

RIVISTA DI

POLITICA

VII-IX

ECONOMICA

CONTRIBUTI SCIENTIFICI:

Alessia Amighini
Sebastiano Bavetta
Chiara Bentivogli
Emiliano Brancaccio
Tommaso Frattini
Francesco Giavazzi
Giovanni Leghissa
Marcello Messori
Massimo Morelli
Pietro Navarra
Eugenia Panicara
Alan R.S. Ponter
Annalisa Rosselli
Jean Tirole
Daniela Venanzi

RIVISTA DI

POLITICA

VII-IX



CONFINDUSTRIA

ECONOMICA

IN QUESTO NUMERO:

Alessia Amighini
Sebastiano Bavetta
Chiara Bentivogli
Emiliano Brancaccio
Roberto Dell'Anno
Marisa Faggini
Tommaso Frattini
Emma Galli
Francesco Giavazzi
Giovanni Leghissa
Marcello Messori
Massimo Morelli
Pietro Navarra
Roberta Occhilupo
Giuliana Palumbo
Eugenia Panicara
Filomena Pietrovito
Alan R.S. Ponter
Roberto Ricciuti
Annalisa Rosselli
Paolo Sestito
Jean Tirole
Daniela Venanzi

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

Carlotta Berti Ceroni

Pierpaolo Benigno

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

Fiorella Kostoris

Elisabetta Iossa

Luca Lambertini

Stefano Manzocchi

Riccardo Martina

Alessandro Missale

Giuseppe Moscarini

Giovanna Nicodano

Francesco Nucci

Luigi Paganetto

Luca Paolazzi

Annalisa Rosselli

Alberto Quadrio Curzio

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 2012

Indice

Invited Paper

Union Integration Paths Massimo Morelli	7
--	---

Invited Policy Paper

Freedom and Redistribution Sebastiano Bavetta, Pietro Navarra	29
--	----

Regolazione decentrata e servizio concentrato: le ferrovie regionali viaggiano su un binario stretto? Chiara Bentivogli, Eugenia Panicara	51
---	----

A New Textbook Approach to Macroeconomics: A Debate Alessia Amighini, Emiliano Brancaccio, Francesco Giavazzi, Marcello Messori	101
--	-----

British and Italian Universities and the Anglo-Saxon Model Alan R.S. Ponter	131
---	-----

Preventing and Resolving Banking and Sovereign Crises: Toward a New Institutional Framework Jean Tirole	143
---	-----

La manifattura in Italia nell'ultimo decennio, sulle spalle dei... nani Daniela Venanzi	155
---	-----

Saggi

Crescita economica ed economia sommersa. Un approccio di simulazione e verifica empirica per l'Italia Roberto Dell'Anno	193
--	-----

Il controllo dell'economia: bilancio e prospettive Marisa Faggini	229
--	-----

Sulla <i>political economy</i> della spesa pubblica nell'Italia liberale Emma Galli, Roberto Ricciuti	275
Le scelte di localizzazione delle opere pubbliche: il fenomeno <i>Nimby</i> Roberta Occhilupo, Giuliana Palumbo, Paolo Sestito	293
Financial Development and Economic Growth. A Theoretical and Empirical Overview Filomena Pietrovito	323
Parole chiave	
Immigrazione Tommaso Frattini	363
Rassegna bibliografica	
RODRICK D. <i>The Globalization Paradox: Democracy and the Future of the World Economy</i> a cura di Giovanni Leghissa	411
BATEMAN W.B., HIRAI T. e MARCUZZO M.C. (eds.) <i>The Return to Keynes</i> a cura di Annalisa Rosselli	425

INVITED PAPER

Union Integration Paths

Massimo Morelli*

Columbia University

This paper proposes a theoretical explanation for the incremental integration path of Europe until the Monetary Union time. The dynamic partition function game suggests that the presence of externalities across coalitions affects the timing of admission of different players, the equilibrium coalition structure, and the equilibrium length of the admission game. In particular, at the early stages of integration it is reasonable to believe that the bargaining power of the founders was not so high as to make the admission process resemble a sequential ultimatum game, and in such a context the admission path is always a sequencing admission path.

[JEL Classification: C70; D70; F15].

Keywords: coalition; union; externalities; side payments; commitment.

* <mm3331@columbia.edu>, Department of Political Science and Department of Economics. I thank Philippe Penelle for our preliminary work together on this topic in the 90s. This paper is based on a reevaluation and update of those notes. Hence any error in the new results is my own.

1. - Introduction

In a world where “acting together” is beneficial, one should see the players “act together” rather quickly, and enjoy the fruits of cooperation as soon as possible. Why then did it take over 30 years for the European Union (EU) to enlarge from the initial six members to fifteen members, and what could instead explain the all-at-once jump to 27 after that?

One simple explanation (which can be labeled as the *exogenous process* explanation) is that some countries were just not “good enough” for the EU in the early sixties, *i.e.*, were *not admissible*, but became “good enough” in the eighties. Any such exogenous process explanation would have difficulties explaining why the process has been incremental, without ever a step back.¹

In the theory developed in this paper the sequencing path is not due to an exogenous process of admissibility. I will investigate conditions under which delaying the admission of *already admissible* countries is consistent with rational behavior. A *necessary*, albeit not sufficient, condition for this to happen is the existence of a *negative externality* suffered by countries left out of an enlarging union. Intuitively, this negative externality affects the *price* the union has to pay to induce a candidate country to join at various stages of the integration process. Since the union is a “veto” player, it can structure the admission path to take advantage of this “price effect”, which may involve sequencing the admission of admissible countries.

One of the main insights of this paper is that the optimality of a “sequencing” strategy depends crucially on the existence and nature of side-payments available to various groups of countries. In a world with no restriction whatsoever on the nature and direction of side-payments, the negative externalities created by the sequential enlargement of the union *are fully internalized* in the price the union has to pay to candidate countries in order to induce them to join. This “price ef-

¹ In the NBER working paper version of MAGGI G. and MORELLI M. (2006), we provided such an explanation of an International Organization expansion in the presence of an exogenous process of incremental eligibility.

While in some contexts it is reasonable to consider an exogenous incremental eligibility process, for example when thinking of human rights violations conditions for admissions to humanitarian organizations, it is less convincing when the topic is an economic one, where we all know that cycles and shocks in all directions are more reasonable than assuming a constantly increasing trend. It is mostly for this reason that the 2006 published version does not include the expansion section, limiting attention to fixed size organizations.

fect” undoes the incentive a forward looking rational union has to sequence admissions, making “bundling” a strictly superior strategy. In the real world, commitment problems may severely limit the role of side-payments as a mechanism to force an enlarging union to internalize the costs it imposes on other countries.² Finding the existence of commitment problems in the absence of enforcement institutions a compelling starting point, all my findings will relate directly or indirectly to potential different degrees of such commitment and enforcement problems that limit the viability of side payments.

I study the equilibrium admission path under various commitment assumptions, finding scenarios such that the sequencing of admissible countries may be justifiable.³ The presence of externalities imposes the adoption of the *partition function approach*. Games in partition function form differ from games in characteristic form in the sense that the payoff available to a coalition depends on the coalition structure (externalities).⁴

Most of the results obtained in the literature on partition function games are valid for identical players, whereas here heterogeneous players are the heart of the matter.⁵

Economic integration of countries is a problem where heterogeneity and dynamics have to be introduced in order to say something meaningful. In this paper I do this, assuming that there exists a subset of countries that act as coalition developers.⁶

² The lack of use of international side-payments to resolve trade disputes has been pointed out by KOWALCZYK C. and SJÖSTRÖM T. (1994). In their lucid discussion of the shortcomings of GATT as an international institution designed to facilitate, and enforce results of trade negotiations, they suggest that GATT *should* encourage and *enforce* agreed upon income transfers across countries, in the context of trade negotiations.

