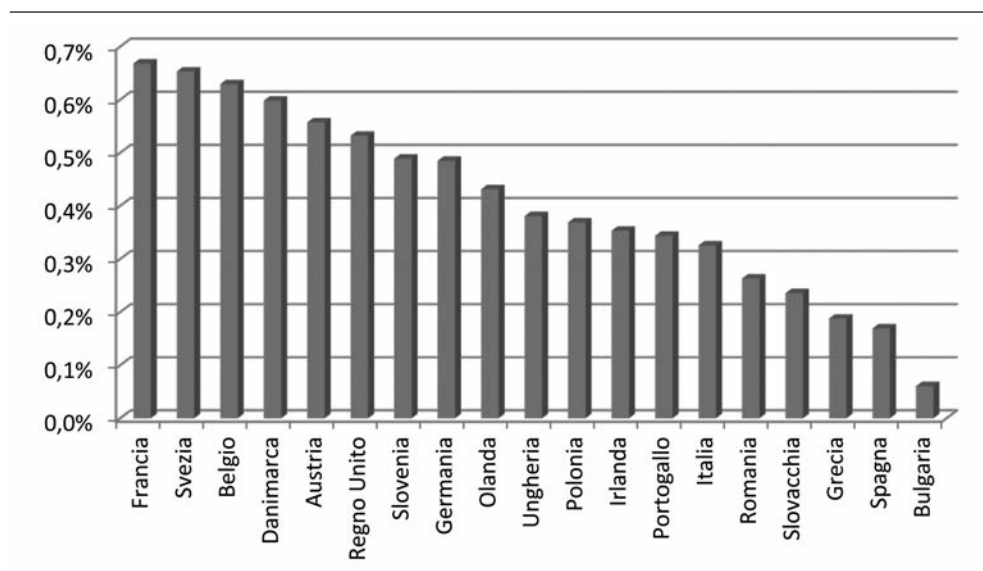
4. - Il mercato postale italiano

4.1 *Il limitato sviluppo del mercato e le sue determinanti*

Il mercato postale italiano è da sempre di dimensioni ridotte, circa 100 invii per abitante/anno contro una media di 300-400 invii degli altri grandi Paesi europei. Di conseguenza, anche l'incidenza del fatturato postale sul PIL è fra i più bassi d'Europa (Graf. 3).

GRAF. 3

INCIDENZA DEL FATTURATO POSTALE NAZIONALE SUL PRODOTTO INTERNO LORDO NEI PAESI EUROPEI
(anno 2010 - valori %)



Fonte: EUROSTAT.

Le ragioni di questo storico differenziale negativo sono strutturali, sia di tipo economico che culturale (Scarfiglieri e Visco Comandini, 2001; Fioramanti, 2006).

Fra i fattori economici, quello più evidente è lo scarso utilizzo del *direct mail* quale strumento di informazione pubblicitaria da parte delle imprese italiane, che preferiscono altri mezzi di comunicazione⁴ (Tav. 1).

TAV. 1

VOLUMI DI DIRECT MAIL INDIRIZZATO, NON INDIRIZZATO E TOTALE PER
1.000 ABITANTI NEI GRANDI PAESI EUROPEI
(anno 2011)

	indirizzato	non indirizzato	totale
Francia	0,06	0,15	0,21
Germania	0,08	0,05	0,13
Italia	0,02	0,01	0,03
Olanda	0,23	0,48	0,71
Regno Unito	0,07	0,15	0,22

Fonte: elaborazioni degli autori su dati dei bilanci degli operatori.

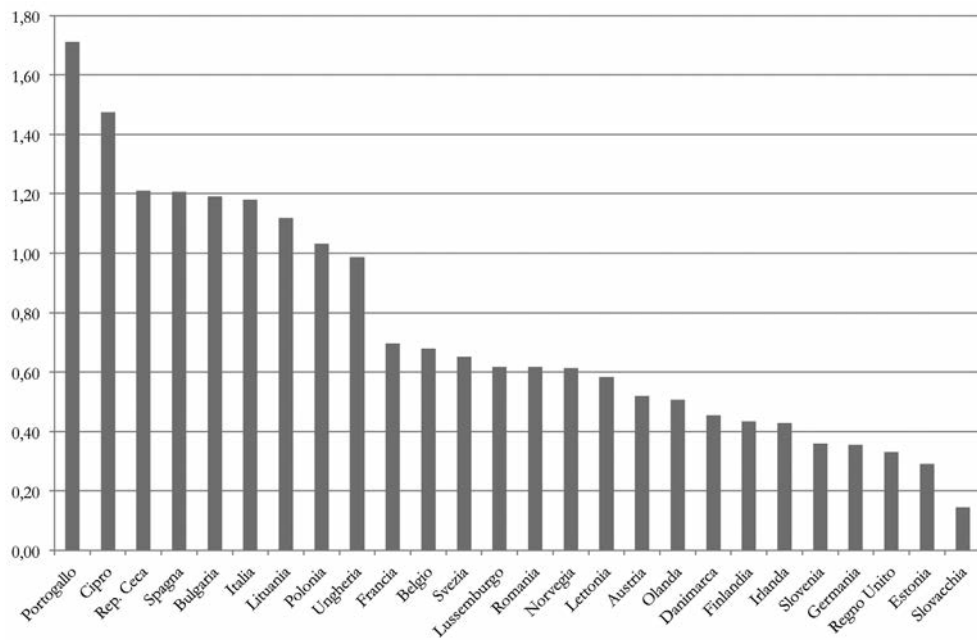
Qui il differenziale negativo rispetto ai grandi paesi è particolarmente accentuato: in Italia il *direct mail* indirizzato è utilizzato circa 10 volte di meno dell'Olanda, quattro volte meno di Francia e Regno Unito, tre volte meno che in Germania. Anche il non indirizzato (ovvero i volantini imbustati o non che troviamo nella nostra cassetta delle lettere) segue il medesimo andamento negativo, anche se il dato risente probabilmente di una sottostima dovuta all'elevata evasione fiscale registrata in questo mercato nel nostro Paese.

Il basso utilizzo del *direct mail* risale fin dalla sua nascita quando, intorno agli anni Sessanta, nacque e si sviluppò l'industria della vendita per corrispondenza (VPC) di cui il *direct mail* è il diretto complemento. Nel nostro Paese la VPC non ha mai conosciuto un reale momento di successo (molte aziende del settore sono ripetutamente fallite), anche perché i consumatori italiani hanno sempre mostrato una forte preferenza per gli acquisti diretti negli esercizi commerciali, che infatti sono da noi molto più numerosi che nei principali Paesi europei (Graf. 4).

⁴ si veda successivo paragrafo 9.

GRAF. 4

NUMERO DEGLI ESERCIZI COMMERCIALI PER 100 ABITANTI IN EUROPA
(anno 2006)

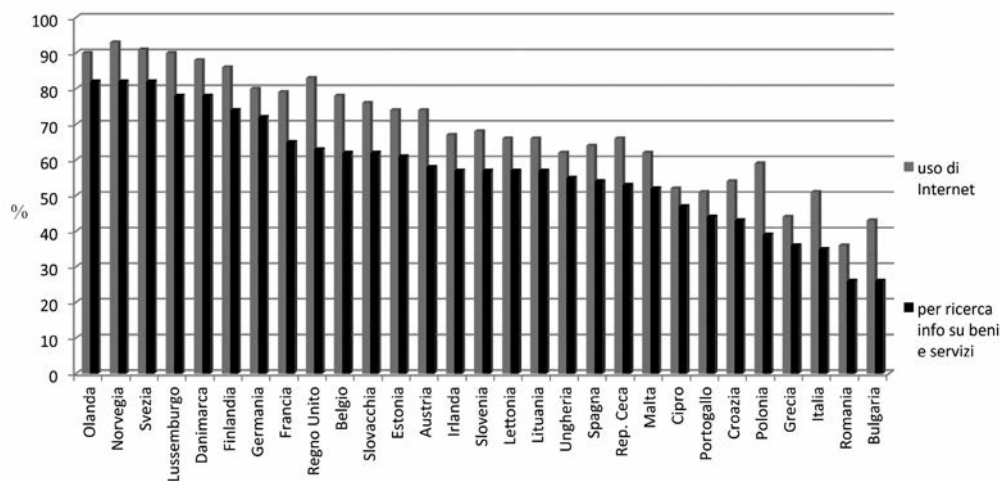


Fonte: EUROSTAT.

La scarsa attitudine degli italiani ad effettuare acquisti a distanza costituisce oggi un vincolo allo sviluppo dell'*e-commerce*, accentuato da un modesto uso che essi fanno di Internet, specie se al fine di informarsi su beni e servizi (Graf. 5).

GRAF. 5

INDIVIDUI 14-64 ANNI CHE UTILIZZANO INTERNET E CHE LO USANO PER
RICERCARE BENI E SERVIZI
(anno 2011 - valori %)

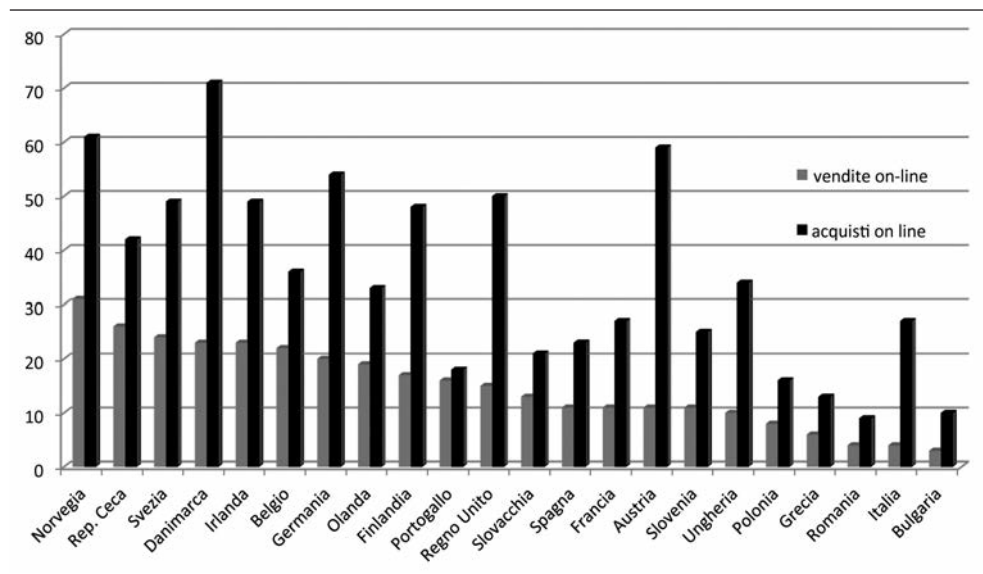


Fonte: EUROSTAT.

Tuttavia, più che da un problema di domanda da parte dei consumatori finali, il modesto sviluppo dell'*e-commerce* e, di conseguenza del *direct mail*, sembra essere determinato dalla mancata offerta di beni e servizi venduti *on-line* da parte delle imprese italiane, in fondo alla classifica europea (Graf. 6).

GRAF. 6

IMPRESE CHE ACQUISTANO E VENDONO ON-LINE IN EUROPA
(anno 2010)



Fonte: EUROSTAT.

La ristrettezza del mercato postale, infine, ha anche alcune specifiche ragioni culturali, quali una scarsa attitudine all'uso della comunicazione scritta⁵ e una antica diffidenza per essa⁶, opposta ad un alto utilizzo di quella vocale via cellulare (il numero di SIM attive ha superato in Italia i 120 mln).

L'insieme di questi fattori, tutti negativi, hanno sempre condizionato il mercato postale italiano che, come vedremo nel paragrafo 7, presenta un *deficit* strutturale, e hanno quindi reso la liberalizzazione del mercato più complessa da realizzare.

⁵ Fioramanti (2006) per spiegare il fenomeno cita un'indagine CIDI di replica di quella multi paese OECD, da cui emerge che poco più del 30% degli intervistati italiani è capace di leggere il testo di un messaggio, ma non è in grado di sintetizzarne il contenuto.

⁶ Lo scrittore ungherese Sándor Márai è vissuto per alcuni anni nella Napoli del dopoguerra. Ecco come ha registrato l'approccio tradizionale degli italiani alle lettere: «Per le gente di queste parti, che abita nel palazzo e nei dintorni, la posta non ha grande importanza. Sono pochi a saper leggere e scrivere, pertanto non fa loro piacere ricevere una lettera. D'altro canto, neanche chi sa leggere e scrivere si aspetta di ricevere posta. Per antica esperienza sanno che la posta porta di rado buone notizie. Il più delle volte si tratta di solleciti di pagamento, o di lettere ufficiali, insomma di cose che è meglio ignorare» (MÁRAI S., 2010, p. 30).

4.2 La concorrenza

Il mercato postale italiano è stato caratterizzato fin dagli anni Settanta da una forma assolutamente peculiare di concorrenza: le fatture commerciali erano fuori riserva perché il codice postale del 1973 le classificava come invii aperti diversi dalle lettere, mentre la posta *business* urbana (traffico interno alle città) a contenuto epistolare era di fatto liberalizzata grazie ad una speciale concessione assegnata ad alcune imprese sotto il vincolo, formale e mai rispettato, di non fare concorrenza sui prezzi a Poste Italiane. All'epoca, il recapito delle lettere raccomandate ed espresse recapitate in ambito urbano veniva inoltre appaltato alle medesime imprese private, tanto che alla fine degli anni Novanta, quando Poste Italiane fu trasformata in società per azioni, nella grande maggioranza dei capoluoghi di provincia risultavano attive numerose imprese concorrenti generalmente operanti in ambito locale.

Il D.lgs. 22 luglio 1999, n. 261 di trasposizione della prima direttiva sulla liberalizzazione del mercato postale (97/67) assegnava a Poste Italiane un livello di monopolio simile a quello della media degli altri paesi europei vietando esplicitamente la concessione di licenze speciali in monopolio a fornitori non concessionari del servizio universale. Tuttavia, ai concorrenti privati veniva consentito di fornire servizi di recapito di lettere di pubblicità diretta (*direct mail*), nonché di fissare accordi individuali con Poste Italiane per proseguire le attività di *outsourcing* svolte fino ad allora.

In Italia il D.lgs. n. 58 del 2011 ha trasposto nell'ordinamento nazionale la direttiva europea 6/2008, mantenendo una limitata area di riserva in favore di Poste Italiane nelle notifiche di atti amministrativi e delle multe stradali, e ha escluso dal servizio universale il *direct mail*.

La liberalizzazione del mercato conseguente al decreto ha rafforzato la concorrenza, che oggi appare piuttosto vivace e di intensità crescente. Nelle Tavole 2 e 3 è mostrata una stima del mercato delle lettere in Italia nel 2010 e 2011, rispettivamente in volumi e ricavi.

TAV. 2

IL MERCATO POSTALE DELLE LETTERE IN ITALIA
(anni 2010 e 2011 - volumi - mln. di invii)

prodotto	2010				2011			
	totale mercato	Poste Italiane	TNT	altri operatori	totale mercato	Poste Italiane	TNT	altri operatori
posta massiva e a data e ora certe	1.854	1.491	255	108	1.796	1.386	309	100
prioritario	1.118	1.118	0	0	1.029	1.029	0	0
direct mail	1.280	1.267	7	6	1.206	1.190	12	4
posta registrata	294	278	8	8	279	262	7	10
prodotti editoriali	674	674	0	0	552	552	0	0
totale mercato								
posta indirizzata	4.546	4.154	270	122	4.862	4.419	328	114

Fonte: stime degli autori su dati e informazioni Poste Italiane e TNT.

TAV. 3

IL MERCATO POSTALE DELLE LETTERE IN ITALIA
(anni 2010 e 2011 - ricavi - mln. €)

prodotto	2010				2011			
	totale mercato	Poste Italiane	TNT	altri operatori	totale mercato	Poste Italiane	TNT	altri operatori
posta massiva e a data e ora certe	944	828	82	35	884	753	99	32
prioritario	789	789	0	0	770	770	0	0
direct mail	318	315	2	1	309	305	3	1
posta registrata	1.155	1.123	16	16	1.132	1.097	14	21
prodotti editoriali	192	192	0	0	158	158	0	0
totale mercato								
posta indirizzata	3.398	3.247	99	52	3.094	2.925	116	53

Fonte: stime degli autori su dati e informazioni Poste Italiane e TNT.

Le Tavole 4 e 5 mostrano gli stessi dati espressi in quote di mercato, rispettivamente in volumi e in ricavi.

TAV. 4

QUOTE DI MERCATO SERVIZI POSTALI IN ITALIA IN VOLUMI
(anno 2011)

prodotto	2010			2011		
	Poste It	TNT	altri operatori	Poste It	TNT	altri operatori
posta massiva e a data e ora certe	80,4%	13,8%	5,8%	77,2%	17,2%	5,6%
prioritario	100,0%	0,0%	0,0%	100,0%	0,0%	0,0%
direct mail	99,0%	0,5%	0,5%	98,7%	1,0%	0,3%
posta registrata	94,7%	2,6%	2,7%	93,8%	2,5%	3,7%
prodotti editoriali	100,0%	0,0%	0,0%	100,0%	0,0%	0,0%
totale mercato posta indirizzata	91,4%	5,9%	2,7%	90,9%	6,7%	2,3%

Fonte: stime degli autori su dati e informazioni Poste Italiane e TNT.

TAV. 5

QUOTE DI MERCATO SERVIZI POSTALI IN ITALIA IN RICAVI
(anno 2011)

prodotto	2010			2011		
	Poste It	TNT	altri operatori	Poste It	TNT	altri operatori
posta massiva e a data e ora certe	87,7%	8,6%	3,7%	85,2%	11,2%	3,6%
prioritario	100,0%	0,0%	0,0%	100,0%	0,0%	0,0%
direct mail	99,1%	0,5%	0,4%	98,8%	0,9%	0,3%
posta registrata	97,3%	1,3%	1,4%	96,9%	1,2%	1,8%
prodotti editoriali	100,0%	0,0%	0,0%	100,0%	0,0%	0,0%
totale mercato posta indirizzata	95,6%	2,9%	1,5%	94,5%	3,7%	1,7%

Fonte: stime degli autori su dati e informazioni Poste Italiane e TNT.

Poste Italiane vede scendere di poco la sua quota totale di mercato in volumi da 91,4% a 90,9%, ma la perdita è quasi interamente concentrata nel *bulk mail* dove scende dall'80,4% al 77,2%. Più alta la quota controllata dall'*incumbent* espressa in ricavi, dove la perdita è limitata a circa un punto percentuale sul mercato totale, e di due punti e mezzo in quello *bulk*.

Per valutare il reale ribilanciamento del mercato a favore dei concorrenti è stata stimata la loro quota in riferimento ai soli prodotti dove c'è concorrenza (*bulk mail*, *direct mail* e posta registrata) e a quelle dei medesimi prodotti nelle aree dove la concorrenza riesce a recapitare in modo profittevole (Tav. 6).

TAV. 6

QUOTE DI MERCATO CONCORRENTI NEI SERVIZI POSTALI IN ITALIA
(anno 2011 - valori %)

	2010	2011	2010	2011
	volumi		ricavi	
quota del mercato totale	8,60%	9,10%	4,40%	5,50%
quota mercato prodotti contendibili	11,40%	13,50%	6,20%	7,30%
quota mercato geografico contendibile	20,80%	20,70%	13,80%	16,00%

Fonte: stime degli autori su dati e informazioni Poste Italiane e TNT.

Fra il 2010 e il 2011 i concorrenti hanno allargato questa area, passando da circa il 5% al 65% degli indirizzi servibili, così che la loro quota del mercato contendibile è rimasta sostanzialmente immutata (circa 20,75%) fra il 2010 e il 2011 se espressa in volumi, mentre è cresciuta quella in valore, dal 13,8% al 16%.

Questa differenza di aumento delle quote è dovuta al fatto che i concorrenti, allargando la loro operatività ad alcune porzioni dell'extraurbano, riescono ad accedere al prezzo più elevato, così come fissato per posta massiva di Poste Italiane, dove è più alto rispetto a quelli per i capoluoghi di provincia e per le aree metropolitane.