³ On a similar topic, see ALESINA A., ANGELONI I. and ETRO F. (2005) and BORDIGNON M. and BRUSCO S. (2001).

⁴ Partition function games have been studied first in THRALL R. (1962); LUCAS W. (1963) and LUCAS W. and THRALL R. (1963). More recently, the prominent role of this class of games has been emphasized in RAY D. and VOHRA R. (1997); BLOCH F. (1996); YI S.-S. (1997) and YI S.-S. and SHIN H. (2000).

⁵ There are sharp results (see BLOCH F., 1996; YI S.-S. and SHIN H., 2000) for the simple case where players are identical and payoff division is exogenously fixed, while only in RAY D. and VOHRA R. can one find some preliminary characterization results for the general case.

⁶ The survey by BLOCH F. (1997) contains references on coalition developers in other contexts, like in principal-agent problems. This setting is fully developed in GENICOT G. and RAY D. (2006). Other less directly related papers on the endogenization of expansion paths of a coalition are ROBERTS K. (1999); BARBERA S., MASCHLER M. and SHALEV J. (2001); ACEMOGLU D., EGOROV G. and SONIN K. (2008).

While the main analysis is conducted assuming complete information, we will also highlight the importance of side payments, commitment and enforcement in a world with shocks and crises. Finally, I will briefly relate the insights of the paper to an ongoing research agenda aimed to explain the desirability of fiscal integration for the future of Europe. Morelli *et al.* (2012); Luque *et al.* (2012) and Herrera *et al.* (2012) develop, respectively, a distributive reason, a volatility related reason, and a cultural integration reason for believing that the fiscal union step is necessary at this point of European integration history. This paper does not add a fourth independent reason, but rather offers a theoretical evaluation framework for the role of side payments, commitment and enforcement, which all have consequences for each of the three separate arguments developed in the parallel work on fiscal integration.

The paper is organized as follows. I first provide a succinct historical account of the incremental integration path that took place until the conception of the Euro, based on the preliminary notes in Morelli and Penelle (1997). Then I will present a formal admission game that delivers in a simple way all the comparisons I aim to illustrate among different institutional scenarios, albeit stylized. Then I will conclude by analyzing the intuitive impact of uncertainty, distributive concerns, volatility and cultural heterogeneity.

2. - A Brief History of EU Integration

The story of EU integration is a story of subtle balance between *widening* and *deepening*. The widening of the union refers to the enlargement of the founding union to new members. The deepening of the union refers to the deepening of the scope of integration, from a common market to an economic union. This section briefly documents some important stages of EU integration, especially with respect to the temporal dimension.⁷

From the signature of the Treaty of Rome in 1957 to the admission in the European Union (EU) of Austria, Finland and Sweden (since January 1, 1995), the evolution of the process of integration has been quite remarkable. The founding six members of the European Community (EC), Germany, France, Italy, Belgium, the Netherlands and Luxembourg, had all the ingredients required to form

⁷ A good reference for historically inclined readers is GEORGE S. (1994). This section draws heavily from FRANKLIN M. (1990).

a customs union by the end of the agreed upon transition period in 1969.⁸ They were joined in 1973 by Denmark, Ireland and the UK, Greece was admitted in 1981, Spain and Portugal in 1986. Joining the EU in 1995, Austria, Finland and Sweden increased the size of the union to 15 countries.

The blueprint for what became the EU was signed in 1987 by the Twelve, and is referred to as the Single European Act (SEA). This blueprint was amended and extended in 1993 with the signature of the Maastricht Treaty.

The case of the United Kingdom is of particular interest as Britain declined to join the European Coal and Steel Community (ECSC) that had been created by French Foreign Secretary Robert Schuman (often referred to as the Schuman Plan). Britain was reluctant to relinquish control over the Coal and Steel industries which had just been nationalized by the Labour Government. The British Government was, however, favorable to the idea of creating a free-trade area, which led to the creation of EFTA. When the UK finally decided to apply for EC membership in 1961, it was more driven by the fear of seeing a successful EC become the main trading partner of the US (a role traditionally assumed by the UK) than by overwhelming enthusiasm for the European idea.⁹

The first two British attempts at joining the EC in 1961 and 1967 were cut short by French President Charles de Gaulle, who claimed that Britain was not ripe for membership and would destroy the union. It took the UK the departure of de Gaulle in the aftermath of the events in 1968, and his replacement with Pompidou, to finally be admitted in the European club.

Widening the union has not always gone as smoothly as hoped for by applying countries, and the case of the UK is barely an exception. As it did for the United Kingdom, it took Denmark and Ireland three applications to be granted membership (1961, 1967, 1970). Norway applied in 1961, 1967, 1970 and 1992, was granted membership twice (1970 and 1992), but did not join following negative *referenda*. Austria applied in 1989 and was told to wait the completion of

⁸ An economic union can take several forms (see BELASSA B., 1961 for example). The weakest form is a free-trade area, which only requires the member states to abolish tariffs, but does not require them to apply identical tariffs to the rest of the world, which is the case for a customs union. A common market is a customs union with the added requirement of free movement of capital and labor. An economic union is a stronger form of common market which requires the participating members to achieve some degree of policy harmonization in order to minimize potential discrimination as a result of policy divergence. Finally, total economic integration presupposes the unification of monetary, fiscal, and social policies.

⁹ The fear of the negative externalities from staying out of the Union is exactly what will be used as primary ingredient in the theory.

the single market (SM). Turkey and even Morocco bid for membership in 1987. They were both denied admission.

The pattern of application for membership described above illustrates an interesting feature of the application/admission process that took place. British applications in 1961, 1967 and 1970 were immediately followed by applications from Ireland, Denmark and Norway.

The UK, Ireland and Denmark were admitted jointly in 1973. Spain and Portugal applied jointly in 1977 and were admitted jointly in 1986. In 1990, the Prime Ministers of Austria and Hungary declared jointly that their common aim was for their countries to become members by the time the SM would be completed. Austria was later admitted, Hungary, however, did not make it.

The EU decision makers have often been criticized in the media for moving at a snail pace. There have been episodes of conflicts, threats, and even major crises jeopardizing the very future of the union. The two most significant crises happened in 1965 and 1979-1984, respectively triggered by France and the UK. Despite the crises and disputes that have paved the way towards European integration, not a single member has ever left the EU.

The first goal of the subsequent analysis is to offer a theoretical explanation for the above described sequencing path, at an abstract level. At the same time, I will argue that the theory can be made consistent with the sudden jump in widening to 27 members and with the risks of break down that started to worry politicians and markets after the financial crisis of 2007 and the debt crises of the following years. Given the importance of commitment, enforcement and side payments that will be elucidated by the theory, I hope to provide some additional arguments in favor of the point of view that economic integration is not independent of political integration. Indeed, total economic integration presupposes the unification of monetary, fiscal and social policies, and requires the setting-up of supranational authorities whose decisions are binding for the member states.

3. - Union and Admission Game

Denote by π any partition of an initial set of N players. Consider as starting point the partition π_0 with all singletons, like in Europe after World War II. If the game is a balanced superadditive coalitional game with transferable utility, then the Core of the game is nonempty and the grandcoalition forms immediately. Given that in reality the grandcoalition very seldom forms, and, in particular, given

the history of the European Union discussed in section 2, we are lead to the conclusion that some, if not all, of the assumptions invoked by the above result must not be satisfied in the economic integration problems we are interested in.

When the Core is empty, the discussion is open on what should be the right way to analyze the situation. The prediction depends on the rules of the game. For example, in a three-player simple distributive game with an empty core, the possibility of *binding agreements* can lead to reasonable predictions: if two countries can form a binding agreement and then bargain with the third country, then an efficient but asymmetric outcome can be predicted. If these kinds of binding agreements are allowed, and if all kinds of transfers of utility are possible, the game between the new player – the union of two players – and the third one is different from the original one and admits a solution. This is indeed what I do here: I identify a simple criterion that may have determined the formation of the initial binding agreement of the first members of the European Community, and then I interpret the rest of the game as an admission game between the initial united members and the rest of the players. This asymmetry between players already bound by common institutions and the rest of potential members is then going to be kept also in the dynamic analysis of the periods after the initial union formation stage.

Denoting by $V(S, \pi)$ the value of coalition S in a partition π (game in partition function form), I assume throughout the paper the following *negative externality* assumption:

ASSUMPTION 1: for any coalition S , $V(S, \pi) > V(S; \pi')$ whenever π is a “refinement” of π' , *i.e.*, whenever to go from π' to π we break some of the coalitions in π' in subsets.

An implication of this assumption is of course that the value of any coalition S is maximized when all other players remain singletons. Since the European Community was born primarily for reasons related to trade, it seems clear that a trade union can rip a larger share of the gains from trade in the absence of other unions.

3.1 Identification of the First Binding Agreement

Let V^i denote the value of remaining autarchic when the partition is π_0 . Denote by $\pi_0(N \setminus S)$ a partition where all the players outside a coalition S remain singletons. Consider the coalition

$$S^* \equiv \arg \max_S \frac{V(S, \pi_0(N \setminus S)) - \sum_{i \in S} V^i}{\# S}$$

Suppose that such S^* is unique.

Given that by construction S^* constitutes (if one limits attention to the players in S^*) a balanced game, there exist binding agreements on how to share $V(S^*, \pi_0(N \setminus S^*))$ that all members of S^* can sign such that S^* can be assumed to act as a unique player after that.

To see that an initial group of countries like S^* can be thought of acting united in the subsequent admission game that will be analyzed next, consider the vector $\alpha_{S^*}^*$ of shares such that

$$\alpha_{S^*}^i V(S^*, \pi_0(N \setminus S^*)) - V^i = \alpha_{S^*}^j V(S^*, \pi_0(N \setminus S^*)) - V^j$$

for every i and j in S^* .

Suppose that there exists a pair (α_T, T) , for some T , such that (α_T, T) would make all the members of T strictly better off than in S^* , of course respecting feasibility. If this was the case, then there would be an obvious contradiction with the definition of S^* .

If the game is superadditive, there is no loss for the countries in S^* to sign the binding agreement, form the union, and then let the union bargain about the division of the remaining *surplus* in what we call the admission game: in fact, the bargaining power of the union can never be lower than the individual one of each member of the union, and hence the share of the remaining surplus that will go to the union when all the gains from superadditivity are used can only be greater.

3.2 The Admission Game

Let S denote an existing union.¹⁰ We treat S as one single player at time 1 of the admission game. Let B and C denote two candidate countries for admission.

¹⁰ The previous section has simply given a suggestion about what may determine the exact composition of the initial Union, but obviously the admission game described in this section does not depend in any way on whether the initial set of players in S is indeed determined by the simple logic of the previous section or by some other reasons.

V_t^i denotes the payoff for player i ($i = S, B, C$) when the three players remain separate at time t ; $V_t^i(j, k)$ denotes the date t payoff for country i when countries $j \neq i$ and $k \neq i$ form a coalition. The aggregate payoff for countries j and k is denoted $V_t(j, k)$ and $V_t(S, B, C)$ denotes the aggregate payoff of the grand coalition at date t .

In each period the rules of the game for that period depend on the degree of commitment to side payments, within and across coalitions, and on the relative bargaining power of the various countries. In the benchmark game form, first the current union (*i.e.*, S at the first period but then for example the new player $S \cup B$ in period 2 if the union $S \cup B$ formed in period 1) makes a take it or leave it sharing offer to a chosen subset of other players not yet admitted; then, if the set of proposed new members accept (or a subset of them do), the old union plus the new accepting members becomes the new union player next period. Those who reject the offer stay as separate players. Beside considering this benchmark *sequential ultimatum game* (SUG henceforth), I will consider two variants of it, characterized by different assumptions about the ability of the other players to commit to side payments among themselves or form separate binding agreements:

A *sequential admission game with external side payments* (SAGESP henceforth) is like the SUG but the players outside the union who are not being offered admission can commit to side payments to the players who are offered admission in case the offer is declined. The motivation to offer such side payments to convince a prospective union member to decline the current admission offer is clear given the assumption of negative externalities, but the power to commit to such external side payments is what is hard to obtain.

Finally, we will consider a third variant, a *maximum outside option game* (MOOG henceforth): here the power of the current union to make a take-it or leave-it offer finds its counterpart in the assumption that if the offer is rejected the remaining players (those who rejected the offers and the others) can form their own most convenient binding agreement among themselves.

To make things comparable, I will assume that in any case

$$V_t(S, i) - V_T^S - V_t^i > V_t(i, j) - V_t^i - V_t^j \text{ for some } i, j$$

that is, I assume that at any point in the admission game the gains from joining the current union in terms of extra *surplus* are larger than from the formation of an alternative binding agreement, at least for one of the remaining players outside the union. With this assumption, I believe that the three types of admission game

rules considered are the most reasonable and consistent alternative scenarios in terms of relative power between the current union and the external members.

An additional distinction will need to be made in terms of *internal compacting* along the admission path: when the union gets larger at time t , I will have to distinguish the case in which the union at time $t - 1$ still dictates the distribution of *surplus* after time t (hence potentially altering the treatment of previously admitted countries) from the case of perfectly transferable utility in which the new enlarged union becomes the effective new take-it or leave-it offer maker of the next period and the internal treatment of previously admitted countries cannot be renegotiated downwards by the initial union player.

3.3 Analysis

Let the partition function be the following:

$$\left. \begin{aligned}
 1) \quad & V_t^S = 3 & V_t^B = 3 & V_t^C = 2 \\
 & V_t(B, C) = 5 + \eta & V_t(S, C) = 9 & V_t(S, B) = 7 \\
 & V_t^S(B, C) = 3 - \varepsilon & V_t^B(S, C) = 2 & V_t^C(S, B) = 0 \\
 & V_t(S, B, C) = 10 + \gamma
 \end{aligned} \right\}$$

This partition function satisfies the negative externality assumption, and we will be able to understand the features of the various types of games by changing the parameters.

For sufficiently small ε , γ and η , the socially efficient coalition structure is $\{(S, C); B\}$, which produces a total of $9 + 2 = 11$.