Il ribilanciamento delle quote di mercato in favore dei concorrenti ha raggiunto livelli simili a quelli osservati nei servizi di connessione a banda larga, dove il processo concorrenziale ha però impiegato più anni a consolidarsi.

L'aumento dei volumi gestiti dai concorrenti privati ha consentito a questi ultimi di aumentare l'area geografica del mercato contendibile, che oggi ha raggiunto come già osservato il 65% degli indirizzi servibili, corrispondenti a circa il 45% dei Cap nazionali⁷. Tale percentuale è destinata a crescere ulteriormente⁸, ma non a raggiungere il 100% della copertura nazionale perché vaste aree del Paese, per ragioni sia di basso reddito dei destinatari che di insufficiente densità abitativa, sono strutturalmente antieconomiche dal punto di vista del recapito, quindi raggiungibili unicamente da un servizio che necessita comunque di tutela pubblica di universalità.

⁷ La copertura territoriale per Cap della concorrenza è minore di quella per punti di recapito serviti, perché la popolazione e le imprese hanno una concentrazione maggiore nei Cap più densamente abitati.

⁸ In Germania e Svezia l'area servita dalla concorrenza ha raggiunto il 70% degli indirizzi.

5. - L'accesso alla rete

L'accesso alla rete nei servizi postali ha un significato tecnico-economico diverso rispetto a quello delle telecomunicazioni su rete fissa o mobile. Mentre nelle TLC garantire l'accesso all'ultimo miglio della rete dell'*incumbent* ai concorrenti è il principale strumento regolatorio proconcorrenziale nella fornitura dei servizi finali a banda larga, nell'industria postale non è altro che la concessione di sconti quantità e/o prelaborazione a favore dei grandi speditori di invii. Infatti, a differenza delle TLC o della distribuzione dei quotidiani a domicilio⁹, nei servizi postali la porzione equivalente dell'ultimo miglio, il recapito, non può essere tecnicamente negata dall'*incumbent* per due ragioni: La prima è da quando esiste il servizio, norme specifiche disciplinano gli obblighi di garantire l'accesso alla rete del fornitore universale a utenti singoli, grandi clienti e, da quando il mercato è stato liberalizzato, ai concorrenti. La seconda è che la rete, perché non è un'infrastruttura fisica cui si può concedere o meno la connessione (ovvero un'*essential facility*), ma solo la parte finale di un modello organizzativo di fornitura facilmente replicabile, specie in ambito urbano, essendo pressoché assenti costi sommersi o investimenti significativi in impianti fissi.

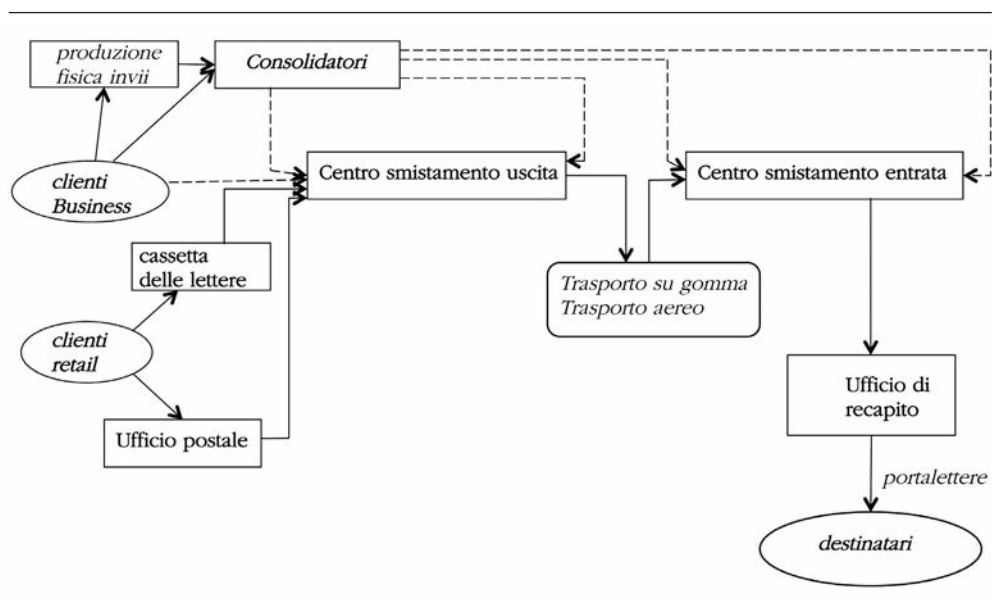
Nei servizi postali l'accesso alla rete dell'*incumbent* avviene in fase di accettazione di grandi quantità scontate (rispetto al prezzo pieno e/o *retail*) di posta *bulk* (ovvero la posta spedita da grandi clienti come banche, *utilities*, pubbliche amministrazioni, imprese, organizzazioni) da recapitare ai destinatari finali.

Nel Grafico 7 è mostrata la catena del valore nei servizi postali. La prima fondamentale distinzione riguarda i clienti, *retail* (coloro che inviano lettere singole), e *business*: i primi sono i principali e, in molti paesi, esclusivi fruitori del servizio universale e accedono al servizio tramite cassetta postale in strada o recandosi all'ufficio postale; i secondi, che rappresentano mediamente il 90-95% del traffico, producono in proprio le lettere *bulk* in formati *standard* e senza materiale affrancatura ovvero si avvalgono di centri stampa esterni, e immettono o direttamente o tramite consolidatori gli invii al livello dei centri di accettazione specializzati, che in Italia sono i centri di smistamento (CMP e CP).

⁹ Corte di Giustizia Europea, caso C-7/97, *Oscar Blonner vs. Mediaprint Zeitung-und Zeitschriften, Verlag* (1988) ECRI-7791. Per una discussione del caso vedi VISCO COMANDINI V. (2006, cap. 5).

GRAF. 7

LA CATENA DEL VALORE NEL SERVIZIO POSTALE



Tre diversi soggetti possono chiedere l'accesso alla rete postale:

- i) un concorrente operante nel recapito che chiede all'*incumbent* di consegnare gli invii dei suoi clienti destinati in aree geografiche non profittevoli a bassa densità, dove non svolgerà mai il servizio perché antieconomico;
- ii) un'impresa di consolidamento che opera nei mercati a monte del recapito;
- iii) un grande cliente.

Dal punto di vista dell'accesso, il concorrente *i*) appare all'*incumbent* come un grande cliente che accede agli sconti se e in quanto dispone dei requisiti qualitativi richiesti e, a meno di previsioni normative specifiche¹⁰, è da questi indistinguibile. Al contrario, *ii*) e *iii*) presentano caratteristiche fra loro in parte diverse che l'*incumbent* deve comunque considerare.

¹⁰ Ad esempio il criterio di accesso anti *cream-skimming* previsto nel Regno Unito: i volumi che gli operatori consegnano alla rete nazionale debbono rispettare determinate quote *standard* di ripartizione fra destinazioni urbane ed extraurbane. Nel caso non vengano rispettate, ovvero l'operatore recapiti in proprio solo quelle urbane consegnando le altre, il prezzo d'accesso sale in modo significativo. Tale previsione non risulta mai essersi verificata.

5.1 I consolidatori

Nei mercati *upstream* del recapito le imprese di consolidamento svolgono tipicamente una o più delle seguenti tipologie di servizi:

- a) attività antecedenti la parte propriamente postale, consistenti nella produzione fisica dell'invio (ad es. gestione dei file, stampa e imbustamento, cellofanatura, creatività del messaggio per il *direct mail*, ecc.);
- b) attività di prelaborazione postale (presmistamento, trasporto);
- c) attività di consolidamento dei flussi posta di più clienti al fine di raggiungere soglie di sconto.

Le attività del tipo a) costituiscono un mercato separato, non regolato dalla disciplina postale europea, mentre quelle b) e c) possono essere esercitate in forme giuridiche diverse (in conto proprio e in conto terzi) e sono finalizzate ad ottenere sconti quantità e/o prelaborazione.

Le norme del trattato¹¹ e della direttiva europea¹² impongono che venga offerto il medesimo trattamento a tutti i soggetti che si trovano nelle stesse condizioni. La non discriminazione riguarda, in particolare, grandi clienti e consolidatori a prescindere dal regime giuridico attraverso cui operano, in riferimento sia agli sconti prelaborazione che, con alcune specificità, a quelli quantità.

Nel Grafico 7 le linee continue mostrano i percorsi degli invii svolti in esclusiva dall'*incumbent* (o da un altro operatore verticalmente integrato), quelle tratteggiate le attività evitabili in quanto gestibili da terzi che danno origine a sconti. Si considerino le linee tratteggiate che partono dal soggetto "Consolidatori": quella più a sinistra indica che l'impresa di consolidamento si limita ad immettere gli invii prima dell'attività di smistamento in uscita al fine di accedere al solo sconto quantità; spostandosi a destra in senso antiorario troviamo il flusso di invii presmistati inseriti a valle del centro di smistamento in uscita, che consente di accedere a sconti sia quantità che prelaborazione di primo livello; proseguendo troviamo l'immissione degli invii già lavorati a monte del centro di smistamento in entrata, dove il consolidatore effettua in proprio anche il trasporto (sconti di prelaborazione di secondo livello); nel caso infine di trasporto e consegna di invii già presmistati per zona di destinazione (tipicamente il Cap) si ha la prelaborazione di terzo livello¹³.

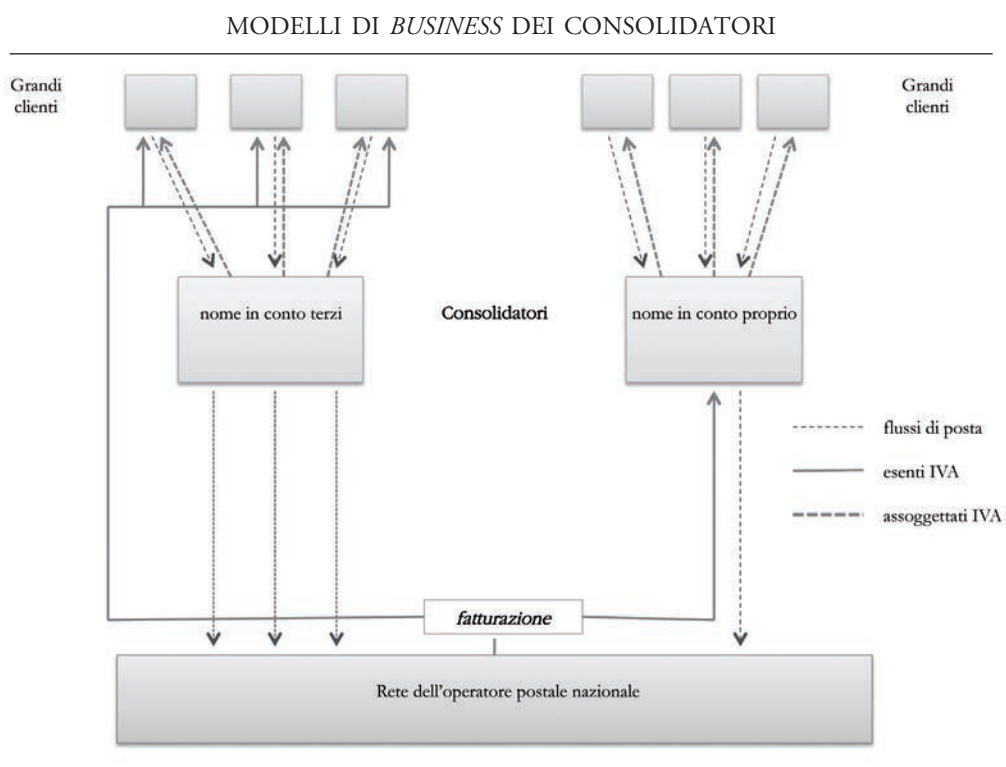
¹¹ Art. 102 TFUE (ex art. 82 TCE).

¹² Considerando 28 e artt. 7 (1) e 12 della direttiva europea 97/67.

¹³ Nel Grafico sono indicate le attività teoricamente prelaborabili da terzi, non quelle effettive che divergono per Paese, normativa vigente e prodotto. Ad esempio in Italia lo sconto del terzo tipo (smistamento per Cap e trasporto) non è previsto se non per alcune tipologie di prodotti editoriali.

Il Grafico 8 illustra i due modelli di *business* dei consolidatori: a nome in conto terzi e a nome in conto proprio. La scelta da parte del grande cliente fra i due regimi dovrebbe essere di puro *make or buy*, in cui questi sceglie il primo se intende mantenere il controllo del processo acquistando separatamente i servizi del consolidatore e dell'operatore della rete postale, il secondo se affida in *outsourcing* all'impresa di consolidamento il proprio flusso di posta che ne diventa soggetto responsabile.

GRAF. 8



La differenza fra i due modelli riguarda anche l'informazione da trasmettere al gestore della rete postale: operando in conto terzi, l'impresa di consolidamento è obbligata a comunicare l'identità dei suoi clienti all'operatore di rete, non così nel caso del conto proprio. L'efficienza di mercato richiederebbe una perfetta neutralità fra i due modelli, che però oggi, almeno in Italia, è violata dalla normativa fiscale vigente.

Se infatti il *business* si riferisce ad un prodotto postale universale (come ad esempio posta massiva in Italia) la scelta fra essi è distorta. Nel nome in conto

terzi l'*incumbent* emette una fattura esente Iva¹⁴ direttamente ai grandi clienti, mentre nel nome in conto proprio la fattura esente Iva è inviata al consolidatore, ma questi deve a sua volta riversarla ai clienti con Iva. Di conseguenza, tutti i grandi clienti che non raccolgono Iva (istituti di credito, assicurazioni, organizzazioni *no profit*, pubbliche amministrazioni) tendono a scegliere il primo modello perché non possono dedurla, anche se per essi potrebbe essere più efficiente affidare l'intero ciclo delle spedizioni postali in *outsourcing*¹⁵. Questa distorsione invece non si genera se il prodotto postale non è universale.

5.2 Sconti quantità

La tradizionale giustificazione economica degli sconti quantità è che i benefici derivanti dalle economie di scala godute dall'operatore postale nel recapito vengono trasferite al grande cliente speditore. Tuttavia, l'aumento della concorrenza, unitamente al calo strutturale dei volumi, ha accresciuto i rischi di un uso anti-competitivo di questo tipo di sconto, perché le imprese che ne beneficiano sono spinte a legarsi ad un solo operatore. Gli sconti quantità, unitamente a quelli di fatturato annuo minimo, sono quindi spesso visti con sospetto dalle Autorità Antitrust e di regolazione perché tendenzialmente leganti e suscettibili di discriminazioni di prezzo (Erpg, 2012).

Due sono i sistemi di sconto quantità generalmente praticati nell'industria postale: gli sconti retroattivi e quelli incrementali. Uno sconto retroattivo si applica all'intera quantità se questa supera un certo livello predefinito. Al contrario, lo sconto incrementale si applica in modo differenziato alle quantità incrementali che eccedono determinate soglie. Si consideri, ad es., un sistema di sconto incrementale con due soglie quantità $T1 < T2$, cui corrispondono sconti $ST1 < ST2$. Con tale sistema lo sconto concesso al cliente A che inserisce in rete una quantità di invii pari a $T1$ per la porzione $0 \rightarrow T1$ è uguale a quello del cliente B che accede con una quantità $T2$, ma quest'ultimo riceverà anche un ulteriore sconto $ST2$ per la porzione $T1 \rightarrow T2$.

Dal punto di vista concorrenziale lo sconto incrementale è quindi preferibile a quello retrospettivo perché tutti i soggetti, clienti o consolidatori sono trattati allo stesso modo a parità di quantità immesse in rete.

¹⁴ Ai sensi della decisione della Corte di Giustizia Europea caso C-357/07, *TNT Post UK Ltd v. The Commissioners for Her Majesty's Revenue and Customs*, (2009) ECR I-3025 i prodotti postali universali non possono essere assoggettati ad Iva, mentre quelli non universali debbono esserlo.

¹⁵ Alcuni concorrenti *end-to-end* di Poste Italiane preferiscono in alcuni casi operare in conto proprio, scontando il maggior costo Iva, pur di non comunicare all'*incumbent* i dati sui loro clienti.

Tuttavia, una recente sentenza della Corte Suprema francese¹⁶, sentiti anche i pareri della Autorità di Regolamentazione Arcep e del Conseil de la Concurrence, sembra consentire all'operatore postale, a condizione di offrire sconti incrementali, la possibilità di differenziare il trattamento offerto a consolidatori e grandi clienti. La decisione ha riconosciuto all'operatore La Poste il diritto a discriminare il trattamento sugli sconti quantità fra consolidatori e grandi clienti. La *ratio* è che gli sconti quantità concessi ai grandi clienti ne stimolano la domanda perché creano al margine un incentivo ad aumentare la quantità di invii da spedire, in particolare per le spedizioni di *direct mail* in cui l'estensione della campagna pubblicitaria è variabile nel tempo e nello spazio. Al contrario, secondo la Corte gli sconti quantità non stimolano la domanda delle imprese di consolidamento, perché queste non fronteggiano una curva di domanda classica, né sono in grado di trasferire l'incentivo ai piccoli clienti di cui hanno acquisito il flusso di posta per consolidarli.

Qui si evidenzia un'altra differenza nelle caratteristiche – e dunque nella regolamentazione – del mercato postale rispetto a quello delle TLC. In quest'ultimo l'apertura del mercato ai *reseller* di rete fissa e agli operatori mobili virtuali (gli equivalenti dei consolidatori postali) ha generato un benessere sociale positivo dovuto ai clienti a bassa disponibilità a pagare che, grazie ai prezzi contenuti dei nuovi servizi, ne sono diventati consumatori mentre prima ne erano esclusi. Al contrario, nel mercato postale la pura attività di intermediazione fra prezzi *wholesale* e *retail* anche a causa della domanda derivata – i consumatori finali ovvero i destinatari degli invii non sono gli acquirenti del servizio – non fa di per sé crescere il mercato generando benessere sociale, ma si traduce in una semplice redistribuzione di profitti fra soggetti che operano lungo la catena del valore.

5.3 Sconti prelaborazione

Gli sconti per prelaborazione (*worksharing*) sono largamente praticati dagli operatori postali, e in alcuni Paesi come il Regno Unito, la Nuova Zelanda e gli Stati Uniti, anche definiti da un sistema di prezzi regolati. In questi Paesi la percentuale di posta che accede alla rete con sconti *worksharing* è stimabile fra il 30% e il 40%. Negli altri Paesi, dove vige invece il sistema di sconti è lasciato alla contrattazione fra le parti, la percentuale è più bassa, intorno al 5% in Francia e al 2,5% in Germania, ancora inferiore negli altri (Copenhagen Economics, 2010).

¹⁶ Decisione 5 maggio 2009.

Contribuendo a migliorare l'efficienza di costo della catena del valore postale, gli sconti prelaborazione vanno applicati in modo non discriminatorio a tutti i soggetti, sia consolidatori che grandi clienti, che rispondano ai requisiti richiesti (Copenhagen Economics, 2010). In particolare, i *worksharing rebates* debbono essere neutrali rispetto ai due modelli di *business* dei consolidatori mostrati nel Grafico 2, come sancito da due recenti sentenze, rispettivamente della Corte di Giustizia Europea (casi *Vedat Deniz*, unificati C-287/06 e C-292/06) e dell'Autorità di regolamentazione postale tedesca Bundesnetzagentur (BNA, decisione del Settembre 2010), entrambe attinenti gli sconti prelaborazione nel settore postale.

Nel caso *Vedat Deniz*, la Corte di Giustizia Europea ha sostenuto che negli sconti per prelaborazione va concesso un uguale trattamento per condizioni simili a consolidatori e grandi clienti. Pertanto, per questi sconti, a differenza di quelli quantità, discriminazioni di prezzo di secondo tipo (ovvero un diverso trattamento per le due categorie) sono da considerarsi abusivi.