Let $\delta \in (0, 1)$ denote the common discount rate used by the three players. Let $M > 2$ denote the number of periods the game is played over. Let $X_t^i(S, i)$ be the payoff received by country $i = B, C$ at date t , if in a coalition with S . Let $X_t^i(S, B, C)$ denote the payoff received by country $i = B, C$ if in the grand coalition.

Suppose that the union S can make admission offers to the countries she cares to admit, where these take-it-or-leave-it offers specify a transfer payment X_t^i for country i in each and every period if the offer is accepted. Each recipient of an offer has to decide whether to accept or reject the offer. Depending on which kind of “contracts” B and C can sign (or, in other words, on whether the game is a *SUG*, a *SAGESP* or a *MOOG*), both the offer and the incentives to accept or reject can obviously vary.

Let $S_0 = S$, $S_1 = S \cup B$, $S_2 = S \cup C$ and $S_3 = S \cup B \cup C$. Each of them (S_k , with $k = 0, 1, 2, 3$) is a possible “end-union.” Let π_k^1 denote the path of admission leading to the end-union S_k that involves admitting $S_k \setminus S$ at once in the first period. Let π_3^2 (π_3^3) denote the path of admission leading to S_3 that involves admitting B (C) in the first stage, and C (B) in the second stage.

Consider first π_1^1 . To induce B to join, S has to offer B

$$\begin{aligned} X_t^B &= 3 \text{ in } \text{SUG} \\ &5 \text{ in } \text{SAGESP} \\ &3 + \eta \text{ in } \text{MOOG} \end{aligned}$$

One can compute the present value of the path π_1^1 for S as:

$$\begin{aligned} W(\pi_1^1, M, \delta) &= 4 \frac{1 - \delta^M}{1 - \delta} \text{ in } \text{SUG} \\ &2 \frac{1 - \delta^M}{1 - \delta} \text{ in } \text{SAGESP} \\ &(4 - \eta) \frac{1 - \delta^M}{1 - \delta} \text{ in } \text{MOOG} \end{aligned}$$

Consider now π_2^1 . To induce C to join, S has to offer C

$$\begin{aligned} X_t^C &= 2 \text{ in } \text{SUG} \\ &3 \text{ in } \text{SAGESP} \\ &2 + \eta \text{ in } \text{MOOG} \end{aligned}$$

Thus, the present value of the path π_2^1 for S is:

$$\begin{aligned} W(\pi_2^1, M, \delta) &= 7 \frac{1 - \delta^M}{1 - \delta} \text{ in } \text{SUG} \\ &6 \frac{1 - \delta^M}{1 - \delta} \text{ in } \text{SAGESP} \\ &(7 - \eta) \frac{1 - \delta^M}{1 - \delta} \text{ in } \text{MOOG} \end{aligned}$$

REMARK 1 *Conditional on desiring to admit one country only, for η small enough the SAGESP scenario is the least desirable by the union player, even worse than the MOOG scenario.*

Consider now π_3^1 . To induce B and C to join, the payment structure has to be such that:

$$\begin{aligned} X_t^B(S, B, C) + X_t^C(S, B, C) &= 2 \text{ in } \textit{SUG} \\ &5 \text{ in } \textit{SAGESP} \\ &5 + \eta \text{ in } \textit{MOOG} \end{aligned}$$

One can thus compute the present value of the path π_3^1 for S as follows:

$$\begin{aligned} W(\pi_3^1, M, \delta) &= (8 + \gamma) \frac{1 - \delta^M}{1 - \delta} \text{ in } \textit{SUG} \\ &(5 + \gamma) \frac{1 - \delta^M}{1 - \delta} \text{ in } \textit{SAGESP} \\ &(5 - \eta - \gamma) \frac{1 - \delta^M}{1 - \delta} \text{ in } \textit{MOOG} \end{aligned}$$

REMARK 2 *In contrast with the previous remark, when the union wants to admit both countries at the same time the worst scenario for her is the MOOG one if $\eta > \gamma$.*

Having completed the computations for the “all at once” paths, I now need to analyze the two “sequencing” paths. Consider first π_3^2 , along which B is admitted first and C second. In the first stage, B has to receive the same treatment as in the computations for π_1^1 (first period) to be induced to join. In the second period onwards the same argument applies as along the path π_3^1 in the case of no internal compacting of the newly enlarged union, whereas in the case of internal compacting the new take-it or leave-it offerer is $S \cup B$ and hence the remaining player can only receive her outside option, without need of distinguishing among the three external scenarios,¹¹ and country B has to continue to receive at least as much as in the first period of partial expansion.

¹¹ The SAGESP scenario could in principle be modified by allowing the external player to have an even greater (and even less realistic) ability to commit to future transfers, *i.e.*, allowing an external player not only to commit to transfers in order to “buy” a rejection today, but also (or in alternative) to “buy” future behavior of an admitted member. This degree of commitment is so unrealistic that I have decided to ignore it altogether.

Denoting by *IC* and *NIC* the two alternative internal compacting and no internal compacting scenarios, it follows that the present value of this path for *S* is:

$$\begin{aligned}
 W(\pi_3^2, M, \delta) &= 4 + \delta(8 + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{SUG} - \textit{NIC} \\
 &4 + \delta(7 + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{SUG} - \textit{IC} \\
 &2 + \delta(5 + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{SAGESP} \\
 &4 - \eta + \delta(5 - \eta + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{MOOG} - \textit{NIC} \\
 &4 - \eta + \delta(7 - \eta + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{MOOG} - \textit{IC}
 \end{aligned}$$

REMARK 3 *The possibility of internal compacting or internal commitment is detrimental for S in the SUG scenario but advantageous in the MOOG one, whereas it doesn't matter in the SAGESP. Moreover, it should be noted that for sufficiently small η and sufficiently large M and δ , $W(\pi_3^2, M, \delta)$ can be greater than $W(\pi_3^1, M, \delta)$ in the MOOG-IC scenario.*

Finally, consider π_3^3 along which *C* is admitted first and *B* second. In the first stage, *C* has to receive one of the treatments of π_2^1 to be induced to join. From the second period onwards the same argument applies as along the path π_3^1 in the case of no internal compacting. It follows that the present value of this path for *S* is:

$$\begin{aligned}
 W(\pi_3^3, M, \delta) &= 7 + \delta(8 + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{SUG} \\
 &4 + \delta(5 + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{SAGESP} \\
 &7 - \eta + \delta(5 - \eta + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{MOOG} - \textit{NIC} \\
 &7 - \eta + \delta(6 - \eta + \gamma) \frac{1 - \delta^{M-1}}{1 - \delta} \text{ in } \textit{MOOG} - \textit{IC}
 \end{aligned}$$

that is,

REMARK 4 *The sequencing path π_3^3 dominates bundling admissions at time 1 (π_3^1), in both MOOG-NIC and MOOG-IC, and for every δ .*

All the above yields a proof by example of the following proposition:

PROPOSITION 5 *Given the existence of negative externalities, the admission game in all its forms can lead to different types of inefficiencies:*

1. *There exist parameter values such that the efficient path π_2^1 is the unique equilibrium path in SAGEST and MOOG situations, with and without internal compacting, whereas the inefficient immediate widening prevails in the SUG (inefficient for $\gamma < 1$).*
2. *There are other parameter values (for sufficiently high δ and γ) for which equilibrium inefficiency prevails also in a MOOG – IC, where the inefficient sequencing path π_3^2 is chosen.*
3. *For $\gamma > 1$ the grand coalition would be total utility maximizing; however, in the MOOG – IC, one of the two sequencing paths is always chosen: π_3^2 for sufficiently high δ and π_3^3 below that threshold.*
4. *In SAGESP there is never sequencing and the equilibrium path is always the efficient one. In SUG there is never any equilibrium sequencing but the equilibrium can be inefficient for low values of γ .*

In words, the generic example of a partition function used so far shows that when the grand coalition is efficient, the presence of negative externalities in a game with commitment within coalitions but no external transfers across coalitions (description in words of the MOOG – IC scenario) can create incentives to inefficient sequencing, and with high δ a particularly inefficient one can prevail. If commitment to transfers is possible even across coalitions (the SAGESP case), efficiency is always guaranteed, with π_2^1 chosen for $\gamma < 1$ and π_3^1 when $\gamma > 1$. Finally, in a SUG the inefficiency goes the other way, making the grand coalition chosen even when it is not efficient.