La decisione BNA del 2010 ha invece affrontato la questione del livello di sconto *workshered* applicato dall'operatore dominante Deutsche Post a clienti e consolidatori. Le poste tedesche erano state accusate dai concorrenti nel recapito non già di praticare uno sconto troppo basso, bensì uno eccessivo rispetto alle attività effettivamente evitate, tale da ridurre la convenienza dei grandi clienti ad avvalersi di operatori integrati alternativi all'*incumbent*. Il Regolatore tedesco BNA, dopo un'attenta analisi dei costi delle attività evitate di Deutsche Post, concluse che il prezzo d'accesso fissato non era da considerarsi predatorio, in quanto copriva una parte significativa dei costi fissi.

5.4 Sistemi efficienti di prezzi d'accesso

Armstrong (2001) ha definito le condizioni di efficienza dei prezzi d'accesso o Efficient Component Pricing Rule (Ecpr)¹⁷, sia quando le attività *downstream* sono un'*essential facility* (se nessuno compete nelle attività downstream del recapito), sia quando c'è concorrenza anche nella distribuzione finale della posta.

Se il recapito è svolto dal solo *incumbent* perché non c'è concorrenza in questo segmento, il prezzo d'accesso Ecpr può essere calcolato nel seguente modo:

$$(1) \quad C1 - C2 = P1 - P2$$

¹⁷ Il modello di Armstrong, che combina prezzi d'accesso e obblighi di servizio universale sotto forma di prezzi geograficamente uniformi, è stato pensato originariamente per le TLC, ed è stato adattato alle caratteristiche del settore postale da VISCO COMANDINI V. (2006).

dove:

$C1$ = costo del prodotto postale *end-to-end*

$C2$ = costo del prodotto parziale di solo recapito

$P1$ = prezzo del prodotto postale *end-to-end*

$P2$ = prezzo del prodotto parziale di solo recapito ($P1$ meno lo sconto)

ovvero:

$$(2) \quad P2 = C2 + (P1 - C1)$$

(1) mostra la stretta aderenza dei prezzi ai costi specie se come nel caso dell'industria postale i margini sono molto ridotti, e (2) identifica le componenti del prezzo d'accesso del prodotto parziale: costi dell'attività di *delivery* più la perdita di profitto subita dal gestore della rete nella concessione dell'accesso.

Se calcolato in modo corretto, ovvero vicino se non uguale all'Ecpr, lo sconto prelaborazione genera un incentivo all'efficienza lungo l'intera catena del valore: il consolidatore accederà alla rete solo se i suoi costi di produzione delle attività *upstream* al recapito sono inferiori a quelli dell'operatore postale verticalmente integrato. Direttamente a monte e indirettamente a valle si genera dunque una pressione competitiva con beneficio sociale positivo.

Il modello di Armstrong definisce il prezzo d'accesso efficiente anche quando il concorrente ha la possibilità di operare nel recapito, ovvero se la distribuzione di ultimo miglio della posta, nei mercati dove c'è concorrenza *end-to-end*, non è un'*essential facility*.

In questo caso è necessario utilizzare insieme al prezzo d'accesso una tassa unitaria sul prodotto capace di compensare le differenze geografiche di costo se il prezzo del servizio alla clientela finale è uniforme¹⁸.

Se dunque c'è possibilità di fare concorrenza sul recapito, gli incentivi sociali e quelli privati relativi all'accesso sono conciliati quando il prezzo d'accesso è uguale al costo e *contemporaneamente* è istituita una tassa t su tutti i soggetti operanti nelle porzioni geografiche profittevoli del mercato pari alla perdita unitaria subita nelle aree in perdita:

$$(3) \quad P2 = C2 \quad \text{e} \quad t = C1_p - P1$$

¹⁸ La medesima situazione vale se, pur vigendo una differenziazione geografica di prezzo, rimane (come in Italia) un costo netto positivo per il servizio universale tale da richiedere un fondo di compensazione.

dove:

t = accisa unitaria sul recapito nelle zone profittevoli

$C1_p$ = costo del prodotto postale *end-to-end* nelle aree in perdita

Un diverso modo di fissare il prezzo d'accesso implica necessariamente inefficienza, di cui vediamo due casi.

Il primo caso è quando, non essendoci concorrenza *end-to-end*, la rete è de facto un'essential facility. Se il prezzo d'accesso viene fissato in modo tale che $P2 < C2 + (P1 - C1)$, si genera un'inefficienza perché costi e prezzi divergono dal momento che il prezzo d'accesso è inferiore ai costi evitati $C1-C2$.

Il secondo caso è quando c'è concorrenza anche nel recapito, in cui ci sono due possibilità *a)* e *b)* di deviazione dal prezzo efficiente. In *a)* la deviazione avviene se i soggetti ammessi al contributo pagano t^{19} , ma $P2 < C2$. Infatti un concorrente non ha convenienza ad entrare nel mercato *end-to-end* anche quando ha propri costi di recapito tali che $c2 + t < C2$ e $c2 + t > P2$. In *b)* si genera inefficienza quando $P2 = C2$ ma il concorrente usa la rete dell'*incumbent* senza pagare t . Questi, infatti, ha un corretto incentivo nella scelta *make or buy* una volta entrato, ma un incentivo sbagliato ad entrare.

Questo semplice modello sembra descrivere efficacemente la realtà del mercato di alcuni paesi, in particolare quelli di Finlandia e Regno Unito.

In Finlandia vige una regolamentazione che ha istituito la tassa, fissata però in modo esogeno, in cui verosimilmente $t > C1_p - P1$. In questo Paese, nonostante la liberalizzazione sia avvenuta nel 1994, non è mai entrato alcun concorrente nel mercato *end-to-end*.

Il Regno Unito ha ugualmente interamente liberalizzato il mercato nel 2006, creando contestualmente un sistema di *zonal pricing*. Anche qui è verosimile ritenere che vi sia stata deviazione rispetto ai valori efficienti nella fissazione dei prezzi d'accesso, ossia che $P2 < Ecpr$. Infatti, pur essendoci forte concorrenza nei mercati *upstream*, non si osserva alcun concorrente nel recapito: alle imprese conviene infatti immettere nelle rete dell'*incumbent* i propri invii anziché recapitarli in proprio.

Quello sopradiscusso è un altro degli elementi di differenziazione della regolazione postale da quella delle TLC: mentre in quest'ultimo settore ogni riduzione delle tariffe d'interconnessione, fino ad arrivare al modello Bill & Keep in cui queste sono pari a zero (Cambini e Valletti, 2003) produce verosimilmente un aumento di efficienza, perché i flussi di traffico tendono a compensarsi (a parità

¹⁹ Ovvero i prezzi sono differenziati geograficamente in modo tale che $P1 - C1 = 0$.

di mansione), nel settore postale prezzi d'accesso inferiori al Ecpr generano invece indesiderate inefficienze perchè i flussi sono monodirezionali.

Empiricamente, vi sono due modi per calcolare gli sconti prelaborazione, il metodo *retail-minus* e quello LRIC dei costi incrementali di lungo periodo. Il primo metodo usa un procedura *top-down* basata sui costi delle attività evitate così come risultano dal sistema di contabilità analitica ABC dell'*incumbent* gestore della rete, il secondo un approccio bottom-up di calcolo dei costi incrementali che un operatore efficiente supporterebbe fornendo i servizi di recapito. In un mercato perfettamente concorrenziale e in sviluppo l'approccio LRIC potrebbe essere teoricamente quello migliore, perché sconta le possibili inefficienze dei servizi in fase di recapito. Tuttavia, nel caso dell'industria postale il metodo *retail-minus* è quello preferibile, per due ragioni. Da un lato è più trasparente (se è ugualmente correttamente strutturata la contabilità analitica dell'operatore) e semplice da calcolare, dall'altro è neutrale rispetto all'*e-substitution* che genera in modo progressivo crescenti insaturazioni di rete, ovvero il problema principale che caratterizzerà l'industria postale nel prossimo decennio.

6- - Le strategie regolatorie nel mercato liberalizzato

Come in tutte le industrie a rete, il passaggio dal regime monopolistico a quello liberalizzato impone una revisione del sistema di regolazione, che ora deve avere fra i suoi obiettivi espliciti la promozione della concorrenza, senza dimenticare la tutela del consumatore, nel caso dei servizi postali meno scontata in quanto chi riceve le lettere non è quasi mai, se non indirettamente e solo se il settore mittente è poco concorrenziale²⁰, colui che paga il servizio. Inoltre, dato il consistente e costante calo dei volumi che generano inevitabilmente crescenti insaturazioni di rete, la nuova regolazione del settore postale dovrà tenere conto del mantenimento dell'equilibrio economico degli operatori, a continuo rischio di deterioramento. Il recente conferimento dei compiti di regolazione postale all'Agcom, avvenuto alla fine del 2011 e divenuto operativo qualche mese dopo, rappresenta una buona opportunità per rivedere, con un buon livello di tecnicità, tale processo.

²⁰ In Italia alcune *utilities*, evidentemente non sottoposte a sufficiente pressione concorrenziale, fanno pagare all'utente il costo della spedizione postale della bolletta.

6.1 I due modelli alternativi di competizione

Escludendo la competizione per il mercato, tipica dei monopoli naturali affidati ad un unico operatore attraverso una gara, inapplicabile al caso dei servizi postali – nonostante le esortazioni più di bandiera che operative di Agcm (2010) e Oecd (1996) – in quanto non sono un monopolio naturale, ma piuttosto un monopolio innaturale non sostenibile (Estrin e De Meza, 1995), rimane la competizione nel mercato, che qui prende due diverse forme, *end-to-end* e *upstream competition*.

Nonostante che esse siano teoricamente fra loro complementari nei modelli di regolazione, come ad es. Armstrong (2006), sussistono ampie evidenze empiriche che le due forme di competizione tendano a creare condizioni di mercato diverse, e dunque modelli di *business* parzialmente incompatibili. Ne segue che promuovere un tipo di concorrenza in un mercato dove vige l'altro tipo tende a distruggerne il modello di *business* adottato, ed è quindi necessario un attento bilanciamento di questo peculiare *trade-off* regolatorio.

La concorrenza *end-to-end* stimola l'innovazione – di cui un esempio sono in Italia i servizi a data e ora certe – di prodotto e/o processo, spinge a limitare e a ridurre le inefficienze dovute all'insaturazione di rete, ma è intrinsecamente basata sulla pratica di *cream-skimming*, fisiologica in questo mercato, e quindi incide sull'Onere del Servizio Universale (d'ora in poi OSU).

Nel 2012 in UK è stato adottato un nuovo schema regolatorio (Ofcom, 2012) che lascia maggiore libertà a Royal Mail nella fissazione dei prezzi *retail* e *wholesale*, mantenendo tramite *margin squeeze test* un controllo sulla concorrenzialità del mercato.

La concorrenza nelle fasi a monte del recapito, ossia l'*upstream competition*, stimola invece la concorrenza in tutte le fasi della catena del valore eccetto il recapito e anche fra fasi – ciascun operatore di una fase desidera acquistare predominanza sulle altre, trasformandole in *commodity* – risolve alla radice il problema del costo OSU. Infatti se il prezzo d'accesso al recapito è inferiore al Ecpr viene immediatamente meno la convenienza a svolgere l'attività finale di recapito dove l'OSU è in prevalenza concentrato, come mostra con chiarezza la decisione BNA discussa al paragrafo 5.3.

Nei Paesi dove c'è concorrenza nel recapito – Italia, Germania, Olanda, Svezia l'*upstream competition* è dunque limitata a sconti quantità e prelaborazione regolate nelle modalità (le tipologie di accesso) ma non nei prezzi lasciati alla contrattazione fra le parti. Viceversa, nei Paesi dove l'accesso è regolato – tipicamente i Paesi anglosassoni USA, UK²¹ e NZ – non esistono reti di recapito alternative a quella del-

²¹ Nel 2012 il 44% dei volumi riguardava operatori diversi dall'*incumbent* Royal Mail.

l'*incumbent*, protetta o dal monopolio (USA) o da un prezzo d'accesso $< Ecpr$.

Nessun Paese è mai passato da una forma di concorrenza all'altra, ciò che conferma la non applicabilità al settore postale della teoria del *ladder of investment*. In riferimento all'Italia, appare quindi opportuno vigilare sulle condizioni di accesso alla rete, come ha del resto fatto recentemente Agcom con la delibera 92/13 CONS.

6.2 la regolazione dei prezzi

La peculiarità italiana nella regolazione dei prezzi dei servizi postali è l'adozione di un modello di *price cap* in presenza di un *deficit* dei prodotti regolati, infatti agganciato ad un *subsidy cap* relativo al trasferimento statale compensativo²², che lascia una certa libertà di aumento dei prezzi, superiore al tasso d'inflazione.

D'altro canto, anche negli altri paesi nonostante la liberalizzazione del mercato, si è assistito ad una crescita dei prezzi dei servizi (Tav. 7).

TAV. 7

CRESCITA DEI PREZZI NELLE TARIFFE LETTERE IN ALCUNI PAESI EUROPEI
(anni 2001-2011 - valori %)

Royal Mail	122%
Post Danmark	100%
La Poste	30%
PostNL	28%
Itella (FI)	24%
Posten (Sve)	20%
Deutsche Post	-2%

Fonte: WIK (2012).

I dati della Tavola 7 mostrano un aumento generalizzato dei prezzi, con la sola eccezione di Deutsche Post, che però nel 2001 presentava il prezzo della lettera più alto in Europa. Questi dati si riferiscono al servizio lettere più veloce (in Italia lettera prioritaria), e non tengono conto dei prezzi della posta *bulk* originata da imprese e organizzazioni, rimasti invece sostanzialmente stabili e, più di recente, in calo sia in Italia che negli altri Paesi.

La ragione è che il servizio prioritario è *retail*, per sua natura poco soggetto alla concorrenza (i competitors sono interessati soprattutto alla posta *bulk* e a quella registrata), gravato da vincoli di universalità, in alcuni paesi come l'Italia

²² Per una discussione sul modello di *price cap* adottato in Italia vedi VISCO COMANDINI V. (2007)

trattato da una rete separata con elevata componente di lavoro manuale, quindi costoso e con profittabilità strettamente legata a volumi elevati che consentono il godimento di economie di scala. Gli operatori francese, tedesco e olandese riescono a presentare margini positivi anche per questo segmento, ma non quello italiano dove i volumi sono scarsi.

Le prospettive future sono quindi di aumenti ulteriori del prezzo per i segmenti postali *retail*, fundamentalmente legati alla crescita del fattore lavoro, e di lievi riduzioni nel segmento *bulk*, dove la concorrenza sia *end-to-end* sia *upstream* li spinge verso il basso.

L'effetto della liberalizzazione del mercato postale sembra dunque tradursi in un iniziale beneficio per le imprese che hanno anche maggiori possibilità di scelta, ma anche in un maggior onere per i consumatori singoli, che vedranno ulteriormente crescere i prezzi dei servizi universali *retail*. In ogni caso, lo scenario dei volumi in calo non potrà non condizionare seriamente la regolazione dei prezzi.

La Tavola 8 è desunta da uno studio di Cohen e McBride (2012) che ha utilizzato con dati aggiornati il modello del medesimo autore (Cohen *et al.*, 2006) in cui si dimostrava come la percentuale dei costi fissi nell'industria postale sia inversamente proporzionale alla dimensione del mercato (espressa in volumi *pro capite*). Vengono mostrate le stime degli effetti del calo dei volumi sulla variazione dei prezzi necessaria per mantenere invariato il saldo di gestione dell'operatore postale ad obblighi di servizio universale inalterati rispetto a quelli mantenuti costanti.

TAV. 8

STIMA DELLA VARIAZIONE DEI PREZZI NECESSARIA A MANTENERE COSTANTE
IL SALDO DI BILANCIO DEGLI OPERATORI POSTALI AL VARIARE
DEI VOLUMI DI POSTA

scenario	% costi fissi su totale costi al 2009	Δ % volumi 2009-2020 (2009=100)	Δ % prezzi a invarianza saldo di bilancio	Δ % annua prezzi	% costi fissi su totale costi al 2020
1	60	125	-16,0	-1,6	54,6
2	60	100	0,9	0,1	60,3
3	60	75	32,2	2,6	67,5
4	60	50	104,5	6,7	75,6
5	60	25	350,0	15,0	85,0
6	40	75	19,4	1,6	48,4
7	40	50	62,9	4,5	59,8

Fonte: COHEN R. - MCBRIDE D.R. (2012).

Nella tavola sono presentati i valori caratteristici del mercato postale italiano, dove a causa dei bassi volumi l'incidenza dei costi fissi dell'operatore postale nazionale è stimata dal modello di Cohen intorno al 60% (scenari 1-5), posti a confronto con quelli tipici dei grandi operatori europei, dove l'incidenza del costo fisso è valutata intorno al 40%.

Si vede subito che se nel 2020 in Italia i volumi si ridurranno del 25% rispetto agli attuali (stima conservativa), sarà comunque necessario un aumento medio annuo dei prezzi pari al 2,6%, ma che se il calo dovesse essere del 50%, l'aumento dei prezzi richiesto salirebbe a 6,7%. Migliore situazione per i paesi in cui il costo fisso è oggi al 40%: a fronte di riduzione dei volumi del 25% e del 50%, l'aumento dei prezzi sarebbe più contenuto, rispettivamente 1,6% e 4,5%. Da osservare, infine, nell'ultima colonna, che l'effetto finale del calo dei volumi implica necessariamente un aumento della componente fissa dei costi.

A tal fine appare quanto mai opportuno aggiornare alle nuove esigenze del mercato liberalizzato il regime di regolazione dei prezzi dei prodotti universali, fissato dal provvedimento Cipe del 2003. In quel provvedimento il *price cap* era stato agganciato ad un più stringente *subsidy cap*: in sostanza veniva lasciato più spazio agli incrementi tariffari in cambio di una progressiva e più che proporzionale riduzione del trasferimento statale. L'aggiornamento del modello regolatorio²³ dovrà tenere conto dell'effetto concorrenza: alcuni prodotti o segmenti di essi (posta massiva recapitata nelle aree urbane dove operano i concorrenti) non necessitano di alcun controllo essendo i prezzi determinati dal mercato. Inoltre, sarà necessaria una scelta fra modelli di *price cap multi basket* (che consentono una maggiore flessibilità nella fissazione dei prezzi) o *single-basket* (che vincolano prodotto per prodotto la crescita del prezzo). Ciascuna delle due soluzioni ha effetti distributivi diversi: la prima consente di accrescere la differenza di prezzo fra prodotti *bulk* e prodotti *retail*, la seconda la limita.

6.3 I prezzi politici

In molti paesi i prezzi di alcuni servizi postali "meritori" come i prodotti per l'editoria e il *non profit* vengono offerti a prezzi sottocosto, con sussidi governativi compensativi trasferiti al fornitore nazionale. In Francia un sussidio statale è garantito a La Poste a completa copertura dei costi. Non così in Italia dove i prezzi, già inferiori ai costi, vengono fissati da decreti ministeriali con approvazione del

²³ Ofcom ha recentemente abolito il controllo prezzi tramite *pricecap*, sostituendolo con forme di verifica meno stringenti.

Parlamento, e mai coperti da adeguate compensazioni concesse a Poste, rimborsate a consuntivo dopo anni, con “sconti” ulteriori richiesti dal Tesoro. Più di recente, tagli alla spesa pubblica hanno quasi azzerato i sussidi ai prodotti postali, ma la logica della fissazione di prezzi imposti per via politica è rimasta pressoché inalterata.

Il mantenimento di questo regime è incompatibile con la liberalizzazione del mercato²⁴, sia perché preclude libere scelte imprenditoriali da parte dell'operatore postale nazionale, tentato di effettuare scambi politici anziché di mercato, sia perché impedisce che altri operatori, che hanno il vincolo della redditività e che non godono della medesima possibilità di accedere ai sussidi, entrino in questo segmento di mercato.

Le nuove norme in materia di tariffe postali agevolate addirittura estendono tale distorsione. In data 20 luglio 2012, è stata pubblicata la legge 16 luglio 2012, n. 103, che ha convertito il D.l. 18 maggio 2012, n. 63, recante «disposizioni urgenti in materia di riordino dei contributi alle imprese editrici, nonché di vendita della stampa quotidiana e periodica e di pubblicità istituzionale». In sede di conversione, detta legge ha inserito nel testo del decreto-legge originario l'art. 5-*bis*, rubricato “semplificazioni in materia di editoria per le associazioni ed organizzazioni senza fini di lucro e le associazioni d'arma e combattentistiche” che estende agli enti *non-profit* le tariffe postali agevolate fissate dal Decreto del Ministro delle Comunicazioni del 13 novembre 2002, già applicabili.

Di conseguenza, si è aperta la possibilità per una parte significativa della clientela utilizzatrice del *direct mail*, come le associazioni e organizzazioni senza fini di lucro iscritte nel Registro degli operatori di comunicazione (ROC), di avvalersi delle tariffe agevolate (ovvero a prezzi ancora inferiori) finora riservate alle imprese editrici di libri, quotidiani e periodici.

Queste norme di fatto costituiscono un forte disincentivo per tutti gli operatori, sia quello dominante sia i suoi concorrenti, a sviluppare il mercato del *direct mail*, che tende a svuotarsi (i suoi prezzi, ancorché inferiori ai costi, sono infatti superiori a quelli fissati in favore dei prodotti editoriali sovvenzionati). In altre parole, per favorire alcune categorie di utilizzatori di posta si precludono le opportunità di sviluppo dell'unico segmento mail che crescerà in futuro.

²⁴ Nel 1993 l'operatore postale svedese accettò la liberalizzazione del mercato a condizione di non essere più gravata da tale imposizione.

7. - Ridefinizione, costo netto e finanziamento del servizio universale

7.1 *Definizione del perimetro*

Le due dimensioni in cui è declinabile il servizio postale universale sono la orizzontale e la verticale. La prima attiene le modalità del suo svolgimento, in primis la frequenza del recapito, e rimane di stretta competenza dell'Unione Europea che ne fissa criteri e parametri nelle direttive, la seconda è lasciata invece alla decisione degli stati membri.

Quanto alla prima, le richieste alla Commissione Europea affinché autorizzi, sia attraverso deroghe *ad hoc* sia attraverso una revisione della direttiva, una maggiore libertà nell'obbligo di consegna della posta su un numero inferiore agli attuali 5 la settimana, si sta facendo pressante. Le ipotesi non sono di allentare in modo generalizzato tale obbligo, ma modulare il recapito al fine di ridurre la crescente insaturazione dei portalettere dovuta al calo dei volumi.

La Commissione Europea ha approvato deroghe parziali ai 5 giorni di consegna settimanale alla Grecia, in considerazione della conformazione geografica in cui una parte importante della popolazione vive in isole sparse, e alla Finlandia in un'area a sud del Paese a bassa densità abitativa.

In Italia il D.lgs. n. 58 del 2011 ha innanzitutto ridotto dai precedenti 6 a 5 gli obblighi di recapito settimanale, ma ha anche dato facoltà a Poste Italiane di introdurre un sistema di recapito postale ordinario a giorni alterni nelle aree con densità di popolazione inferiore ai 200 abitanti per Kmq e per un massimo di popolazione di circa 7,5 milioni.

La *ratio* di tale norma è evidente, se si considera che in gran parte del Mezzogiorno la media annua delle lettere per abitante non supera le 40, ossia una lettera ricevuta mediamente ogni 9 giorni, a fronte di un servizio oggi dimensionato per servire quotidianamente ciascun punto di recapito sul territorio. Si tratta di uno spreco sociale di risorse incompatibile, nel medio periodo, con l'obiettivo di pareggio nel *business* postale di Poste Italiane e con il mantenimento di un servizio universale efficiente.

Tuttavia, nelle decisioni di ristrutturazione della rete (recapito e/o uffici postali) vanno tenuti presenti gli atteggiamenti negativi da parte della classe politica e dei media locali: in tutto il mondo si registrano proteste, come negli Stati Uniti in cui il Congresso ha stigmatizzato negativamente l'esclusione del recapito il sabato, o le recenti forti critiche sollevate in Italia sui progetti di chiudere un centinaio di piccoli uffici postali in grave *deficit* (quelli che svolgono meno di 30 operazioni a settimana).

Quanto alla dimensione verticale, l'esclusione dal servizio universale dei prodotti mail tipicamente *bulk* – ossia quelli che non vengono accettati allo sportello degli uffici postali ma in grandi quantità direttamente ai centri di smistamento – è un'opzione già esercitata in diversi Paesi fra cui Germania, Belgio, Olanda, Regno Unito (in Francia solo per i pacchi), ma per ragioni più di tutela della concorrenza e di maggiore libertà nei giorni di consegna che per reali risparmi di costo conseguibili nel breve periodo.

Tale scelta è oggi ulteriormente rafforzata dai principi contenuti nella Comunicazione della Commissione del 20 dicembre 2011 sull'applicazione delle norme dell'Unione europea in materia di aiuti di Stato alla compensazione concessa per la prestazione di servizi di interesse economico generale, secondo cui non è desiderabile attribuire obblighi specifici di servizio pubblico ad un'attività che è già fornita in modo soddisfacente dal mercato. Il *bulk mail*, che nella gran parte dei paesi è il prodotto che consente al fornitore nazionale di realizzare profitti, presenta proprio tale caratteristica, e in linea teorica dovrebbe, come suggerisce l'Agcm (2012), essere escluso dal servizio universale.

Il mantenimento di posta massiva nel servizio universale presenta la seria implicazione anticoncorrenziale determinata dall'assoggettamento asimmetrico dell'Iva (in Europa esclusa dal servizio universale a seguito della citata sentenza *TNT Post UK* della Corte di Giustizia Europea). Infatti, i clienti che utilizzano posta massiva (non gravata da Iva) e che, come banche, assicurazioni, pubblica amministrazione e organizzazioni *non profit*, non la raccolgono, hanno un incentivo a continuare ad utilizzarla anche se il prodotto equivalente dei concorrenti, il servizio a data e ora certe fuori del servizio universale assoggettato ad Iva, ha un prezzo inferiore.

Tuttavia, la semplice esclusione di posta massiva dal servizio universale presenta alcuni svantaggi di politica: costringe immediatamente i soggetti che non sono attivi Iva ad aumentare sensibilmente i propri costi per le spese postali.

Una possibile soluzione viene dalla stessa sentenza europea sopracitata che così recita: «l'esenzione Iva va interpretata in senso strettamente restrittivo, applicabile solo agli operatori che forniscono servizi postali rivolti a soddisfare bisogni essenziali della popolazione»²⁵, ossia in quanto fornitori di «servizi sostanzialmente diversi da quelli offerti da altri operatori»²⁶. Questa interpretazione della Corte di Giustizia è coerente con il principio della neutralità fiscale, chiaramente violato

²⁵ Corte di Giustizia Europea, caso C-357/07, paragrafi 31 e 36.

²⁶ *Idem*, paragrafi 37-39.

dall'attuale normativa Iva italiana, e potrebbe concretizzarsi in una distinzione interna al medesimo prodotto di posta massiva, che verrebbe quindi sdoppiato in due prodotti diversi.

Tale soluzione che si applicherebbe così sia fra prodotti sia fra aree geografiche e/o Cap, e aggiusterebbe le disparità fiscali oggi vigenti, minimizzando peraltro l'aggravio dell'Iva per i clienti che non la possono dedurre nei loro bilanci, riportando condizioni di parità concorrenziale per tutti gli operatori di mercato. Infatti nei Cap dove c'è concorrenza, tutti gli operatori fornirebbero un servizio con Iva, in quelli antieconomici in cui opera il solo fornitore universale, invece, il servizio sarebbe, come oggi, esente Iva.

7.2 *Il costo del servizio universale*

L'allegato 1 della direttiva postale 6/2008 prevede che il costo OSU sia calcolato utilizzando il metodo del costo netto evitato e cioè come la differenza tra i costi netti del fornitore del servizio universale che è soggetto ad un obbligo di servizio universale (scenario base) e lo stesso costo netto in assenza di tali obblighi (scenario controfattuale). Esso definisce tre fattori ulteriori da aggiungere a questa differenza:

- i) i benefici intangibili e commerciali derivanti;
- ii) un rendimento sugli investimenti ragionevole;
- iii) incentivi ai miglioramenti di efficienza.

In questo caso, il costo del servizio universale può essere rappresentato dalla seguente formula:

$$OSU^N = \Pi_{OSU} - \Pi_{NONOSU} - B + R - \alpha$$

dove:

OSU^N = costo netto dell'*OSU*;

Π_{OSU} = profitto realizzato dal fornitore del servizio universale con gli attuali obblighi di servizio universale;

Π_{NONOSU} = profitto realizzato dal fornitore del servizio universale dove non ci sono obblighi di servizio universale;

B = benefici tangibili e intangibili generati dagli attuali obblighi di servizio universale;

R = ritorno sugli investimenti effettuati;

α = parametro di efficienza.

Dal momento che lo scenario base e lo scenario controfattuale, a meno di un progetto di chiusura del servizio, non differiscono in modo significativo in termini di qualità del servizio e ampiezza della rete, la differenza $\Pi_{OSU} - \Pi_{NONOSU}$ può essere semplificata come differenza tra i profitti in caso di obblighi di servizio universale e in caso di assenza di obblighi di servizio universale.

Nei servizi postali, generalmente, le perdite imputabili agli obblighi di servizio universale dipendono da tre fattori:

- (1) Obbligo di recapitare su tutto il territorio nazionale, anche nelle aree non profittevoli;
- (2) Obbligo di mantenere aperti gli uffici postali, anche quelli non profittevoli;
- (3) Obbligo di fornire dei servizi a prezzi regolati che non coprono i costi totali.

Lo scenario controfattuale quindi misura i profitti nel caso in cui il fornitore del servizio universale abbia libertà di ridurre i giorni di recapito alla settimana nelle aree non profittevoli e di ottimizzare la rete di uffici postali.

B è un parametro importante perché deve tener conto dei benefici economici ricevuti dal fornitore del servizio universale nel suo espletamento. Si tratta quindi di quantificare i profitti, se esistenti, realizzati in servizi forniti grazie all'ubiquità della rete (in recapito e in accettazione con gli uffici postali). Nel caso italiano, appare difficile non considerare i profitti realizzati da Poste Italiane nei settori finanziari del Bancoposta e in quelli assicurativi di Posta Vita.

R è il ragionevole tasso di rendimento del capitale investito, che in prima misura potrebbe essere il WACC (attualmente non calcolato per Poste Italiane).

Infine α che dovrebbe rappresentare un parametro di promozione dell'efficienza, scontando nell'ammontare del costo netto riconosciuto al fornitore universale una determinata percentuale di inefficienza. La misura dell'efficienza degli operatori postali nazionali è piuttosto complessa e non sempre il *benchmarking* fra paesi è corretto se non si tengono conto delle differenze strutturali esistenti fra essi (Visco Comandini *et al.*, 2010). È comunque importante distinguere correttamente l'inefficienza produttiva, legata a decisioni o azioni sub ottimali, dalla semplice insaturazione di rete, data da un insufficiente output ma da fattori produttivi scarsamente modificabili.

7.3 Progetti innovativi finalizzati alla riduzione del costo OSU

Diversi operatori esteri – in particolare quelli finlandese e svizzero – stanno studiando soluzioni capaci di mantenere la qualità del servizio riducendo però i costi. L'obiettivo è ridurre la frequenza del recapito fisico e migliorare la qualità

del servizio sostituendolo parzialmente con il recapito elettronico su PC e dispositivi mobili.

Il nuovo servizio (*reverse hybrid mail*) consiste nella scansione delle buste destinate al recapito, le cui immagini in formato elettronico vengono inviate al destinatario il giorno (o i giorni) precedente al recapito fisico, in modo da avvertirlo del prossimo arrivo. Al destinatario è offerta quindi l'opzione fra il recapito fisico o quello esclusivamente elettronico, con la scelta supplementare dell'apertura della busta e la scannerizzazione completa del suo contenuto (a pagamento).

Il *reverse hybrid mail* è stato sviluppato e posto sul mercato qualche anno fa da Earth Class Mail negli Stati Uniti, e adesso anche in due Paesi europei, la Svizzera e la Finlandia, che hanno investito risorse ed esperienze su questo servizio, e sono passati dalla fase della sperimentazione all'offerta rivolta a tutti.

Il progetto ha un modello sostenibile di *business*, perché quello che l'operatore postale perde come minori recapiti, lo può recuperare giocando il ruolo di portale che raccoglie, distribuisce e intermedia i servizi.

Date le prospettive di calo del mercato, l'innovazione costituisce la scelta obbligata per gli operatori postali, che hanno l'opportunità di governare l'*e-substitution* anziché subirla passivamente, e di utilizzarla per rendere più efficace e moderno lo stesso servizio universale, che ha il suo punto nodale nel recapito. È quindi auspicabile che anche in Italia vengano lanciati, sul modello della Finlandia, progetti sperimentali di sostituzione del recapito fisico con quello elettronico, approfittando proprio dei bassissimi volumi oggi recapitati (ancora più bassi proprio nelle aree rurali) e della attuale grave diseducazione della popolazione di Internet, e della domanda potenziale di molti servizi su Internet che Poste intermedierebbe. I progetti potrebbero essere complementari agli investimenti infrastrutturali di base necessari per collegare comunità locali oggi in digital divide in agenda del Governo e a programmi mirati di educazione all'utilizzo di Internet nelle medesime comunità, da cofinanziarsi attraverso programmi comunitari e nazionali.

7.4 *Il finanziamento del servizio universale*

Legato alla misura del costo netto del servizio universale c'è il problema del fondo di compensazione, istituito fin dal D.lgs. del 1999, ma di fatto mai implementato, perché il suo finanziamento, peraltro parziale, era interamente lasciato al sussidio pubblico. Gli stretti vincoli al bilancio pubblico impongono ora una revisione che implicherà certamente rendere operativo il fondo, che dovrà essere finanziato da tutte le imprese postali che operano nelle aree profittevoli e che of-

frono anche parzialmente servizi sostitutivi di quello universale. Sorge quindi il problema della definizione della base imponibile del fondo, nonché quello dell'aliquota e del criterio di contribuzione, che certamente non è risolvibile semplicemente utilizzando il parametro della quota dei ricavi.

Di fatto l'ammontare del fondo dipenderà dal calcolo del costo netto, mentre non appare praticabile l'ipotesi, prevista dalla raccomandazione dell'AGCM (2012) di mettere a gara il servizio universale. La perplessità non è di ordine concettuale, bensì di concreta praticabilità dello strumento.

Infatti, come peraltro spiega la teoria economica con il teorema della privatizzazione di Sappington e Stiglitz (1987) i vantaggi dell'affidamento di una concessione di servizio pubblico tramite asta competitiva dipendono in modo cruciale da tre requisiti:

- i)* diffusione fra i partecipanti all'asta delle informazioni rilevanti;
- ii)* un numero di partecipanti sufficiente ad evitare rischi di collusione;
- iii)* neutralità al rischio dei partecipanti.

Nel caso dell'industria postale, il primo criterio è facilmente assolto, ma non il secondo e il terzo: essendo oggi Poste l'unico soggetto tecnicamente in grado di offrire il servizio quanti soggetti potrebbero partecipare ad un'asta al ribasso per l'aggiudicazione remunerata della sua gestione in Italia? Probabilmente due, nel migliore dei casi tre: non vi sarebbero quindi apprezzabili ribassi, a meno di non stimolare l'assunzione di rischi crescenti. Non a caso, l'opzione dell'assegnazione della concessione tramite gara fino ad oggi non è mai stata implementata in alcun Paese al mondo. I vantaggi tipici dell'asta competitiva che la rendono in teoria preferibile all'assegnazione diretta della concessione, nel caso in questione non sembrano essere verosimilmente conseguibili.

8. - Il corriere espresso

8.1 *Nascita e sviluppo dell'industria*

L'industria del corriere espresso è sorta nei primi anni Ottanta, quando imprese private iniziarono ad offrire servizi nazionali ed internazionali di consegna di plichi e documenti ad elevato valore aggiunto, come la garanzia di recapito in 24 o anche 12 ore, la tracciatura del percorso o l'accettazione al domicilio del mittente. Dopo un iniziale periodo di conflitto con i ministeri e gli operatori postali nazionali che consideravano illegittima l'attività dei corrieri perché in violazione del monopolio, furono emanate norme che autorizzavano queste imprese

ad operare, a condizione che le prestazioni fossero sostanzialmente diverse e i prezzi decisamente superiori a quelli del servizio ordinario, allora riservati all'operatore statale.

Nei primi anni il mercato del corriere espresso era rappresentato dalla domanda di nicchia di professionisti e imprese che, nel caso di invii importanti, potevano scambiarsi documenti con assoluta garanzia di celerità e sicurezza, specie nelle tratte internazionali. Questo mercato si presentò subito come decisamente profittevole, specie per le imprese che disponevano di una propria infrastruttura di trasporto (flotta aerea e mezzi su gomma) capace di assolvere gli elevati *standard* di servizio richiesti.

Fin dai primi anni Novanta le migliori imprese di corriere espresso divennero oggetto d'attenzione da parte degli operatori postali nazionali europei che, preparandosi all'apertura del mercato, vedevano nell'integrazione con i servizi di corriere espresso un'opportunità di crescita della competitività e dell'area geografica servita. In questo decennio tutti i grandi operatori europei acquisirono i maggiori operatori di corriere: Deutsche Post acquistò DHL, le poste olandesi il corriere TNT, Royal Mail l'operatore GLS, i francesi la rete internazionale di consegne DPD, Poste Italiane acquisì il corriere nazionale SDA e, per alcuni anni, anche una partecipazione di minoranza del corriere Bartolini (oggi BRT).

Negli Stati Uniti, invece, fin dall'inizio i mercati postali e di corriere espresso furono tenuti rigorosamente separati per proprietà, servizi svolti e assetti regolatori: le poste sottoposte al monopolio statale, i corrieri lasciati liberi di operare con regolamentazioni minime. Negli Stati Uniti le due maggiori imprese, oggi divenute multinazionali, sono Federal Express e UPS.

8.2 *I modelli di business*

Dallo scambio di documenti importanti e urgenti i corrieri si sono progressivamente allargati a quello delle consegne B2B fra imprese, alla distribuzione dei pacchi e, negli ultimi anni, al *fulfillment* della catena distributiva degli esercizi commerciali che, in tutta Europa, stanno sempre più spesso restringendo lo spazio di magazzino per aumentare quello di vendita. I documenti che nel 2000 costituivano il 42% delle spedizioni, oggi rappresentano solo il 20% (Luiss Business School-Aicai, 2010), anche perché risentono, come le lettere ordinarie, dell'effetto dell'*e-substitution*.

I corrieri hanno saputo rispondere all'esigenza delle imprese di ridurre il magazzino, adottando la cosiddetta tecnica del *just-in-time* (Astrid, 2012), in cui si

accorcia la catena che va dalla produzione alla vendita finale. In Italia, in particolare, l'aumento di valore degli immobili siti nei centri urbani ha fatto lievitare il costo degli spazi commerciali, con conseguente riduzione delle superfici utilizzate in cui il magazzino è addirittura assente. Di qui la necessità di un rifornimento di merci rapido e praticamente continuo: oggi un tipico esercizio commerciale sito nel centro urbano e operante nel settore dell'elettronica/elettrodomestici o nell'abbigliamento utilizza le imprese di trasporto e logistica per i grossi carichi, quelle di corriere espresso per coprire rapidamente i picchi delle richieste.