3.4 Interpretation of EU History Given the Equilibrium Analysis

As already mentioned, the existence of enforcement of (and commitment to) transfers between players or coalitions that are not formally bound by any supra-national institution is questionable, hence the always efficient outcomes of the SAGESP scenario can be ignored in this interpretation section.

The MOOG case seems to be the best suited to interpret the first 40 years of EU history: as the discussion of the UK perspective and the history of initial integration steps highlighted in section 2 more generally witness, the bargaining power of the initial members was limited by the best outside option opportunities for the other admissible countries, which were plentiful at the initial stages of the integration process. Thus, one way to explain the experienced sequencing admission path is to say that even though the coalition of 15 could already deliver the maximum total *surplus* (superadditivity in gains from trade and economies of scale) much earlier, sequencing prevailed because sequencing of one kind or another is always preferred by the initial Union members in equilibrium if the admitted members have to receive their maximum outside option and admission implies a subsequent internal compacting, with implicit commitment not to renegotiate downwards the terms of *surplus* sharing at the time of admission.

As integration kept deepening as well as widening, the strength and bargaining power of the enlarged Union grew over the years. The creation of a single currency and the use of such a more stable currency in all economic transactions arguably made the Union a stronger and more attractive entity for outsiders, or at least a stronger provider of negative externalities for the countries remaining outside. It is perhaps for this reason that the largest expansion step happened more or less at the same time or immediately after the creation of the Euro. In terms of the theory, one could argue that the increased strength provided by the Euro transformed the game in a *SUG*, hence explaining the enlargement at once, perhaps even including countries violating superadditivity.¹²

Given the reduced form of the partition function games and the assumption of certainty and complete information, the above interpretation hints are the maximum that we can extrapolate from our theoretical framework so far. However, we can now turn to analyze the potential implications of introducing un-

¹² MORELLI M., YANG H. and YE L. (2012) provide an alternative explanation for the expansion to 27 countries of the 21st century: while the poor should always prefer fiscal integration and the rich should always oppose it on distributive grounds, the middle class preferences for fiscal integration *versus* fiscal autonomy depend on mobility and income distribution within the Union. Arguably the increased mobility and all other conditions before the enlargement were such that the median voter of Europe was shifting her preferences in the direction of fiscal integration. The one-step admission of many poorer countries has determined a discontinuous jump in such median voter's preferences, away from the fiscal integration adversed by the elites. We will come back later to the desirability of fiscal integration in contexts more general than a purely redistributive one.

certainty or shocks in the value of the various countries and/or coalitions, and with such an extension it will be possible to provide additional hints on the nature of the dynamic possibilities for the Union under different scenarios of institutional integration.

3.5 *Uncertainty, Shocks and the Desirability of Fiscal Integration*

Suppose now that every period $V_t(S, \pi)$ can be subject to shocks. More precisely, interpret V in the computations above to represent the expectation of value, whereas the realized value Y_t could be higher or lower. If a Union becomes also a fiscal union, *i.e.* coordinating on a unique fiscal as well as monetary policy, the individual shocks of its various components do not affect the overall value of the Union, so that only the aggregate shocks of the Union can alter its potential *vis a vis* the external countries.¹³ On the other hand, if a Union shares exclusively the currency, but not fiscal policy, then also the individual members' shocks matter,¹⁴ and the integration path itself can be altered. We can illustrate this point through the language of the partition function game analyzed above.

First of all, even in the case with maximum internal commitment among those considered so far, namely the MOOG-IC scenario, there is a commitment not to renegotiate downwards the *surplus* shares in case of new admissions, but there was no specification of commitment in case of shocks to individual countries' values. For example, if *ex ante* the chosen path is π_3^3 but in between the first two periods there is a shock that reduces temporarily the value of country C , nothing prevents S from dropping country C in the second period, even though in expected value the chosen path remains the most desirable one.

At a very abstract level, therefore, the addition of fiscal integration provides just enough additional commitment and insurance to guarantee that only aggre-

¹³ A fiscal union means, for me, nothing short of complete integration of fiscal policy, with a European Minister managing both federal taxes and federal expenditures with a countercyclical function, like for the Federal Government in the US, which manages at least 80 percent of the total fiscal policy activities. For the abstract modeling of the paper a fiscal union step is equivalent to the elimination of all the internal commitment problems. Intuitively, it is hard to obtain commitment to transfers or commitment to certain behavior given transfers when such transfers have to happen between sovereign States, whereas a central Government's function is precisely that of countering local shocks but with a common set of procedures and agreed upon checks and balances.

¹⁴ LUQUE J., MORELLI M. and TAVARES J. (2012) provides a more general argument in favor of fiscal union when volatility of individual country shocks increases, and provides a normative analysis of the necessary changes in political weights in order to achieve the desirable fiscal union.

gate shocks matter, and that therefore no temporary shocks to individual values that still leave cooperation *ex ante* desirable could translate in expulsions or punishments or costly renegotiations.

Note also that the presence of shocks could make even the identification of the Core of the Union, namely S^* , unstable. Only if S^* become effectively one country through fiscal integration they can continue to play the admission game in their role of formateur.

4. - Concluding Remarks

In this paper I have primarily focused on a positive game theoretic analysis of the European integration history. Even though the partition function game analysis is obviously very much in reduced form, I have highlighted the implications of different kinds of commitment and enforcement of side payments assumptions on the rationalizability of different admission paths, hence reaching some conclusions in terms of the most reasonable interpretations of the observed behavior throughout the various phases of European integration.

I have then only sketched the potential enrichments of our framework that could lead to normative conclusions in terms of desirability of fiscal unification, where such normative implications are intimately connected with the introduction of realistic possibilities like the existence of uncertainty on future countries' values. In separate work with Guiso *et al.*, (2012) we focus on the individual citizen's perspective, both in terms of reasons for observing integration of different cultures and for the consequent normative desirability of fiscal unification even from the point of view of such an individual level cultural integration problem.

The message that I hope to eventually convey on the conditions for the desirability and realism of further steps in European integration is, in a nutshell, that further widening steps are unwarranted, whereas important deepening steps are necessary, even for the full appreciation of the benefits of the widening steps made so far. The incremental history of European unification is at risk without further deepening steps in the direction of fiscal integration, as the potential shrinking of the Euro zone considered in 2012 discussions on Greece witness.

The results of this paper are to be seen as a stepping stone towards a more complete understanding of Europe's prospects.