Il modello produttivo dei corrieri in Italia è basato sull'*outsourcing*²⁷ selettivo a imprese individuali e cooperative specializzate della fase del trasporto e smistamento, con contratti basati sul n. spedizioni effettuate e/o dei km percorsi. L'efficienza è perseguita ottimizzando quotidianamente la rete di trasporto e distribuzione di pacchi e documenti.

Nella Tavola 9 viene mostrata l'attuale ripartizione, secondo il consueto schema origine-destinazione, del traffico gestito dai corrieri espressi in Italia.

TAV. 9

RIPARTIZIONE DEL FLUSSO DI DOCUMENTI E MERCI GESTITE DAI CORRIERI
ESPRESSI PER ORIGINE-DESTINAZIONE
(anno 2012 - valori %)

<i>origine</i>	<i>destinazione</i>			nel complesso
	Imprese	Privati	Istituzioni	
Imprese	81	8	1	90
Privati	3	4	1	8
Istituzioni	1	1	0	2
nel complesso	85	13	2	100

Fonte: stime degli autori su dati AICAI.

In tutto il mondo il traffico del corriere espresso è originato per il 90% da imprese, per l'8% da privati e in via residuale dalle Istituzioni. In Italia questo traffico si svolge in modo preponderante nel segmento B2B per soddisfare le esigenze del just-in-time, mentre è ancora poco sviluppato (solo 8%) quello B2C, in cui rientra l'*e-commerce*.

Nell'*e-commerce* l'Italia presenta un gap rispetto agli altri paesi (Tav. 10), che potrebbe ridursi nei prossimi anni considerate le forti prospettive di crescita in tutto il mondo.

²⁷ Il 47% dei costi dei corrieri deriva dall'*outsourcing* del trasporto (38%) e dello smistamento-facchinaggio (9%) (fonte: LUISS BUSINESS SCHOOL - AICAI, 2010).

TAV. 10

VENDITE NELL'E-COMMERCE B2C, N. ACQUIRENTI ON-LINE E VENDITE PER ACQUIRENTE NEI PRINCIPALI PAESI INDUSTRIALIZZATI (anno 2011)

Paese	totale vendite (mld. \$)	Acquirenti on-line (mln.)	vendite per acquirente (\$)
USA	143,4	302	2.104
Regno Unito	33,0	109	3.304
Germania	38,2	38	997
Francia	24,5	30	1.211
Spagna	13,4	16	1.165
Italia	10,5	13	1.230
Olanda	7,3	7	960
Svezia	4,1	5	1.298
Norvegia	2,2	5	2.259
Finlandia	2,3	5	2.064
Danimarca	2,5	4	1.430

Fonte: EMARKETER (2013).

In Italia nel 2011 gli acquisti *on-line* sono stati solo 13 milioni, contro i 16 della Spagna, i 30 della Francia e i 38 della Germania, fino ai 109 del Regno Unito. Il valore medio di acquisto in Italia non è basso, indizio che gli acquirenti su Internet, ancorché ancora poco numerosi, seguono però comportamenti del tutto simili a quelli dei Paesi più sviluppati. Stime recenti (*fonte*: Osservatorio B2C della School of Management del Politecnico di Milano) indicano una crescita del 19% del valore delle vendite online fra il 2011 e il 2012, in completa controtendenza con PIL e consumi in deciso calo.

I corrieri hanno interesse a gestire la parte logistica dell'*e-commerce* incluse le consegne, specie se si tratta di acquisti *cross-border* fra paesi diversi, ma stanno sempre più spesso trovando un'accesa concorrenza negli operatori postali tradizionali. Questi ultimi da sempre gestiscono, ovunque in perdita, il servizio pacchi universali in cui vengono servite le fasce più povere del segmento dei pacchi a basso prezzo (ben lontano da prezzi e prestazioni offerte dai corrieri), e tentano, con maggiore o minore successo, di evitare il rischio di marginalizzazione attraverso forme di integrazione con le aziende di corriere acquisite. Poste Italiane, ad esempio, gestisce mediamente circa 30-35 mln di pacchi universali, ma circa un terzo delle perdite subite nel settore postale deriva proprio da questo prodotto.

8.3 *Il mercato del corriere espresso*

Per le ragioni sopradescritte, il mercato del corriere espresso è in costante crescita in tutto il mondo, compresa l'Italia dove, nonostante la crisi economica, è aumentato fra il 2010 e il 2011 del 2,4% in valore (4% nel caso dei corrieri internazionali, fonte: Aicai).

In Italia il mercato del corriere espresso ha raggiunto dimensioni ragguardevoli (poco meno di 4 mld di € in valore nel 2011), e ormai da qualche anno ha superato il valore del mercato postale tradizionale, fermo a circa 3,5 mln di €.

Nell'industria postale si sta assistendo, in tutto il mondo, ad un'inversione di ruoli, in cui il settore acquisito (corriere espresso) ha preso il sopravvento in termini di ricavi su quello dell'acquirente (i servizi postali tradizionali). Evidenze sono il cambio di denominazione dell'operatore nazionale tedesco che ha recentemente modificato il proprio nome in Deutsche Post - DHL, o la separazione proprietaria avvenuta fra le poste olandesi (Post NL²⁸) e il corriere TNT Express, che la statunitense UPS ha recentemente tentato di acquisire, senza successo data la decisione negativa della Commissione Europea sulla fusione.

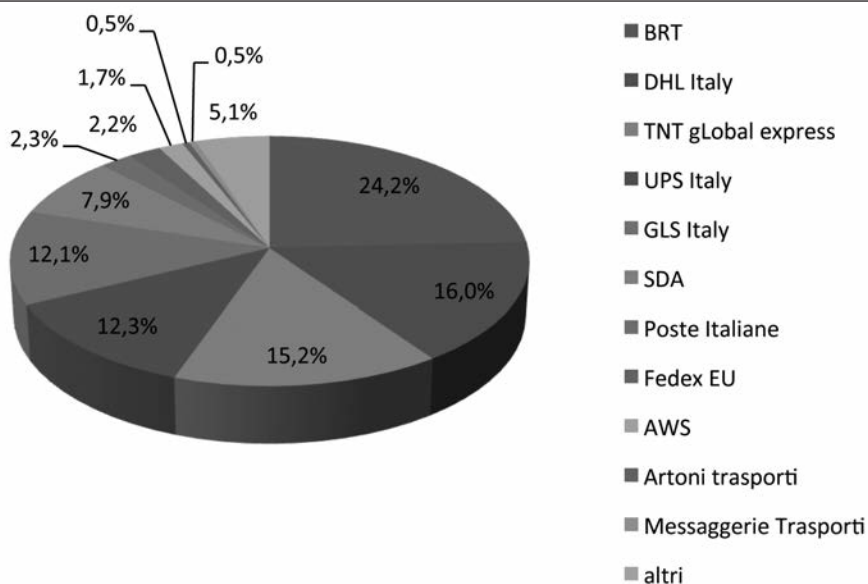
Il mercato è composto dai due segmenti geografici nazionale e internazionale: fra il 2008 e il 2011, il primo risulta in calo del 2,3% l'anno, il secondo in crescita pari a +5,9% (fonte: Aicai).

Nel Grafico 9 sono mostrate le quote del mercato italiano nel 2011 dei principali operatori di corriere espresso secondo stime che tengono conto di dati Databank, ed elaborazioni degli autori.

²⁸ In Italia e Germania opererà ancora per qualche anno con il nome di TNT-Post.

GRAF. 9

INDUSTRIA DEL CORRIERE ESPRESSO QUOTE DI MERCATO
(anno 2011 - valori %)



Fonte: elaborazioni degli autori su dati DATABANK.

Come si può osservare, in Italia il mercato appare piuttosto concentrato – i primi 10 operatori rappresentano quasi il 95% del fatturato specifico totale – ma nessuna impresa appare in posizione dominante.

9. - La concorrenza multimediale

9.1 *Il direct mail fra i grandi media di investimento pubblicitario*

Nella maggior parte dei grandi paesi industrializzati il *direct mail* (d'ora in poi, DM) è da sempre considerato uno dei media importanti su cui vengono effettuati investimenti pubblicitari. Nel 2006, ad esempio, negli Stati Uniti l'investimento nel DM era pari all'85% di quello nella TV, una percentuale più o meno costante registrata in quel Paese negli ultimi vent'anni; nel Regno Unito, sempre nel medesimo anno, incideva per il 12,2% sul totale degli investimenti.

Il mondo dell'*advertising* distingue tradizionalmente due tipi di investimento pubblicitario: 1) quello effettuato *above-the-line* (ATL), ovvero su veri e propri media che dispongono di una specifica piattaforma, rivolti ad una generalità di

consumatori secondo la classica relazione broadcast *one-to-many*, in cui l'obiettivo è diffondere un messaggio sul marchio e sul valore dell'impresa pubblicizzata, creando le condizioni per un acquisto non necessariamente immediato, e 2) quelli *below-the-line* (BTL) ovvero su mezzi, o piuttosto attività (come le sponsorizzazioni, il *direct marketing*, le promozioni) dove c'è un rapporto diretto *one-to-one* che spinge il destinatario ad effettuare un acquisto immediato.

L'avvento della tecnologia digitale ha reso meno cogente questa tradizionale ripartizione degli investimenti pubblicitari, in particolare in relazione a Internet che presenta caratteristiche sia ATL (è un media con propria piattaforma che mette in relazione utenti, editori di contenuti/fornitori di servizi e investitori pubblicitari, secondo il tipico modello dei mercati multiversanti²⁹) che BTL (dal momento che la relazione, grazie alla capacità di segmentazione dell'utenza, è di tipo *one-to-one*).

Ma anche il DM si adatta ormai male a questa distinzione, per almeno due motivi. Il primo è che, date le sue potenziali caratteristiche di attrattività grafica (se bene confezionato), non è necessariamente finalizzato a generare un impulso immediato di acquisto³⁰, ma può essere utilizzato più volte (ossia può consistere anche in un catalogo o in un mezzo per raggiungerlo). Il secondo è che opera attraverso una vera e propria piattaforma dedicata (il circuito di distribuzione postale) e mette in comunicazione consumatori e inserzionisti pubblicitari che senza di essa non potrebbero scambiarsi informazioni: è infatti anch'esso classificabile come *two sided transaction market* (Filistrucchi *et al.*, 2012; Visco Comandini *et al.*, 2009).

Nelle statistiche dei vari paesi sugli investimenti pubblicitari per media a volte il DM compare insieme ai mezzi ATL, altre volte no. In Francia, Adexpress include il DM fra i media, nel Regno Unito WARC fa altrettanto, in Italia Nielsen Media Research pubblica regolarmente il fatturato pubblicitario del DM unitamente a quelli degli altri media; Agcom (2013), invece, non effettua rilevazioni dirette sul DM, ma lo include, senza identificarlo come media autonomo, nella più ampia classe del *direct marketing* (incluse quindi le telefonate dirette ai consumatori potenziali) all'interno dei BTL.

Il DM è un media costoso, con un elevato costo per contatto (se confrontato a quelli tradizionali ATL), costi variabili determinati dalla componente di recapito,

²⁹ Per una discussione e rassegna delle applicazioni della teoria dei *multisided market* si veda RYSMAN M. (2009).

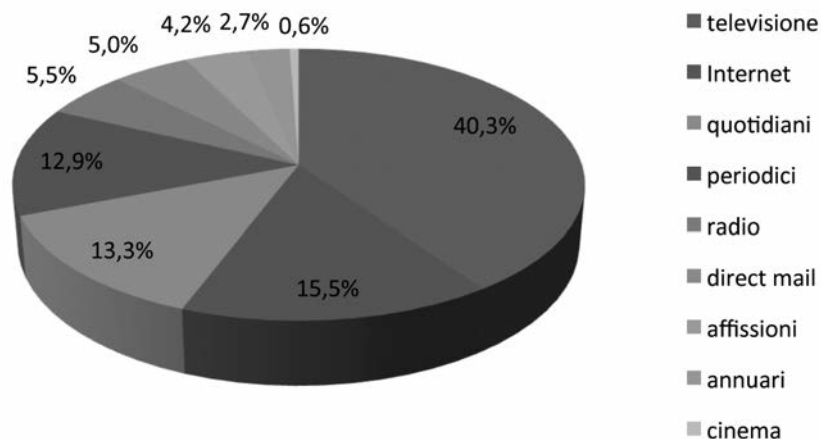
³⁰ La finalità di impulso immediato all'acquisto del DM è invece prevalente in quello di bassa qualità, il cosiddetto *junk mail*, simile allo spam dell'*email*, che riesce a interessare solo il consumatore marginale a basso livello d'istruzione, mentre è considerato con fastidio dal consumatore evoluto.

in tutto il mondo in parte sussidiata (i prezzi postali del DM sono inferiori a quelli delle lettere, nonostante abbiano costi simili) e costi della parte *upstream*, comprendenti stampa, creatività e gestione liste clienti, riconosciuti nel prezzo finale pagato dal cliente solo nei paesi dove questo media è fortemente utilizzato. In Italia, dove il suo sviluppo è invece limitato, le imprese che operano nella parte *upstream* fanno fatica a veder remunerato il valore aggiunto da loro apportato al prodotto.

Nel graf. 10 sono presentate le quote di fatturato al 2011 dei media ATL più il DM, considerando però i valori rilevati da Agcom (e Nielsen per il DM). Accomunare valori di media diversi ha senso se esiste un certo grado di sostituibilità fra di essi, in particolare il DM, come del resto osservato empiricamente per il Regno Unito da Soteri *et al.* (2009) che ne ha studiato la relazione con Internet, da Ambrosini *et al.* (2009) e da Visco Comandini *et al.* (2009). Tale sostituibilità è invece negata da Agcom, secondo cui ATL e BTL non competono fra loro.

GRAF. 10

INVESTIMENTI PUBBLICITARI IN ITALIA PER MEDIA
(anno 2011 - valori %)



Fonte: elaborazioni degli autori su dati NIELSEN MEDIA RESEARCH e AGCOM.

Mancano ad oggi studi sistematici che abbiano analiticamente esaminato il complesso fascio di relazioni che legano fra loro gli investimenti pubblicitari dei diversi media. In effetti la relazione esistente fra essi non è quasi mai lineare di semplice complementarità o sostituzione, perché l'effetto netto dipende dal contesto, dalla dimensione dell'investimento da effettuare, dal punto di osservazione

nonché dal settore merceologico (Ambrosini *et al.*, 2009). La sostituibilità è però certamente più alta al margine fra media che pure hanno investimenti medi per cliente molto diversi. Se infatti è difficile sostenere che un forte investimento, in TV in prima serata sia facilmente spostabile su altri media senza modificare la struttura della campagna, è invece verosimile che un investimento in TV in ore marginali a basso ascolto sia facilmente sostituibile con quello di pari limitato import su quotidiani, periodici, DM o ancor più su Internet, ormai divenuto stabilmente anche in Italia il secondo media per investimenti.

9.2 I profili di privacy

Come si è già osservato, il DM è l'unico prodotto postale con prospettive positive di crescita nel prossimo futuro. Tale vantaggio potenziale è però messo in discussione da una legislazione sulla *privacy* che, almeno in Italia, lo ha finora fortemente penalizzato e che rischia di comprometterne lo sviluppo. Solo di recente, infatti, con la Legge 12 luglio 2011, n. 106 (che ha convertito in legge il cosiddetto "Decreto Sviluppo" - D.l. 13 maggio 2011, n. 70) è stato modificato il regime di *privacy* relativo alla pubblicità per corrispondenza (*direct mail*).

La nuova normativa ha esteso anche agli indirizzi postali il regime dell'opt-out (il consumatore riceve inizialmente il messaggio e, in caso non lo desidera, può chiederne il blocco dell'invio) di recente introdotto nell'ordinamento in materia di trattamento dei numeri telefonici degli abbonati per l'esercizio del *marketing* telefonico. Quindi anche per il DM si applica il registro delle opposizioni, per cui gli operatori di *marketing* diretto cartaceo potranno utilizzare anche gli indirizzi degli abbonati contenuti nell'elenco telefonico per finalità promozionali senza bisogno di chiedere il consenso, con la sola condizione che questi ultimi non abbiano richiesto l'iscrizione del proprio numero telefonico e del proprio indirizzo presso il registro delle opposizioni di recente istituito dalla Legge n. 166 del 2009 e gestito dalla Fondazione Ugo Bordoni.

Chiaramente l'obbligo di consultare il registro delle opposizioni riguarda solo coloro che abbiano intenzione di usare indirizzi tratti dagli elenchi telefonici. Chi usa recapiti di fonte diversa non è, ad oggi, tenuto a tale adempimento.

Va al riguardo segnalato che la norma più rilevante – quella relativa all'estensione del registro delle opposizioni al mailing cartaceo – richiede l'emanazione di un regolamento che definisca come dovrà essere gestita l'attività di consultazione del registro, che il MISE inspiegabilmente ad un anno di emanazione della norma non ha ancora pubblicato. Ritardi nella regolamentazione sulla *privacy*, generando

incertezza nei grandi investitori pubblicitari, privano l'unico segmento della posta con prospettive positive di crescita della possibilità di svilupparsi.

10. - Conclusioni

Il calo dei volumi dovuto all'inesorabile e lenta sostituzione dei messaggi fisici su carta con quelli elettronici impone una riconsiderazione sull'ampiezza del servizio universale, ancora oggi la modalità di fornitura preferita da cittadini e imprese.

Questo è già in via di riduzione in tutti i paesi: quelli in cui, come l'Italia, il recapito era organizzato su sei giorni la settimana, hanno abbassato lo *standard* a 5 giorni, con numerosi progetti tendenti a ridurli ulteriormente, anche per rispondere alla sfida competitiva dei concorrenti, in genere organizzati per fornire un servizio di recapito due o tre giorni la settimana.

Diverso è invece il problema della riduzione del numero degli uffici postali sul territorio, la cui localizzazione è stata storicamente basata sul principio di massimizzazione dell'accessibilità dell'utenza, non del profitto. In quasi tutti i paesi, il 30% degli uffici postali genera l'80-90% dei ricavi, così che il principio del MEIP (*market economy investor principle*) imporrebbe la chiusura del rimanente 70%.

Molti operatori postali, compresa Poste Italiane che su tale profilo rappresenta un caso di successo con PosteVita e Postemobile, hanno cercato di ovviare alla non profittabilità di una vasta porzione della rete degli uffici postali accrescendo il numero dei prodotti e servizi offerti. Questa opzione, specie se coniugata con un *deficit* rilevante conseguito nei servizi postali tradizionali, presenta però criticità di *antitrust* – i rischi di sussidi incrociati illegittimi, o di prezzi predatori nel settore in perdita, aumentano sensibilmente – e di strategie di lungo periodo che i governi sono chiamati a risolvere.

BIBLIOGRAFIA

- AMBROSINI X. - BREVILLE S. - CORNEE J. - KLARGAARD O., «Mail Order Business Demand: A Conceptual Framework», in CREW M. - KLEINDORFER P. (eds.), *Progress in the Competitive Agenda in the Postal and Delivery Sector*, Edward Elgar publ., Cheltenham & Northampton, 2009, pages 196-208.
- ASTRID, *I servizi di distribuzione urbana delle consegne espresse in Italia fra tutela ambientale, riduzione del traffico e ricerca dell'efficienza*, VISCO COMANDINI V. (a cura di), mimeo, marzo, 2012.
- AUTORITÀ GARANTE PER LA CONCORRENZA E IL MERCATO, *Segnalazione in merito a Proposte di riforma concorrenziale ai fini della Legge annuale per il mercato e la concorrenza anno 2013. Il settore postale*, Roma, ottobre, 2012, p. 46.
- AUTORITÀ PER LE GARANZIE NELLE COMUNICAZIONI, *Indagine conoscitiva sul settore della raccolta pubblicitaria*, Allegato A alla Delibera n. 551/12/CONS, Roma, 2012.
- .-, *Osservatorio sulla pubblicità. Risultati della prima edizione*, Roma, gennaio, 2013.
- BOSTON CONSULTING GROUP, *The Postman Always Brings Twice. Receiver-driven Transformation of the Mail Business Model*, March, 2011.
- CAMBINI C. - VALLETTI T., «Network Competition with Price Discrimination: “Bill-and-Keep” is not so Bad After All», *Economics Letters*, vol. 81, no. 2, 2003, pages 205-213.
- COHEN R. - ROBINSON M. - SCARFIGLIERI G. - SHEEHY R. - VISCO COMANDINI V. - WALLER J. - XENAKIS S., «The Role of Scale Economies in the Cost Behavior of Posts», in DIEKE A. - NIEDERPRÜM A. (eds.), *Regulating Postal Markets - Harmonised Versus Country-specific Approaches*, Papers presented at the 8th Königswinter Seminar, 16-18 February 2004, *WIK Proceedings*, no. 11, Bad Honnef, 2006.
- COHEN R. - MCBRIDE D.R., *Implications of Declining Mail Volumes on the Financial Sustainability of the US Postal Service and Other Posts*, George Mason University, Washington, 2012.
- COMMISSIONE EUROPEA, *Report from the Commission of 23 March 2005 to the Council and the European Parliament on the Application of the Postal Directive (Directive 97/67/EC as amended by Directive 2002/39/EC)*, COM(2005)102, 2005.
- .-, *Comunicazione della Commissione del 20 dicembre 2011 sull'applicazione delle norme dell'Unione europea in materia di aiuti di Stato alla compensazione concessa per la prestazione di servizi di interesse economico generale*, 2012.
- COPENHAGEN ECONOMICS, *Main Developments in the Postal Sector (2008-2010)*, Final Report, Study commissioned by the European Commission, November 29, 2010.
- ERGP, *Report on “Access” to the Postal Network and Elements of Postal Infrastructure*, November 30, 2012.
- ESTRIN S. - DE MEZA D., «Unnatural Monopoly», *Journal of Public Economics*, no. 57, 1995, page 471.
- FAULHABER G., «Cross-Subsidization Pricing in Public Enterprise», *American Economic Review*, vol. 65, no.5, 1975, pages 966-977.

- FILISTRUCCHI L. - GERARDIN D. - VAN DAMME E., «Identifying Two-Sides Markets», Dipartimento di Scienze Economiche, Università di Firenze, *Working Paper*, no. 2, 2012.
- FIORAMANTI M., «La domanda di servizi postali», in VISCO COMANDINI V., *Concorrenza e regolazione delle reti postali*, Roma, Carocci, 2006.
- GLOCUS, *Il futuro del mercato postale in Italia. Analisi e proposte*, VISCO COMANDINI V. (a cura di), Roma, dicembre, 2012.
- LIEBOWITZ S. - MARGOLIS S., «Path Dependence, Lock-in, and History», *Journal of Law, Economics and Organization*, vol. 11(1), April, 1995, pages 205-226.
- LUISS BUSINESS SCHOOL - AICAI, *Il modello di business dei corrieri aerei internazionali. Dalla catena del valore alle indicazioni di policy*, Roma, gennaio, 2010.
- MÁRAI S., *Il sangue di San Gennaro*, Milano, Adelphi, 2010.
- OECD, *The Essential Facilities Concept, Policy Roundtables*, OCDE/GD(96)113, Paris, 1996.
- OFCOM, *Securing the Universal Postal Service Decision the New Regulatory Framework*, March 27, 2012.
- PRICEWATERHOUSECOOPERS, *The Impact on Universal Service of the Full Market Accomplishment of the Postal Internal Market in 2009*, Final Report, Study commissioned by the European Commission, Internal Market and Services Directorate General (DG MARKET/2005/03/E), May, 2006.
- RYSMAN M., «The Economics of Two-Sided Markets», *The Journal of Economic Perspectives*, vol. 23, no. 3, Summer, 2009, pages 125-143.
- SAPPINGTON D. - STIGLITZ J., «Privatization, Information and Incentives», *Journal of Policy Analysis and Management*, no. 4, 1987, page 567.
- SCARFIGLIERI G. - VISCO COMANDINI V., «Postal Profit Arise where People Are», in CREW M. - KLEINDORFER P. (eds.), *Future Directions in Postal Reform*, Kluwer Academic Publishers, Boston - Dordrecht - London, 2001, pages 255-266.
- SOTERI S. - FÈVE F. - FLORENS J.P. - RODRIGUEZ F., «Internet Advertising and Direct Mail: Trends and Analysis for the UK», in CREW M. - KLEINDORFER P. (eds.), *Progress in the Competitive Agenda in the Postal and Delivery Sector*, Edward Elgar publ., Cheltenham & Northampton, 2009, pages 209-222.
- VISCO COMANDINI V. (a cura di), *Concorrenza e regolazione delle reti postali*, Roma, Carocci, 2006.
- ., Servizi postali, *Rivista di Politica Economica*, luglio-agosto, fasc.7-8, 2007, pp. 257-314.
- VISCO COMANDINI V. - CONSIGLIO A. - GORI S. - PICCININ E. - PIERLEONI M.R., «The Altmark Ruling and Approaches to Measuring Efficiency of Postal Operators», in CREW M. - KLEINDORFER P. (eds.), *Heightening Competition in the Postal and Delivery Sector*, Edward Elgar publ., Cheltenham & Northampton, 2010, pages 70-84.
- VISCO COMANDINI V. - LINTELL M. - GORI S. - PIERLEONI M.R. - TISDAHL B., «Postal Price Elasticities and Intermedia Competition - A Multisided Market Approach», in CREW M. - KLEINDORFER P. (eds.), *Progress in the Competitive Agenda in the Postal and Delivery Sector*, Edward Elgar publ., Cheltenham & Northampton, 2009, pages 163-183.

RASSEGNA BIBLIOGRAFICA

Recensione del volume:
Inequality and Instability[◇]

JAMES K. GALBRAITH

a cura di

Paolo Liberati*

“Roma Tre” University

James K. Galbraith has surely written an unconventional book on inequality, under various profiles. The first uncommon point of view is that inequality within countries must be interpreted as a macroeconomic phenomenon. Among the many numbers produced in this book, his estimates in chapter 3 give powerful evidence that inequalities within countries are not a national “choice”, rather they are shaped by global dominant forces. On this basis, Galbraith provides his unconventional reading of the evolution of inequality from the Seventies until the first decade of the new millennium. After a period of declining inequality through the 1970s, a sudden reversal of inequality started at the beginning of the Eighties. As Galbraith argues «this is the onset of the debt crisis, marked by high real interest rates and commodity price deflation, a collapse of economic growth, and mass unemployment throughout the world» (page 72). Quite interestingly, the global component of inequality continued to push inequality upward for other twenty years (the *superbubble* period in world financial markets), with regional crests in Latin America, Central and Eastern Europe and Asia, to stop only after the technology crash of 2000. However, following a period of falling interest rates, the revival in commodity markets peaked with the oil bubble of 2008 and with a boom in home financing that collapsed at the end of the same year. Thus, inequality cycles, according to Galbraith, are strictly dependent on global cycles, «vitiating national policy autonomy in the determination of inequality levels» (page 72). Neutralising the global component, he argues, inequality would have declined.

[◇] Oxford University Press, 2012.

* <paolo.liberati@uniroma3.it>, Department of Economics.

A second unconventional point of view is that, by using new datasets and new inequality measures, Galbraith provides a re-interpretation of the Kuznets's original hypothesis. With data drawn from the Utip-Unido dataset on pay inequality, he actually finds that most of the results lie on the downward-sloping segment of the Kuznets curve; yet, he also provides evidence that for China and the richest countries the relationship may reverse, yielding rising inequality as income increases. To this respect, the message is that changes in inequality can be guided by concentration of income in specific sectors of the economic activity, that may evolve and change over time, and still be consistent with the Kuznets's hypothesis.

But why does this happen? Galbraith, and this is the third unconventional view, identifies inequality as driven by what happens at the top of the income distribution. But what happens at the top of the income distribution, he argues, is mainly affected by the functioning of the capital. In other words, inequality is driven by "forces that are largely financial in character", beyond being global. Particularly interesting, in the Galbraith's book, is that these forces operate not only with respect to stable democratic and industrialised countries, but also with respect to countries that are ideologically committed like China and Cuba (chapters 10 and 12, respectively) or with respect to less developed countries like Argentina and Brazil (chapter 11). Galbraith claims that Argentina and Brazil made similar transitions to open market economies, and that "both experienced the instability and stress associated with the neoliberal economic climate". The result being that after large increases, inequality fell when countries retreated from neoliberal globalisation policies. In other words, the Galbraith message is that "no country is an island" in a deeply integrated world, and if inequality depends from macroeconomics, every country can be affected by it. This is the reason why, in his view, macroeconomic forces becomes more important than the microeconomic foundations of inequality changes.

This position is path-breaking with respect to more traditional views of what forces can drive inequality changes. On the one hand, it is here rejected the mainstream argument that growing volumes of trade would increase the demand for high skilled workers in developed countries and for low skilled workers in developing countries. As a consequence, inequality in developed countries would increase, as low-skilled wages would be depressed, and inequality in developing countries should reduce – at least to some extent. This latter effect, however, would strongly depend on the technological content of the tradable sector. Moving high technology to developing countries would also put a premium on the more skilled

labour; at the same time it reduces the demand for unskilled labour, with the result that inequality may increase also in developing countries. As Galbraith says, the skill bias argument, for the mainstream, «... had two comforting political features: it subsumed rising inequality into inevitable and desirable technical progress, and it placed the onus on individual workers to improve their lot through education and training. It thus deflected those who might be tempted to call for a public action» (page 125). According to this view, technological changes are good and, consequently, inequality must be accepted as a price that societies have to pay for flexible markets. However, the “skill bias” argument is rejected in the Galbraith’s book. This would make justice of a well established argument that has populated a significant part of the economic literature since the Nineties. This position implies that education and training may not be the best (and the only) route to fight inequality. He instead argues that especially in the last twenty years, income has concentrated in specific sectors. In 1990s, the technology revolution has conveyed enormous masses of financial investments; in the late 2000s, at least in the US, the real estate sector has replaced the technology venture. In both cases, Galbraith argues, it is this concentration of incomes and rents that must be avoided and the corresponding harmful credit cycle.

On the other hand, the view that trade openness and international competition from low-wage countries significantly contributed to inequality is also rejected, especially because wage inequality has not dramatically increased around the world. Rather, it is non-wage inequality driven by stock benefits and financial investments that has expanded and widened the gap between poor and rich countries. While fascinating, these arguments are not entirely new. The thesis that the liberalisation of financial markets would have contributed to increase speculation and to create a new class of *rentiers*, by this way favouring increased income inequality, flows through the literature at least from the same time as globalisation issues were addressed. However, the Galbraith’s merit has been to feed this idea with new power and to bring it in from the cold of the recent times, challenging the mainstream position that inequality is a natural product of well-behaving markets, and the orthodox leading line of argument that there is no link between inequality and financial instability. Indeed, it comes as a surprise that the relationship between credit cycles, inequality and financial instability has been systematically under-stressed in the economic literature.

This paves the way to the fourth unconventional view of the Galbraith’s book. He argues that greater inequality increases unemployment. While he finds that more egalitarian regions have less unemployment than unequal regions, he also

argues that there is no evidence that being more unequal or having more wage flexibility attracts jobs to a region. This sounds like a hammer for those who strongly believe in the virtues of the flexible labour markets in curing unemployment. Particularly important and worth of further investigation is the comparison between the US and Europe as a whole. The conventional wisdom in this field is that the US have lower unemployment as they have greater wage flexibility; while European countries have higher unemployment as they pursue more egalitarian policies. Galbraith shakes this conventional wisdom by observing that looking at Europe as a transnational labour market, pay inequality is much higher in Europe than in the US. As a consequence the required policy would not be to increase the degree of flexibility of the labour market, but to reduce the degree of inter-country pay inequality. As sharply argued in the conclusions to chapter 9, «the exercise casts grave doubt over the idea that unemployment in European countries can be explained by a failure of their wages to fall, or that unemployment can be remedied in general by policies aimed at cutting relative wages» (page 213). To Galbraith, the widespread thesis, among economists, that Europe chose equality over efficiency, while America made the opposite choice is wrong once a proper level of aggregation of the labour market is chosen.

The fifth unconventional view directly leads to the title of the book. Inequality and instability. Drawing from the experience of a rich country, as the United States, Galbraith finds that the credit cycle has been a fundamental driver of the economic performance since 1980. Indeed, inequality measured across counties in the United States is found to be highly correlated to the proportional movement of the Nasdaq stock index. To this regard, Galbraith argues that a very large share of the rise in the topmost incomes was concentrated in just a handful of counties closely associated with the technology boom (Silicon Valley, Seattle, and their bankers in Manhattan). The consequences for the functioning of markets and democracies are more serious than one may think. In such cases, not only «financial booms and busts drive the performance of employment, and thus prosperity is associated with rising income inequality», but one can find «practically everywhere the same signature of a rising share of total income passing through the financial sector» (page 290). With the consequence that when movements are global, national economic policies may hardly deal with inequality in an effective way. Galbraith reminds us that small countries are very vulnerable to global financial forces, and there are limits to what the national policies of these countries can do to address the unwanted outcomes of this process. On the other hand, the countries that control the bulk of financial forces have an enormous power,

not least because big countries have more opportunities for domestic control. As Galbraith concludes, the economics of inequality is in large measure an economics of instability, where inequality is a barometer of the instabilities that global credit relationships create (page 18).

Overall, the book is a wide and detailed examination of different facets of the economics of inequality, as well as a valuable effort to provide an unconventional view of the relationships between economic inequality and global macroeconomic forces. For sure, it is good food for all readers that are undeceived with current economic orthodoxies. Even though the book does not fully describe what remedies should be activated to face such a challenging issue, its implications are significant. First, the faith on self-adjusting markets is not always well grounded, especially in the case of financial markets, that may cause instability and growing inequality. Second, the ability of single countries to counteract the unwanted effects of these forces is low. This severely undermines the autonomy of national policies, and weakens the effectiveness of public and social policies that are not market-based. Third, if global forces operate, global remedies must be activated, yet agreements on global policies are not easily achieved. Fourth, the flexibility of labour markets is not necessarily the proper remedy to inequality; in the same vein, national efforts in improving education and training may vanish under the power of global forces. Thus, the message is as clear as troubling. Inequality is a global phenomenon that require global and complex remedies. Galbraith's book provides a strong incentive to embark in this new perspective of the study of inequality issues.

Recensione del volume:
**The Basel Committee on Banking
Supervision. A History of the Early
Years 1974-1997**[◇]

CHARLES GOODHART

a cura di
Giorgio Szegö*
"Sapienza" University of Rome

1. - General Comments

This is a very detailed presentation of the history and of the decision process in the period 1974-97 of what is currently called «The Basel Committee on Banking Supervision». It reads like a murder story in which, unfortunately, the victim is the economy of most western countries, in particular of the EU. This Committee was created in March 1975 by the Central Bank Governors of the Group of 10. The original mandate was to reach a harmonization on Banks' Capital. This mandate was soon expanded into that of designing an efficient bank regulatory and supervisory system. All G-10 countries plus Switzerland and Luxembourg joined the new Committee which took over the role of an informal *Groupe de Contact*, which was established by middle-rank managers of European Central Banks. The *Groupe* had its first meeting in June 1972 and ended its activity in February 1975. All members of the *Groupe* were initially appointed to the new Committee.

Prof. Goodhart was granted access to some material of the Basel Committee, but not to the minutes of the meetings that, on the contrary, were secretized. At page 137 of the book the author comments one of the most striking *modus operandi* of the committee «in the best tradition of bureaucracies, the issue was passed on to a subgroup to reconsider». Indeed, I have counted 29 among subgroups, subcommittees and task forces; it seems that everybody who was anybody

[◇] Cambridge University Press, 2011.

* <gsz.jbf@fastwebnet.it>, Emeritus Professor of Monetary and Financial Economics.

in the committee wanted to have his name attached to one! It is interesting to consider the list of topics assigned to these subgroups: most deal with the implementation of the rulings of the Committee, and on the analysis of sources of possible losses (country, exchange, offshore, interest rate, off-balance-sheet, etc.), mistakenly labeled as risk. Not much attention was paid to the more crucial questions, like the macroeconomic consequences of capital requirements, and a proper definition and measure of risk. These issues were not considered until the late 1990s, when, under the chairmanship of Tommaso Padoa-Schioppa, a working group of experts was established. This group, chaired by Patricia Jackson of the Bank of England, had the mandate of analyzing the macroeconomic consequences of the mandatory capital requirements. Indeed, only in 1999 the interesting conclusions reached by the group were published in a working paper entitled «Capital Requirements and bank behavior: the impact of the Basel Accord». It reached the conclusion that capital requirements during downturns «may have limited bank lending ... and contributed to economic weakness». These statements seem to have had zero impact of the decisions of the committee: following the best bureaucratic practices, they were swept under the rug! Unfortunately, the book covers events only up to 1997, when the so-called «Market Risk Amendment to the 1988 Accord» and the «Core Principles of Banking Supervision» were released. The limitation of the time span did not allow to take into account the strong critiques against the mandatory Capital Requirements imposed by the 1988 Accord that were made in the year 1998 by many public organizations. (Iosco, Isda, Ifi, Fed).

In my opinion the most interesting results in the book are contained in Chapter 7 on «The Market Risk Amendment». Indeed the 1988 accord focused on credit risk and ignored interest rate risk. The study on this unaccounted risk begun in 1985, before the 1988 Accord. Chapter 7 presents two important results contained in two excellent reports produced by the «Models Task Force» chaired by Christine Cumming of the Federal Reserve Bank of New York. The first deals with the high dispersion among VaR figures resulting from the application of individual bank models to identical test portfolios (page 251). Appendix J of chapter 7 presents the «summary of results to the test portfolio exercise». This summary is rather difficult to interpret. The dispersion of VaR figures is very likely caused by different techniques used in the construction of profit and loss distributions, thus vanishing all efforts in harmonizing risk measures. The refusal to use competent statistical advice is evident. The second Cumming's report deals with the risk models that were already used by some banks. While the Basel group was fiddling with obsolete and wrong methods to measure risk, the main international banks had already de-

veloped and applied new advanced techniques (see Federal Reserve System, 1998). It seems that the committee never heard about the concept of co-dependence, and had swallowed hook, line and sinker the intuitive, but wrong, idea of VaR (Value at Risk). Unfortunately, VaR is not a measure of risk. Again, it is not understandable why the committee decided to operate as a closed shop without any competent statistical support. While only in 1997 Artzner, Delbaen, Eber and Heath published in the journal *Risk* a paper on the properties that must be shared by all measures of risk, the instabilities that were experienced in VaR tests should have warned the committee that it was on the wrong path: like measuring a distance with a rubber band. A few years after the release of the Cumming reports, in 2001 Daniélsson *et al.* in their «Academic Response to Basel II» stated that «VaR can destabilize an economy and induce crashes when they would not otherwise occur». We had to wait until 2009 for a regulatory agency (the UK Financial Services Authority, FSA) to share this view. Indeed The Turner Review released by FSA in March of the same year states that «VaR measures of risk may not only fail adequately to warn of rising risk, but may convey the message that risk is low and falling at the precise time when systemic risk is high and rising». Unfortunately, this incorrect risk measure is still used by many of the less sophisticated credit institutions and blessed by regulators. For instance, the EU adopted it in the European Capital Adequacy Directive (Dir. 93/6 EEC); indeed, we are currently suffering of a regulation-induced crisis. The second Cumming report opened the way to the permission granted to regulated financial institutions to use internally generated models. The committee somewhat hid this total surrender by requiring that the possibility of using internal models would be conditional upon explicit approval of the competent supervisory authority. We are anxiously waiting for a new committee dealing with the harmonization of the approval criteria!

Throughout the book, the author mentions the number of meetings. For instance at page 263 he notes that the September 9-10 1996 meeting was the seventy-ninth (almost) quarterly meeting. Does the author wish to point out the inefficiency of the Committee?

In the Epilogue, written in 2007, after the start of the current crisis, the author takes the opportunity to express his personal views on the “Basel rules” and points out their failings summarizing them in the following 5 points.