BIBLIOGRAPHY

- ACEMOGLU D. - EGOROV G. - SONIN K., «Dynamics and Stability of Constitutions, Coalitions and Clubs», *NBER Working Paper*, no. 14239, 2008.
- ALESINA A. - ANGELONI I. - ETRO F., «International Unions», *American Economic Review*, no. 95, 2005, pages 602-615.
- AUMANN R. - DRÈZE J., «Cooperative Games with Coalition Structures», *International Journal of Game Theory*, no. 3, 1974, pages 217-237.
- BARBERA S. - MASCHLER M. - SHALEV J., «Voting for Voters: A Model of Electoral Evolution», *Games and Economic Behavior*, no. 37, 2001, pages 40-78.
- BELASSA B., *The Theory of Economic Integration*, Richard D. Irwin, Inc., Homewood, Illinois, 1961.
- BLOCH F., «Sequential Formation of Coalitions with Fixed Payoff Division», *Games and Economic Behavior*, no. 14, 1996, pages 90-123.
- .-, «Noncooperative Models of Coalition Formation in Games with Spillovers», in CARRARO C. - SINISCALCO D. (eds.), *New Directions in the Economic Theory of the Environment*, Cambridge University Press, 1997.
- BORDIGNON M. - BRUSCO S., «Optimal Secession Rules», *European Economic Review*, no. 45, 2001, pages 1811-1834.
- FRANKLIN M., *Britain in the European Community*, The Royal Institute for International Affairs, Council on Foreign Relations Press, New York, 1990.
- GEORGE S., *An Awkward Partner: Britain in the European Community*, Oxford University Press, Second Edition, 1994.
- GENICOT G. - RAY D., «Contracts and Externalities: How Things Fall Apart», *Journal of Economic Theory*, no. 131, 2006, pages 71-100.
- GUISSO L. - HERRERA H. - MORELLI M., *A Cultural-Based Theory of Fiscal Union*, mimeo, 2012.
- KOWALCZYK C. - SJÖSTRÖM T., «Bringing GATT into the Core», *Economica*, no. 61, 1994, pages 301-317.
- LUCAS W., *On Solutions to n-Person Games in Partition Function Form*, Ph.D. Dissertation, University of Michigan, Ann Arbor, 1963.
- LUCAS W. - THRALL R., «n-Person Games in Partition Function Form», *Naval Research Logistic Quarterly*, no. 10, 1963, pages 281-298.
- LUQUE J. - MORELLI M. - TAVARES J., *A Volatility-based Theory of Fiscal Union Formation*, mimeo, Columbia University, 2012.
- MAGGI G. - MORELLI M., «Self-Enforcing Voting in International Organizations», *American Economic Review*, no. 96, 2006, pages 1137-1158.
- MORELLI M. - PENELLE P., «Economic Integration as a Partition Function Game», *CORE Discussion Paper*, 1997.

- MORELLI M. - YANG H - YE L., «Competitive Nonlinear Taxation and Constitutional Choice», *American Economic Journal*, no. 4, 2012, pages 142-175.
- RAY D. - VOHRA R., «A Theory of Endogenous Coalition Structures», *Games and Economic Behavior*, no. 26, 1997, pages 286-336.
- .-.-. «Equilibrium Binding Agreements», *Journal of Economic Theory*, no. 73, 1997, pages 30-78.
- ROBERTS K., *Dynamic Voting in Clubs*, London School of Economics, mimeo, 1999.
- THRALL R., *Generalized Characteristic Functions for n-Person Games*, in Recent Advances in Game Theory, Princeton University Conferences, 1962, pages 157-160.
- YI S.-S. - SHIN H., «Endogenous Formation of Research Coalitions with Spillovers», *International Journal of Industrial Organization*, no. 18, 2000, pages 229-256.
- YI S.-S., «Stable Coalition Structures with Externalities», *Games and Economic Behavior*, no. 20, 1997, pages 201-237.

INVITED POLICY PAPER

Freedom and Redistribution

Sebastiano Bavetta*

Università di Palermo e CPNSS
London School of Economics

Pietro Navarra*

Università di Messina e CPNSS,
London School of Economics

We study the determinants of people's attitudes toward income inequality and their economic consequences. We argue that attitudes toward inequality depend on the extent of freedom of choice and control over life outcomes an individual enjoys. We construct a two-stage empirical model where people choose the level of income transfers first and their optimal level of effort, then. We find that the higher the extent of an individual's free choice and control over life outcomes, the greater the probability that he supports larger income differences as incentives for individual effort. This relationship bears consequences on the individual's willingness to work.

[JEL Classification: D31; E62; H2; P16].

Keywords: freedom autonomy; inequality; taxation; redistribution; political economy.

* <sebastiano.bavetta@unipa.it>, Ordinario di Economia Politica, Dipartimento di Scienze Economiche, Aziendali e Finanziarie - Università degli studi di Palermo.

<navarrap@unime.it>, Ordinario di Economia Pubblica, Dipartimento di Scienze Economiche, Aziendali e Metodi Quantitativi - Università di Messina.

1. - Introduction

A branch of the literature in economics argues that people's beliefs about the causes of wealth and poverty shape citizens' preferences for redistribution. Within this line of research a crucial role is played by the relationship between the perceived fairness of the division of wealth and the legitimization of the income dynamics that brought it about. Some sources of inequality are considered legitimate and others unjust on the basis of the perceived fairness in the process through which social mobility occurs. The smaller the perception that citizens face divergent probabilities of improving their economic conditions, the lesser their preferences for flattening income differences through redistributive schemes (Banbou and Tirole, 2006; Piketty, 1995).

Several scholars argue that individuals consider income inequality fair if the *pre-tax* distribution of income is perceived to be caused by factors under their volitional control such as commitment and effort. On the contrary, they consider the *pre-tax* distribution of income unfair if it is perceived as caused by circumstances beyond individual control such as luck and/or privilege (Alesina and Glaeser, 2004; Alesina and Angeletos, 2005). The individual's control over the determinants of income distribution, either through the working of a meritocratic society or through the functioning of an extensive welfare state seems, therefore, to inspire fairness considerations about inequality.

In this context, however, one important question still remains unanswered: when are individuals in a position of voluntarily affecting and, therefore, controlling the *pre-tax* distribution of income? In this paper we argue that the development of a person's autonomy is closely connected with his ability of making choices that express volitional control over the way her life turns out. The fuller the exercise of a person's autonomous behaviour, the more the individual is in a position of voluntarily affecting the *pre-tax* distribution of income and the lesser her support for redistribution. Differently, in societies where individuals are not autonomous and do not voluntarily determine the source of their incomes, state intervention via redistribution schemes is instrumental in order to guarantee social justice.

The concept of autonomy central to this paper is borrowed from the freedom of choice literature and specifically refers to the notion of autonomy freedom proposed by Bavetta and Guala (2003). According to this notion, a person is autonomous to the extent that she has options to choose from and she exercises control over the outcomes of her choices. People who enjoy autonomy have, to a large extent, control over their achievements so, what they perceive as fair or

unfair is likely to depend upon the degree of autonomy freedom they enjoy. Being in control of their own outcomes, people with high levels of autonomy are likely to perceive the game in life as a fair process where individual skills and abilities dominate luck and privilege. On the contrary, people with low levels of autonomy are likely to perceive economic and social outcomes as largely dependent upon elements outside their control. *Ceteris paribus*, the former group of individuals are less likely to favor the use of redistribution as a corrective tool for the unfairness in life.

We use individual level data from the World Value Survey project to assess empirically the validity of our theory. Our data set contains information about individual attitudes toward redistribution and individual perceptions about the extent of their autonomy freedom other than a variety of demographic characteristics as well as socioeconomic indicators assembled at country level. We show that the higher the extent of autonomy perceived by an individual, the higher the probability that she supports the view that larger income differences are needed as incentives for individual effort. Conversely, the lower the extent of autonomy perceived by an individual, the higher the probability that she supports the view that incomes should be made more equal. Further, we demonstrate that the relationship between an individual's level of autonomy and her attitudes toward income redistribution has important consequences on her perceptions concerning the returns from effort and her willingness to work.

The paper is organized as follows. In Section 2 we review the literature and lay down the theoretical hypotheses. In Section 3 we outline the methodology for the empirical analysis and describe the data used in our investigation. In Section 4 we carry out the estimation, show the econometric results and comment on our findings. In Section 5 we draw some concluding remarks.

2. - Theory and Hypotheses

2.1 *The Related Literature*

Redistributive policies differ substantially across countries (Alsina, Glaeser and Sacerdote, 2001). Four sets of explanations have been offered for this observation in the literature. The first body of literature explains redistributive policies postulating that the costs and benefits of redistribution are evaluated differently in different countries. The second uses historical differences in political institutions across countries as the main determining factor. The third suggests that social groups value

income inequality as a signaling mechanism so that observed redistributive policies reflect the intensity of the desire for such signaling. The fourth, which is more related to this paper, concerns with behavioral motives that explain the relationship between individual preferences for redistribution and welfare spending.

Even in the context of selfish utility maximization, income inequality can be argued to have both positive and negative effects on an individual's incentive to put forth effort. The positive effects typically appear when inequality is perceived as a measure of differential rewards for differential effort. On the contrary, the negative effects often appear when income inequality is perceived to emanate from privilege or means deemed to be unjust (Benabou and Tirole, 2006). When we introduce other-regarding terms into an individual's preferences (altruism, reciprocity, etc.), inequality produces additional negative effects on well-being. Both the perceptions of inequality and the extent of other-regarding terms in preferences can differ across countries, leading to differing redistributive policies.

The second class of explanations is based on political institutions. Comparing the European and the American welfare systems, Alesina and Glaeser (2004) argue that differences with regard to redistribution can be traced back to institutions such as majoritarianism, a federal structure of government and checks and balances – and to the ideological premises that guided their historical evolution. Further, Europe is ethnically less fragmented than the US and geographically denser. These characteristics have substantially influenced the gulf in ideological structure and redistributive policy between the two shores of the Atlantic.

The third class of explanations argues that economic inequality has an informational value for socially motivated decisions (Corneo and Grüner, 2000; 2002). This is because it allows for a greater degree of separation across different social groups, *i.e.*, it increases the diversity of consumption patterns across social groups. In such models, consumption patterns are seen as signals of group identity and individuals use them to make social matching decisions (*e.g.*, whom to marry). Individuals with a high matching value, profit from such information since it allows them to select the persons with whom to mix in society. In contrast, individuals with a low matching value suffer a cost from such information because they are more easily ostracized from the most attractive social groups. Higher redistributive taxation that evens the structure of consumption reduces the amount of information available in society. In turn, this alters the extent to which individuals of different wealth classes make successful matching decisions. Therefore, as social structure varies across countries, the voting (or political) attitudes towards redistribution may be affected.

The fourth line of research stresses the role of individual's perceptions about fairness in the process of social mobility. One of the earlier paper in this tradition is given by Piketty (1995). He argues that individuals support redistribution not only to improve their own financial situations, but owing to a sense of justice which entails that citizens should be compensated through redistributive taxation. However, this redistribution should take place only to the benefit of those citizens whose low incomes are due to bad luck, rather than to those whose low incomes are due to lack of effort. Alesina and Angeletos (2005) propose a theory that links people's attitudes toward redistribution to their view regarding the causes of wealth and poverty and the extent to which they believe themselves responsible for their own fate. Their model lead to the co-existence of low- and high- redistribution societies. In a recent and influential book Alesina and Glaeser (2004) bring this perspective into the data and show that the remarkably differences in direct and indirect redistribution between US and Europe can be significantly explained by different individual beliefs concerning the fairness of the income dynamics.

Beliefs regarding the social and moral legitimacy of holding wealth may have an important impact on the amount of wealth levelling that a society undertakes (Putterman, 1996). These sociotropic preferences are based on the concept of "moral reciprocity", *i.e.*, people are willing to transfer income to others conditionally on the guarantee that the latter are not free-riding on the system. The idea that the demand for redistribution cannot be exclusively attributed to selfish motives, but rather significantly accompanied by preferences concerning other's people income and opportunities has been empirically validated in the economic literature. Fong (2001 and 2006) shows that the dynamic of income distribution, the equality of opportunities as opposed to equality of results seem to enter in the individual's utility function.

2.2 *The Theoretical Hypotheses*

The notion of freedom that we use in this paper replicates the concept of freedom as autonomy proposed in the freedom of choice literature (Bavetta and Guala, 2003; Bavetta and Peragine, 2006): a person is free if she faces options to choose from and if she retains control over the option that she has chosen. In this perspective, freedom is based on the existence of two elements: the availability of options to choose from and the individual control over the option that has been chosen. If options are not available to the chooser, she cannot make her

choice. Once options are available, the individual can exercise her freedom only if the choice that she makes is her own. Therefore, if an individual faces options to choose from, her freedom declines if the choice that she makes is determined by circumstances beyond her control (Bavetta and Navarra, 2012).

In this section we formulate a theoretical claim whose validity will be tested empirically in the second part of the paper. We argue that the extent of freedom enjoyed by an individual affects her preferences for redistribution and her willingness to work. More precisely, we hypothesize that the greater an individual's freedom of choice and control over life outcomes, the less she favors redistribution and the higher she is willing to work.

Let us consider the case of a young musician, Robin, who faces a set containing two career alternatives: trust his talent, join a rock band and touring across the country or following the suggestions of his parents, both of them playing at the London Symphony Orchestra, and enroll himself in a school of music. Let us assume that in the social environment in which Robin lives is largely shared the view that success in life depends on effort and merit rather than on family connections. If Robin is not an autonomous person, he may be affected by his friends' preferences and decide to study music in the nearby university. This would imply that, since the choice he makes crucially depends on the social environment in which he lives, changing circle of friends he will most likely change his choices regarding whether to tour the country with the rock band or to enroll himself in a school of music. Further, since the underlying deliberative process that brings Robin to attend music classes at the university is not rooted in his moral and personal qualities, he will not be strongly committed to the choice that he makes. On the contrary, let us suppose that in the same decision situation, Robin carefully evaluates the *pros* and *cons* of the two above alternatives and that, eventually, he decides to join the rock band and carry out a tour of concerts across the country. Having taken a truly autonomous decision, Robin is now likely to be more committed to his final decision because that is the way he wants to determine the course of his life.

From this simple example clearly emerge three important aspects of the decision process that leads Robin to make his choices. First, Robin faces a choice situation in which he has options to choose from. Without options no choice can be ever taken. Therefore, the availability of options is a necessary element to exercise a person's freedom to choose. Second, choices that result from an autonomous deliberation process provide stronger commitment over the way Robin steered his professional life. Third, a thorough deliberative process implies that

Robin is prepared to undertake a responsible and healthy process of outcome revision in the unfortunate and yet possible case in which he could not achieve what he prefers (in this case, a successful tour with the rock band).