1. The lack of any theoretical basis.
2. The focus on the individual institutions, rather than the system.
3. The failure to reach an accord on liquidity.
4. The lack of empirical analysis.

5. The unwillingness to discuss sanctions or crisis resolution.

I wish to add four more points:

6. Biased country representation.

7. No provision for assets concentration/diversification.

8. Total disregard for any dynamic aspect.

9. Lack of any cross-country harmonization of accounting, fiscal and bankruptcy rules.

The Committee decisions not only did not have any theoretical basis, but they totally disregarded the many theoretical results against the (pseudo) «risk-adjusted capital requirements».

Regarding the second point, on the focus on institutions instead that on the system, we must point out that very seldom monetary authorities are capable/willing to break these institutional walls. In my memory, it happened in the US in 2008 after the Bear-Stern default and the crisis of AIG, Merrill-Lynch, and Lehman Bros. The then US Secretary of the Treasury, Hank Paulson, was able to save the system by breaking institutional walls.

Concerning the failure to reach an accord on liquidity, after reading Chapter 9, my comment is “thank God”!

Regarding the lack of empirical analysis, the committee was clearly not interested in facts, which could have been the outcome of empirical analysis; it would have been in for some surprise as the results obtained by some academic trouble-makers proved that the introduction of costly capital requirements increases banks' risk propensity.

The members of the Committee were chosen among the founding members of the EU, no efforts were made to address the unbalance. Why Belgium and Luxemburg without Spain and Poland?

The main criticisms to the 1988 Agreement centered exactly on the lack of any diversification premium. This was considered an unfair burden to small firms: bank-granting loans to many different small firms is less risky than granting big loans to large firms. The regulatory capital allocation was however the same: smaller firms are penalized. Unfortunately, the Basel committee seems to have been unaware of Markowitz portfolio theory and of the importance of diversification. Markowitz results were published in 1956 and by 1970 became quite accepted and applied by the financial service industry. In spite of that, Basel rulings discouraged diversification. The financial world expected a fast easy correction of these mishaps. Unfortunately, when the obese Basel 2 proposal was published in 2004, (16 years after Basel 1) it contained no provisions on diversification. From the 2004, Michael Gordy of the Board of Governors of the Federal Reserve

System was a member of a tall force on Concentration that developed a very advanced measure, labeled “granularity”, but it was not given any space in the new rulings. It is evident that a prompt correction of the 1988 Agreement would have been tantamount to an admission of error from the Committee!

The committee totally disregarded the “transition cost”. Indeed nobody can deny that more capitalized institutions are more capable to sustain losses than less capitalized ones. Unfortunately with the exception of the Patricia Jackson report on «Capital Requirements and bank behavior: the impact of the Basel Accord» the committee choose to disregard the general impact of a recapitalization of banks and to impose it during the current severe crisis, thus inflicting additional damage to the economy. Is there any rationale in this behavior?

The lack of harmonization in bankruptcy rules makes the implementation of the Basel Capital Requirements equivalent to installing doors in a new building, before the walls!

Unfortunately the motto of the Committee seems to have been «to reach an accord at all costs, no matter its content», forgetting the general principle that no agreement is better than a bad agreement. This principle, of course, is against the interest of the most prominent committee members that would obtain positive publicity and personal benefits only in case of an international accord.

2. - The Structure of the Book

In this section I shall briefly summarize the content of the book. The Volume is structured in 17 chapters: chapter 1 is a general Introduction and chapter 17 an Epilogue.

The introductory chapter recalls the changes in the structure of the committee. It was founded in 1975 by the G-10 Central Banks Governors with the title of Basel Committee on Banking Regulations and Supervisory Practices, which became known as Basle Supervisors’ Committee. In 1998 it was agreed to adopt the local spelling of Basle and to call the committee «Basle Committee on Banking Supervision, BCBS».

The second Chapter presents the antecedents of the BCBS. In 1972 a group of middle-ranking officials of the banks’ supervisory authorities of the six original EEC countries (Belgium, France, Germany, Italy, Luxemburg, and the Netherlands) established an informal “*Groupe de Contact*” with the task of discussing international supervisory issues. The need for some agreement on international banking was ev-

ident: the end of the Bretton-Wood system in August 1971 introduced a strong exchange risk in all cross-currency operations. This risk caused the failure of New York's largest commercial bank: Franklin National Bank in May 1974 and of a private German bank, Bank Herstatt in June 1974. This latter collapse seems to have inflicted the maximum possible damage to the international financial system. This small bank had taken large speculative positions in the foreign exchange market, in particular in US\$/DM trades. The German authorities shut the bank after the German markets had closed, but before the US market had closed. Thus, US banks that had sold DM to Herstatt, and expected a payment in US dollar in USA did not receive it. The losses supported by US and UK banks (Morgan Guaranty, First National Bank of Seattle, and Hill Samuel) exceeded \$ 20 millions. Unfortunately, the German authorities instead of following the consolidated procedures of all financial transactions, insisted in enclosing the amount due to the US and UK banks in all other debts of the bank to be settled after liquidation. Was the timing carefully planned by the Germans? We shall never know.

The two banks failures of 1974, which were both caused by wrong bets on the US\$-DM exchange rates, took place in a world destabilized by the quadruplicating of oil prices which followed the Arab-Israeli war of October 1973. The need of some agreement on international financial transactions became evident. The positive experience of the *Groupe de Contact* was taken into account; all its members were included in the new Committee. It is difficult to evaluate how much the positive results of the *Groupe de Contact* influenced the G-10 decision to establish the Committee or if this decision was dictated by the financial turmoil of 1974. The new Committee was much larger than the *Groupe*. The *Groupe* members were all involved in banks' supervision, but only in the case of two countries (Italy and Netherlands) they were central bank officers: the majority of countries has delegated banks' supervision to *ad hoc* authorities. At the end of 1974, the G-10 (G-10 plus Luxemburg and Switzerland) governors decided to organize a formal Committee of Central Banks on Banking Regulation and Supervision to be hosted by the BIS in Basle. This project was strongly pursued by the Governor of the Bank of England, Lord Richardson. The first two Chairmen of the Committee (George Blunden and Peter Cooke) were both high level officials of the Bank of England. In order to accommodate in the Committee all members of the *Groupe*, and in particular those that were not officials of Central Banks, it was decided to invite in the Committee two members for each country. The first meeting of the BCBS took place in February 1975, almost simultaneously to the establishment of a new EEC Advisory Committee on Banking Regulations.

Chapter 3 of the book deals with the *Modus Operandi* of the Committee. The decisions taken by the Committee were strongly influenced by the personality of the Chairmen, in particular Mr. Peter Cooke (1977-1988), Mr. Gerald Corrigan (1991-1993) and Mr. Tommaso Padoa-Schioppa (1993-1997). Peter Cooke, who was head of The Banks' Supervision Department of the Bank of England is regarded to be the main proponent of the infamous «risk-adjusted capital requirements» that became known as “Cooke ratios”. Unfortunately, risk was not defined and the “risk weights” a typical exercise of “financial fantasy”. One example may strengthen this statement: in the 1988 accord, the sovereigns of all OECD countries and of Saudi Arabia were regarded as risk-free!

Chapter 4, entitled “The Concordat”, begins with the analysis of the home/host relationships *i.e.* cases in which some banks set up their operations in different countries in such a way as to hide from supervisors some of those activities. It analyses the case of the Banco Ambrosiano (1982) and of the bank BCCI (1991). While the first bank was incorporated in Italy, and thus supervised by the Bank of Italy, it set up a non-bank holding company in Luxemburg via some Latin American subsidiaries. Banco Ambrosiano was involved in loans to IOR, a Vatican bank. The establishment of a non-bank holding company prevented an efficient consolidated supervision. BCCI, which was essentially controlled by Abu Dhabi nationals, was incorporated in Luxemburg and in the Cayman Islands as a non-bank holding company. It established several subsidiaries abroad, and started a large volume of activities out of London. No banking supervision authority could ever inspect a consolidated balance sheet.

Indeed a clarification of the power of different bank supervision authorities for banking institutions operating across borders was needed. Another issue is the case of poorly supervised foreign banks subsidiaries. In the book it is mentioned the rather obscure case of the Atlanta Branch of Banca Nazionale del Lavoro (BNL) as well as the fall of Barings Bank. It seems that without a formal authorization from the Rome headquarter, the BNL Atlanta branch manager lent more than \$ 2 bn. to Iraq during the period 1985-1989, *i.e.* at the time of the Iraq-Iran war. It seems that this was a part of a covert US operation of support of Iraq. The loan was never returned and this was the beginning of the decline of BNL, at the time the largest Italian bank. Barings Bank was the oldest merchant bank in London until its collapse in 1995 after one of the bank's employees at the bank's Singapore office lost £827 million (\$1.3 billion) due to speculative investing, primarily in futures contracts. How can banks improve control on their foreign operations?

Chapter 5 deals with «External and foreign exchange issues». Foreign exchange problems had always an high priority in Basel. We recall that the formal establishment of the BCBS was triggered by the Herstatt crisis of 1974. The problems connected with cross country money transfers and foreign exchanges were obviously at the top of the committee agenda. In particular the BCBS tackled the project of developing an early warning system of international financial crisis. While it was evident that the Herstatt crash was caused by speculative excesses of the bank's traders in the forward market, the Herstatt fall could have been caused by any domestic problem, and not by foreign exchange. The committee passed this hot potato to the national supervisory authorities. This problem sired two subgroups: the first gave some thoughts to the liquidity problem; the second reached the conclusion that commercial banks should make their own decisions, but supervisors should check on their internal control mechanisms. The whole exercise seems the have forgotten that the Herstatt crisis and the consequent foreign currency losses were caused by exceptional events: firstly by the German authorities that did not allow to follow the consolidated praxis in international financial transactions to allow the closure of any operation, secondly by the insufficient internal control in Hertatt itself that allowed traders to perform some unauthorized (?) operation.

Chapter 6 presents the capital adequacy rule and the Basel Accord of 1988. According to Prof. Goodhart, BIS was the first institution to call the attention on the need for a uniform capital adequacy rules in a paper issued in 1974 on «Rules and practices to protect the banks' solvency and liquidity». This report includes a comparison of the capital requirements across different countries ranging from the Continental European Countries that adopted some sort of risk-weighted ratios and the UK and US having no formal requirements. In these early years, two questions seemed to dominate the discussions: the first on the possible role of subordinated debt in the regulatory capital and the second on how far foreign banks branches should be endowed with capital to be held in the host country. On both questions there were different opinions (from 0% subordinated debt of Germany, up to 50% of Luxemburg) no attempt was made to reconcile the different views. The chapter includes as Appendix A an excellent paper on «Subordinated debt in bank capital structures in the United States» submitted in 1976. Traditionally, in US bank capital was related to total deposits, the generally accepted norm being a 10% ratio. Only in the '50s the ratio of bank capital to total assets was taken into account: it fell sharply during the '70s. Regarding the question of the enclosure of subordinated debt in bank regulatory capital, in

1962 the US Comptroller of the Currency ruled that banks' regulatory capital could include up to 50% of subordinated notes and debentures.

At international level, the strong growth of international lending called the attention of G-10 Governors in 1980 on the question of capital adequacy. Unfortunately, the Governors did not take any action, and left the Basel Committee to "sink or swim". There was a disagreement among different nations on how to compute capital requirements. The US were strongly in favor of a simple leverage ratio, other countries had different views. During the years 1985-86, in spite of the strong action from the Basel Committee Chairman, Peter Cooke, no agreement was reached. The US Federal Reserve bank, under the Chairmanship of Paul Volcker, was very eager to reach some formal agreement on capital requirements due to the growing competition from the less capitalized Japanese banks. In September 1986 during a UK-US meeting in London the hosts were able to convince Volcker that the only possible accord would be on the basis of (so-far undefined) risk-weighted sum of all assets instead of the simpler, (and well defined) leverage or gearing ratio. The outcome of this agreement was the joint US-UK Agreement on bank capital adequacy of 1987 that paved the way to the 1988 Accord.

Chapter 7 deals with the "Market Risk Amendment" of 1996.

The LSE Financial market group led by Prof. Goodhart produced some very interesting research on this theme. It is not surprising that Chapter 7 is possibly the most documented chapter of the book. This chapter begins with the remark that «BCBS was fully aware that the 1988 Accord focused almost entirely on credit risk, and indicated that it would undertake further work on interest rate risk and investment risk on securities». The initial paper on interest rate risk was produced in 1985. This study was followed by another paper of 1986. While both papers pointed out that under the general heading of market risk include two different typologies: the risk that interest rate changes could affect the current accounting value of some assets (net interest income risk), and the position (or investment) risk *i.e.* a variation of the asset value in case of sale. This theme suggested the establishment of a galaxy of subgroups which dealt on the following problems: net income and property risk, investment/position risk in securities, off-balance sheet exposure and foreign exchange risk. Prof. Goodhart points out that many subgroups had interlocking goals. This almost hysterical urge to create new subgroups was not working in favor of efficiency and soundness of results and explains the reason for which it took 11 (eleven) years (1985-1996) to reach an agreement.

While the BCBS was debating on this matter, the European Community was also working on a directive «On the Capital adequacy of investment firms and

credit institutions” (CAD Directive) (Dir. 93/6/EEC). It seems that Gerard Corrigan, Chairman of the BCBS in the early ‘90s asked Sir Leon Brittan, then EC Commissioner, to slow down the works on the CAD Directive in order to allow BCBS more time to advance its work on Market Risk Amendment. Note that the CAD Directive was centered on the “level playing field” mantra. In my opinion, this is one of the most hypocrite ideas: while claiming to protect market fairness, it is a blatant defense of the existing players against new entries. The Market Risk Amendment on the other end is tainted by the wrong approach to measure risk. Instead of a precise “duration” or “economic” method, the committee adopted a “building blocks approach” using different buckets of risky assets. The rejected “economic” approach would have led to assessing the effect of shocks on banks’ portfolio as a whole. The adopted “building block” approach ignores the risk co-dependence among blocks. Another big mistake is the use of Value-at-Risk (VaR) as a measure of risk, which is not.

Chapter 8 analyses the «Core Principles of Banking Supervision» of 1997.

The work for this “soul-searching exercise” started in April 1988, but immediately stopped due to the very sensible objection of the German representative Schneider of Bafin that the BCBS did not have the mandate to define its own goals and limits; this task should belong to some higher monetary authority. In December 1988 the President of the Deutsches Bundesbank, Karl-Otto Pöhl, after consultation with the German Ministry of Finance withdraw the reserve.

In 1989 a new group, the «Financial Action Task Force on Money Laundering» (FATF) was established. This group, that had a much wider membership than the Committee, had the goal to operate against criminal money laundering and terrorist financing. The secretariat of FATF was hosted by Oecd in Paris, and its chairman was the very competent Swiss diplomat, Alexis Lautenberg. The bank supervisory authorities of all major countries, all already represented in the Basel Committee, nominated the members of this important group. In order to avoid work duplication and possible dissenting opinions, BCBS was induced to abandon any work on money laundering.

In November 1996 the Committee decided to draft a short list of Guidelines consisting of key principles for a sound supervisory system. These key basic principles should be applicable by any country, G-10 as well non-G-10. The IMF was chosen as an ideal point of contact with the non-G-10 banking authorities. The Guidelines in 1997 was given the stronger form of «Core Principles of Banking Supervision» and was open for comments. The final form of the Core principles was finalized in April 1997 in the form of 25 commandments worded in

the form: «Banking supervisors must ...». They are all very reasonable homilies.

Chapter 9 deals with liquidity problems.

We had to wait until 1984 for BCBS to start addressing the problems of bank liquidity: an *ad hoc* subgroup was established. This group was given the task of analyzing both from the conceptual as well as from the practical viewpoint the problem of liquidity of international banks. The measurement of liquidity level was perceived to be very difficult and too much dependent from market conditions and country rules. For instance the willingness of a central bank to accept a certain asset as collateral has a deep impact on an asset liquidity. The group reached the conclusion that the liquidity issue «shall be a matter for individual national authorities».

In January 1987, under the prodding of the subgroup-friendly BCBS Chairman, Peter Cooke, another liquidity subgroup was set up. It produced a report on «Measurement of liquidity adequacy» in September 1990. This subgroup continued its work and at the end of 1992 presented two papers, the first on «A framework for measuring and managing liquidity», the second on «The supervision of the liquidity of foreign establishments». After completion of the two reports, the subgroup ceased its activity. This subgroup had the merit to point out the difficulties connected with the concept and the control of liquidity. One main obstacle in drafting a common proposal was the correct (in my view) position of US and UK banks of first identifying a “critical time horizon” *i.e.* the length of time before which cash reserves from maturing and liquefiable assets were exhausted. I have previously mentioned that different national procedures make it very difficult to propose a common ground for liquidity control. Note that however, without an international agreement, any national central bank that sought unilaterally to tighten liquidity requirements of its banks would be hit by the standard level-playing field argument of weakening their competitiveness. The chapter lists the sources of “core liquidity”: cash and liquid assets (undefined); access to central banks credit; access to standby lines of credit from other sources, particularly in the interbank market. As shown in the 2007 bank crisis, with the exception of cash all other core liquidity sources are very shaky. The recent sovereign stress has increased the importance of liquidity control, but made its management more uncertain.

Chapter 10 deals with off-balance-sheet exposures and derivatives.

During the November 1984 Committee meeting, it was pointed out that «the increasing ingenuity of banks to develop techniques for removing business from their balance sheets and the appearances of new forms of contingent liabilities sug-

gests conducting a survey of the various off-balance-sheet activities conducted by banks». After that, the Committee secretariat prepared a list of six queries to be addressed to the Committee. The first of them was «Should irrevocable commitments under Note Issuance Facility be brought within the scope of capital adequacy measurement?». In July 1985 a study group under the chairmanship of Sam Cross of the FRBNY was established. Its conclusions formed a long (270 pages) report. While the work of the Cross group was still in its early stages, in early 1985, the subgroup-maniacs of BCBS set up another subcommittee. Its conclusions were not made public, but were taken into account in the 1988 Accord. In Autumn 1985, the latter subgroup presented a first report. Its conclusions were that «most of the risks associated with off-balance-sheet business were no different in substance from those of on-balance-sheet exposures and could not be treated in isolation». The report recommends a review of the existing reporting systems. It also suggests supervisors to warn banks of the risk involved in some option contracts. During the same period, (early '80s), the Basel Committee investigated two other problems: “Interbank netting schemes” and “Derivatives”. As usual a subgroup was established. The produced report was published five years after the publication by BIS of the important Lamfalussy report on netting problems! As we know the netting problem is that at the end of the day some counterparties may be left without sufficient liquidity, thus triggering complicated insolvency procedures. Nowadays this problem can be completely eliminated via the introduction of “gross payment systems” in which each transaction is transferred in real time to the interested party. This solution has been made possible by the great improvements in telecommunication technology. Regarding the second theme “Derivatives”, joined to “netting” by some mysterious reason, BCBS was forced to co-ordinate its work with Iosco, the International Organization of Securities Commission that supervises the major users of derivatives, investment banks.

Chapter 11. This chapter presents other topics addressed by the BCBS, most of them proposed by some members. These topics were possibly discussed, but not considered worth a particular study. This operational system was very active in the early years of life of BCBS, the topics and the possible following discussions (if any) were recorded in the (randomly produced) Reports. We can list the following suggested topics:

1. Bank secrecy.
2. Deposit protection schemes.
3. Lender of last resort techniques.
4. Bank profitability.

5. Prevention of criminal use of the banking system.
6. Large credit exposures.
7. Risk in computer usage and in telecommunication.
8. Supervision of conglomerates.
9. Bank insolvency regimes.

Some topics were clearly outside the scope of supervision, others, like 1. (Bank secrecy) and 6. (Large credit exposures) were not. In particular, the fact that this latter topic was not considered worth of any consideration is justified only by the technical incompetence of the Committee members. Indeed the adopted risk weights were a linear combination of assets volumes; unfortunately, they did not take into any account the percentage of each asset, *i.e.* its concentration.

Chapter 12 discusses the relationship of BCBS with banks and other banking regulators. The most interesting part of the chapter deals with the relationships with the non-G10 banking regulators. After intensive work the BCBS was able to draw a list of all centers offering “offshore” banking and financial services. They are divided into three groups: countries offering “offshore” financial services within their jurisdiction, Island Centers, and Mainland Centers.

The first group is composed by the following six locations: Ireland (Dublin), Morocco (Tangier), Taiwan, Thailand, UK (London), US (Colorado, Delaware, Miami, New York). The second contains 34 Island Centers from Anguilla to West Samoa. It also includes San Marino, that in spite of its name does not have any marine contact! The last group contains 12 countries from Andorra to Switzerland.

Under the leadership of Mr. Colin Powell, Economic Adviser to the State of Jersey, a Committee of 14 offshore centers was formed in 1980 and started a formal relationship with the BCBS. That led to the approval of a *memorandum* of understanding and to the establishment of joint working parties. Another non G-10 Banking supervisory group was the Gulf Co-operation Council (GCC), founded in 1981 by all states of the Arab peninsula. BCBS held a few meetings with GCC in which general issues were discussed.

In addition to the European Banking Supervision Committee, now European Banking Authority, and to IMF, in the '90s a *plethora* of regional Banking Supervisory groups was born (the book lists eight of them existing at the end of 1997, all that in addition to the International Conference of Banking Supervisors that begun its activity in 1979. Altogether, I have counted 14 (not including the many US banking regulators). Such groups do not seem to have been very efficient in preventing financial crises!

Chapter 13 discusses the relationships with other non-bank oversight and supervisory bodies. We must recall that the BCBS is not an independent group, but got its mandate from the G-10 central bank governors to whom it reports. The relationships between the BCBS and BIS, that provides a venue and secretarial services to BCBS, “is somewhat intricate”, but seems to be working due to the periodic meetings between Committee Chairmen and BIS Directors.

The BCBS was neither the first nor the only standing subcommittee set up by the G-10 governors. Perhaps the most important such committee has been the Euro-Currency Standing Committee, currently named Committee on the Global Financial System, set up in 1971. This is a very important committee with the main purpose to monitor macro-financial developments.

The BCBS established relationships with the two main accounting organizations, the International Accounting Standards Committee (Iasc), which mainly concentrated its activity on accounting issues, and the International Federation of Accountants (Ifac), that focuses its activity on auditing issues. The relationship with these two groups has become very important. Unfortunately, the co-ordination with the accounting organizations has not been very fruitful: the adoption of the Mark-to-Market rules has strongly increased the procyclicality of the Basel capital requirements. The BCBS had to co-ordinate with Iosco, the International Organization of Security Commission, with which relationships were almost dysfunctional. As we know Securities Commissions, like SEC in US, regulate investment banks and markets, and have been very lenient. Only in 1996 Iosco and BCBS issued a joint statement of intent.

Chapter 14 presents the legal position of the BCBS and points out the lack of any legal status: contrary to the Financial Stability Board who had a formal structure, the BCBS was just a committee of specialists set up by the G-10 governors. This may explain the reluctance of the Basel Committee to enlighten and inform the G-10 governors of the macro consequences of their proposals. Possibly the governors, or some of them, would have warned the monetary authorities of the respective countries about these dangers, and these monetary authorities could have stopped some follies.

Chapter 15 deals with the international relations of the BCBS and discusses five occasions of interface with the rest of the world.

Chapter 16 presents the use made by BCBS of social sciences. It made hardly any use of outside professional services. It recalls only two cases, when Chairman Tommaso Padoa-Schioppa corresponded with Robert Merton on the design of the Market Risk Amendment and when Chairman Gerard Carrigan asked lawyer

Ernest Patrikis, then general Counsel of the Federal Reserve Bank of New York, to write a report on cross-border bankruptcy issues following the BCCI collapse of July 1991. It is appalling the lack of any advice on technical statistical problems. This confirms my opinion that the BCBS was essentially a group of supervisors, interested only in reaching a consensus on possibly wrong positions. It is reasonable to suspect that the committee purposely hid the negative macro consequences of its proposals. The decision process in any case was not designed to take into account wider economic problems. The “owner” of the BCBS, the G-10 group of central banks governors almost never rejected any proposal from BCBS, especially if hidden by technicalities. The Emperor’s New Clothes syndrome prevailed.

Chapter 17 is the epilogue already commented in the first part of this review.

3. - Gender Discrimination and Other Arbitrary Decisions

Appendix B of Chapter 3 of the book contains important information on the *modus operandi* of the committee and the complete list of all participants to the committee meetings in the years 1975,1997. It shows that:

The six members of the *Groupe de Contact* were almost completely “disposed of” by 1976 with the sole exception of the German representative, Mr. Manfred Schneider, who clearly had a strong support by his home institution. It would have saved some taxpayer money to limit the country representation to just one member and to invite the six members of the *Groupe de Contact* as observers in the early meetings.

From the analysis of the distribution of the attendance to the meetings of various delegates, we can deduce that some countries did not take very seriously the role of the Committee. The case of Japan is extreme: in the period 1975-1997 it sent to Basel 49 different delegates. Only two of them attended at least 10 meetings (and possibly had some weight in the committee decisions). Was the attendance to Basel regarded a sort of paid holyday in Japan? All other countries provided at least two delegates that attended 20 or more meetings, thus providing a proper continuity in the committee operations.

Appendix B lists more than 400 different delegates to the committee in the years 1975-97. Among them only 12 were women coming from US, France, Belgium, Canada, and EU Commission. In addition, two of them attended only one meeting. This blatant discrimination contradicts the opinion expressed in the book about the quality of the reports produced by the many subgroups, task

forces and working parties. The three reports produced by the two groups chaired by women, namely Christine Cumming of the Federal Reserve Bank of New York and by Patricia Jackson of the Bank of England, both presented very relevant results, much superior to the other committee studies. Unfortunately, as already previously mentioned, they were completely disregarded by the wise men of the Committee!

BIBLIOGRAFIA

- ARTZNER P. - DELBAEN F. - EBER J.-M. - HEATH D., «Thinking Coherently», *Risk*, no. 10, 1997, pages 33-49.
- DANIELSSON J. - EMBRECHTS P. - GOODHART C. - KEATING C. - MUENNICH F. - RENAILT O. - SHIN H.S., «An Academic Response to Basel II», *www. RiskResearch.org*, 2001.
- FEDERAL RESERVE SYSTEM, *Credit Risk Models at Major US Banking Institutions: Current State of the Art and Implications for Assessment of Capital Adequacy*, Federal Reserve System Task Force on Internal Credit Risk Models, Washington, May 1998.
- FSA, *The Turner Review: A Regulatory Response to the Global Banking Crisis*, Financial Service Authority, London, March, 2009.
- GORDY M., «Granularity Adjustment in Portfolio Credit Risk Measurement», in SZEGÖ G.P. (ed.), *Risk Measures for the 21st Century*, Wiley, Chichester, 2004.
- IFI, *Report of the Working Group on Capital Adequacy*, Institute of International Finance, March, 1998.
- IOSCO, *Objectives and Principles of Securities Regulation*, International Organization of Securities Commissions, September, 1998.
- ISDA, *Credit Risk and Regulatory Capital*, International Swaps and Derivatives Association, March, 1998.
- JACKSON P. *et AL.*, «Capital Requirements and Bank Behaviour: The Impact of the Basel Accord», *BCBS Working Paper*, no. 1, Basel BCBS BIS, April, 1999.
- MARKOWITZ H.M., «Portfolio Selection», *The Journal of Finance*, no. 7, 1953, pages 77-91.
- .-, *Portfolio Selection*, Wiley, New York, 1959.

RIVISTA DI POLITICA ECONOMICA

Trimestrale

ABBONAMENTI

Gli abbonamenti (annuali) decorrono dal mese di gennaio al mese di dicembre. A coloro che dovessero abbonarsi durante l'anno, saranno inviati automaticamente i fascicoli precedenti relativi all'anno di sottoscrizione.

Gli abbonamenti non vengono rinnovati automaticamente, pertanto l'abbonato è tenuto a rinnovare il proprio abbonamento e saldare la relativa quota di sottoscrizione di anno in anno.

- Un fascicolo (anno in corso di pubblicazione): Euro 30,00 più spese di spedizione

- Arretrati: Euro 40,00 cad. più spese di spedizione

I numeri non pervenuti devono essere richiesti entro e non oltre il termine di sei mesi dalla stampa del numero che segue, altrimenti verranno considerati come arretrati.

Si prega di segnalare le variazioni di indirizzo via email a

laura.mori@licosa.com o via fax al numero: +39 055 641257

INFORMAZIONI E SOTTOSCRIZIONI:

LICOSA SpA

Via Duca di Calabria 1/1

50125 Firenze - Italy

Responsabile: Laura Mori

tel. +39 055 6483201 - 055 64831

fax +39 055 641257

e-mail: laura.mori@licosa.com - licosa@licosa.com

QUOTE ABBONAMENTO

Euro 120,00 Italia

Euro 145,00 Altri Paesi

MODALITÀ DI PAGAMENTO:

- Assegno non trasferibile intestato a: LICOSA S.p.A.

- c/c postale 343509 intestato a:

LICOSA S.p.A.,

Via Duca di Calabria 1/1

50125 Firenze

- Bonifico bancario intestato a:

LICOSA S.p.A.

c/o Monte dei Paschi di Siena

IBAN: IT88 Z 01030 02869

000004160064

SWIFT: PASCITM1W04

NORME EDITORIALI PER GLI AUTORI

La Redazione della *Rivista di Politica Economica* dovrà ricevere via e-mail a: rpe@confindustria.it o su CD-ROM il *file* pdf ed il corrispondente *file* in formato Word con le formule elaborate in Equation Editor delle versioni definitive da pubblicare. È altresì indispensabile inviare alla Redazione il documento per la cessione del *copyright* firmato dall'Autore/dagli Autori senza il quale non si potrà procedere alla pubblicazione dell'articolo.

Possono essere presentati per la pubblicazione contributi e saggi scientifici inediti, in lingua italiana o inglese. Tutti i lavori vengono sottoposti ad una prima valutazione del Direttore Responsabile. Superata questa, nel caso di saggio scientifico, il lavoro viene sottoposto in forma anonima a due *referee* scelti fra accademici ed economisti dal Direttore Responsabile, sempre in forma anonima. Il contributo scientifico o *invited paper* viene valutato solo dal Direttore Responsabile.

La responsabilità degli articoli e delle opinioni espresse è da attribuire esclusivamente agli Autori. I diritti relativi agli scritti contenuti nella Rivista di Politica Economica sono riservati e protetti a norma di legge.

E' vietata la riproduzione in qualsiasi lingua degli scritti, o anche di una sola parte di questi, apparsi sulla Rivista di Politica Economica, salvo autorizzazione del Direttore Responsabile.

IMPOSTAZIONI GRAFICHE

Prima pagina:

1) Il titolo del testo va reso al centro della pagina. Sotto di esso deve apparire il nome e cognome dell'Autore, con asterisco sul cognome e richiamo in nota che riporti l'indirizzo e-mail seguito eventualmente dalla Facoltà o Dipartimento o Istituto per le *affiliations* universitarie o dal Servizio in caso di Enti. Non va specificata in nota la carica dell'Autore/degli Autori né recapiti postali e/o telefonici/fax che vanno comunque comunicati separatamente alla redazione.

2) Nella terza riga, centrato sotto il nome dell'Autore deve apparire il nome dell'Ente/degli Enti, seguito/i dalla città.

3) Nel caso di più Autori di uno stesso saggio, i cognomi devono rendersi in ordine alfabetico, mentre l'asterisco di richiamo deve apparire esclusivamente sull'ultimo cognome. Nella nota vanno resi in sequenza gli indirizzi di posta elettronica, ognuno seguito eventualmente dalle specifiche sopra citate. Sempre nella stessa nota, vanno indicati i ringraziamenti ed il *disclaimer*.

4) I saggi vanno corredati da una bibliografia e da un *abstract* in inglese di lunghezza massima di 100 parole seguito dal/dai codice/i di classificazione del JEL (Journal of Economic Literature) da indicarsi tra parentesi quadre Es: JEL Classification [...]. I codici sono consultabili sul sito:

<http://www.aeaweb.org/journal/elclasjn.html>.

Si prega di indicare alla fine dell'*abstract* le *keywords* per l'indicizzazione del testo sul sito RePEc.

NORME GENERALI

- 1) Le note dovranno essere numerate progressivamente.
- 2) Il cognome di Autori citati in nota va reso in carattere maiuscoletto con a seguire l'iniziale del nome e l'anno di pubblicazione del testo citato. La sequenza nella Bibliografia dovrà rispettare l'ordine alfabetico ed è indispensabile che tutte le voci bibliografiche citate nel testo trovino riscontro nella bibliografia e viceversa, le citazioni incomplete devono essere eliminate.
- 3) La bibliografia deve riportare in maiuscolo maiuscoletto il cognome dell'Autore e l'iniziale del nome, il titolo dell'opera tra virgolette « », città, casa editrice ed anno di pubblicazione. In caso di citazione di più Autori senza nomi specifici, questi vanno resi come....*et al.* se all'interno del testo oppure.....*et AL.* se in nota.
 - a) Se l'opera è inclusa in altra pubblicazione va resa come segue:
KRUGMAN P., «Pricing-to-market when the Exchange Rate Changes», in SWEN W.A. - RICHARDSON D.J. (eds.), *Real Financial Linkages Among Open Economies*, Cambridge (Mass), MIT Press, 1987, pp. 49-70.
 - b) Se l'opera è inclusa in una rivista, va resa come segue:
KNETTER M., «Price Discrimination by U.S. and German Exporters», *American Economic Review*, March, no. 79 (1), 1989, pp. 198-210.
 - c) Se l'opera fa parte di Working Papers o simili, compresi mimeos, va resa come segue:
DARVAS Z. - ROSE A.K. - SZAPÁRY G., «Fiscal Divergence and Business Cycle Synchronization: Irresponsibility is Idiosyncratic», Cambridge (MA), *NBER Working Paper*, n. 11580, 2005.
- 4) Negli elaborati in italiano **le citazioni di brani di Autori stranieri andranno rese in italiano**, mentre i termini in lingua inglese vanno resi in carattere corsivo.
- 5) Le **Tavole**: i titoli delle tavole (obbligatori) devono apparire al centro delle stesse, in carattere maiuscoletto così come la numerazione progressiva che va indicata alla loro destra (Es: TAV. 1 in italiano e TABLE 1 in inglese). All'interno del testo, le citazioni delle tavole vanno rese con la sola iniziale maiuscola.
- 6) I **Grafici**: i titoli dei grafici (obbligatori) devono apparire al centro delle figure, in carattere maiuscolo così come la numerazione progressiva che va indicata alla loro destra (Es: GRAF. 1 in italiano e GRAPH 1 in inglese). All'interno del testo, le citazioni dei grafici vanno rese con la sola iniziale maiuscola.
- 7) Le **Formule**: devono seguire una numerazione progressiva (da indicare alla loro sinistra, tra parentesi tonde, in carattere corsivo).
- 8) La dicitura **Enunciato/Ipotesi**: va resa in carattere maiuscolo/maiuscoletto seguita dai due punti mentre il testo composto in tondo.
- 9) La **Proposizione**: va resa in corsivo seguita dai due punti ed il testo composto in tondo.
- 10) La **Dimostrazione di un'Equazione/Analisi**: va resa in carattere grassetto tondo così come il numero dell'equazione/analisi da indicarsi tra parentesi, seguito dai due punti.
- 11) L'**Esempio**: va reso in corsivo seguito dal testo composto in tondo.
- 12) Il **Presupposto**: deve riportare un titolo e seguire una numerazione progressiva da indicarsi in alto a sinistra.

RIVISTA DI POLITICA ECONOMICA

Fondata nel 1911

La Rivista di Politica Economica è stata fondata nel 1911 come «Rivista delle società commerciali» ed ha assunto l'attuale denominazione nel gennaio 1921. È una delle più antiche pubblicazioni economiche italiane ed accoglie analisi e ricerche di studiosi appartenenti alle varie scuole di pensiero. Gli articoli pubblicati nella Rivista sono citati in Econlit, e-JEL, JEL ON-CD, in RePec e nella International Bibliography of the Social Sciences.

La Rivista è trimestrale dal 2009 e dal 2010 si articola in due numeri ordinari, un fascicolo monografico ed il numero speciale pubblicato in inglese e dedicato ai saggi vincitori del Premio per Tesi di Laurea in Economia «Angelo Costa» e alla Lezione omonima.

I saggi proposti per la pubblicazione dovranno essere redatti conformemente alle «Norme editoriali per gli Autori» indicate nella Rivista ed inviati alla Redazione sia in forma stampata sia in formato digitale (una copia completa di nomi e recapiti degli autori ed una copia anonima senza dati):

Redazione RIVISTA DI POLITICA ECONOMICA

Viale Pasteur, 6 - 00144 ROMA - ITALIA

e-mail: rpe@confindustria.it - Tel. +39.06.5903 601 - Fax +39.06.5903 349

Sito internet: <http://www.rivistapoliticaeconomica.it>

Coordinatore editoriale: Adriana Leo

a.leo@confindustria.it - Tel. +39.06.5903 793

Garanzia di riservatezza. Il trattamento dei dati personali che riguardano l'abbonato viene svolto nell'ambito della banca dati elettronica della società LICOSA spa e nel rispetto di quanto stabilito dalla direttiva ex artt. 9/10/11, D.lgs. n. 196 del 30 giugno 2003 sulla tutela dei dati personali. Il trattamento dei dati, di cui garantiamo la massima riservatezza, è effettuato al fine di aggiornare l'abbonato su iniziative e offerte. I dati non saranno comunicati o diffusi a terzi e per essi l'abbonato potrà richiedere, in qualsiasi momento, la modifica o la cancellazione, scrivendo a LICOSA spa.



Servizio Italiano Pubblicazioni Internazionali S.p.A.
Viale Pasteur, 6 - 00144 Roma

Autorizzazione Tribunale di Roma n. 29 del 24-10-1950

Impaginazione: D.effe comunicazione - Roma

Stampa: Saro Italia Srl

Via Serafino Belfanti n°8, 00166 Roma

Finito di stampare nel mese di marzo 2013

Prezzo € 30,00



La Rivista di Politica Economica è stata fondata nel 1911 come Rivista delle Società Commerciali ed ha assunto l'attuale denominazione nel gennaio del 1921. È una delle più antiche pubblicazioni economiche italiane ed accoglie analisi e ricerche di studiosi appartenenti alle varie scuole di pensiero. Come 100 anni fa, la Rivista di Politica Economica nutre ancora "la fondata speranza di rendere un servizio utile all'educazione civile del nostro paese" grazie alla pubblicazione di scritti economici di valore e rilievo.

Saggi scientifici:

L'economia non osservata fra evasione e crimine: una rivisitazione del *Currency Demand Approach* con una applicazione al contesto italiano
**Guerino Ardizzi - Carmelo Petraglia
Massimiliano Piacenza - Gilberto Turati**

Previsioni delle spese del bilancio dello Stato attraverso i flussi di contabilità finanziaria
Giuseppe Bianchi - Tatiana Cesaroni - Ottavio Ricchi

On the Fuzzy Boundaries between Public and Private in Health-Care Organization and Funding Systems
Alessandro Petretto

Alcune implicazioni fiscali di scenari alternativi alla proibizione della cannabis
Marco Rossi

Disuguaglianze, beni relazionali ed elezioni nelle periferie di Roma
Federico Tomassi

ANNO CII - SERIE III
gennaio/marzo 2013
Fascicolo I-III

Trimestrale - Poste Italiane S.p.A.
Sped. abb.post. - D.L. 353/2003
(conv. in L. 27/2/2004 n.46) art. 1 co. 1
DCB Roma - ISSN: 0035-6468