Take again the choice situation illustrated above. If Robin is heteronomous, he cannot make sense of why he decided to join the rock band and play music in live concerts across the country. On the contrary, let us consider the case in which Robin carefully examined the available opportunities and eventually decided to tour the country with the rock band. In the unfortunate circumstance of an unsuccessful experience, he will most likely be able to pin down his faults (lack of effort or determination, for instance) and will also be ready to stick to his preferred course of action were a new opportunity of a tour opens to him.

The decision process described above shows that an autonomous individual who faces options to choose from (*i.e.*, free choice) connects his choices with his view of the good. This implies that an autonomous person is more likely to be committed to what he has chosen and ready to undertake a processes of conscious revisions of his decisions in the face of adversities that might occur in life. Commitment and learning by doing trigger a behavioral process of identification that the decision maker develops between the actual choices he makes and the direction he wants to impress to the course of his life (Weiner, Heckhausen and Meyer, 1972; Rotter, 1954 and 1990). Put it differently, the decision maker exercises control over the way his life turns out through the choices he takes. Therefore, free choice and control over life outcomes are two components that make an individual free.

How does this affect the individual's preferences for redistribution and his willingness to work? To answer this question we hypothesize an individual's decision process structured in two stages. In the first stage the individual chooses whether to support or oppose redistribution. In the second stage, the individual chooses how much effort to carry out in his job. The effort decision is made on the basis of the transfer decision undertaken in the first stage of the decision process.

The Transfer Decision

Let us take again the example discussed above. If Robin makes his choices out of an autonomous decision process he feels responsible over the way his professional life turns out. On the contrary, if Robin is not involved in his career choices since they are determined by say his family suggestions, he will less likely attribute

to his responsibility any future professional outcome, no matter whether it is good or bad. Therefore, greater freedom of choice and control over life outcomes implies greater accountability over the choices an individual takes as well as over the outcomes of these same choices. Thus, individuals who enjoy greater freedom of choice and control over life outcomes are more likely to ascribe experienced successes or failures to themselves.

What we argued so far brings about important consequences on how individuals perceive fairness in the game of life and, therefore, on their attitudes about inequality and redistribution. These consequences can be specified by the following theoretical hypotheses:

H1: *The higher the extent of an individual's freedom of choice and control over life outcomes, the stronger his belief that the pre-fiscal distribution of income is determined by factors under his control such as effort, merit and hard work and the greater his opposition to income redistribution.*

H2: *The lower the extent of an individual's autonomy and control over life outcomes, the stronger his belief that the pre-fiscal distribution of income is determined by factors beyond his control such as connections, privilege and luck and the stronger his support for income redistribution.*

It is important to note that the relationship between freedom and preferences for redistribution we describe in this paper casts light upon the concept of fairness individuals rely upon in shaping their attitudes toward inequality. Consistently with Frey, Benz and Stutzer (2004), we argue that social mobility may be interpreted in procedural terms: if people believe that society offers equal opportunities of actual income mobility, they may be less concerned with inequality because they see social processes as fair. It follows that, if individuals perceive themselves as autonomously determining their income dynamics, they might feel that the mobility process is fair as society offers equal opportunities for all. In contrast, those who perceive that their income dynamic is not autonomously determined might see social mobility as a biased process in which opportunities are open to some but not to all.

The Effort Decision

Let us take again Robin's choice example. As argued above, if Robin enjoys free choice and control over life outcomes, he is more likely to feel himself re-

sponsible for the successes and failures that he experiences in his professional life and opposes income transfers. What are the consequences regarding Robin's choices about how much commitment and work effort to elicit in his professional life? As stated above, if he believes that hard work, rather than luck and connections brings a better living, his support for redistribution is low. This implies that he is not likely to feel income transfers as a safety net in case of an unsuccessful professional career. Therefore, he is more willing to unfold high effort and commitment in his work activities. On the contrary, if Robin does not enjoy free choice and control over life outcomes, he does not feel himself accountable for successes and failures in his professional activities and is more likely to support income transfers. This leads Robin to believe that hard work and commitment are not rewarded in life and his support to high income transfers represents an insurance against adverse work experiences. In these circumstances, Robin is less willing to unfold high commitment work effort in his job.

On the basis of the argument described above, we can derive the two following testable theoretical hypotheses:

H3: *The higher the level of an individual's freedom of choice and control over life outcomes, the lower his preferences for income transfers and the higher the effort he unfolds in the income generation process.*

H4: *The lower the level of an individual's freedom of choice and control over life outcomes, the higher his preferences for income transfers and the lower the effort he unfolds in the income generation process.*

Before turning to the empirical part of our study it is important to stress that the role exercised by freedom on the individual preferences for redistribution sheds further light upon the economic analysis of redistributive policies. It has been suggested in the literature that preferences for redistribution are affected by the racial composition of a society: people do not like to give their money to individuals who do not have their skin color so, where ethnic fragmentation prevails, redistribution is lower (Alesina, Glaeser and Sacerdote, 2001; Alesina and Glaeser, 2004). But individual freedom allows a more fine-grained analysis: rather than looking at race or ethnic characteristics, it enables us to focus on whether people retain control over their lives, independently of any other distinctive feature. Preferences for redistributive policies would therefore depend on individual rather than group characteristics. Recognition of such *intra*-group preference heterogeneity is likely to be the basis for more efficient policy design (Bavetta and Navarra, 2012).

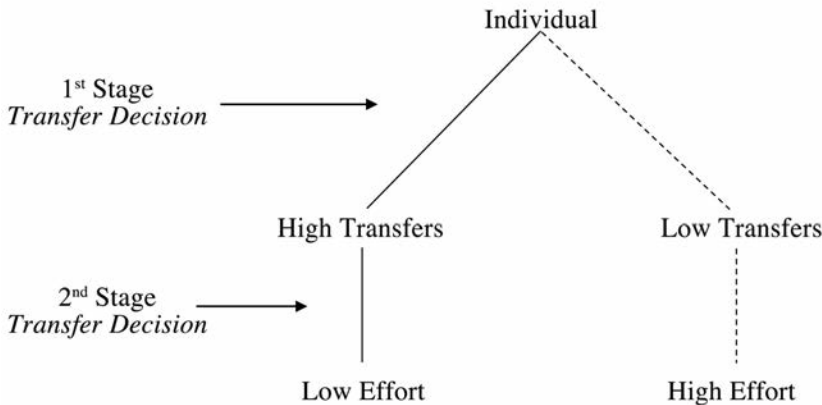
3. - The Empirical Analysis

3.1 *The Methodology*

In Graphs 1A and 1B we show the two-stage decision tree that depicts the individual's sequential decision process. In both graphs the first stage describes the individual's transfers decision while the second stage refers to the same individual's effort decision. However, while in Graph 1A we focus only on those individuals who choose high transfers in the first stage and low effort in the second (see the thick lines in the graph), in Graph 1B we concentrate only on those individuals who choose low transfers and high effort in the first and the second stage, respectively (see the thick lines in the graph). In the empirical part of this paper our objective is to estimate the probability that free choice and control over life outcomes affects the decision process as described in Graphs 1A and 1B in the manner indicated by the theoretical hypotheses described in Section 2.

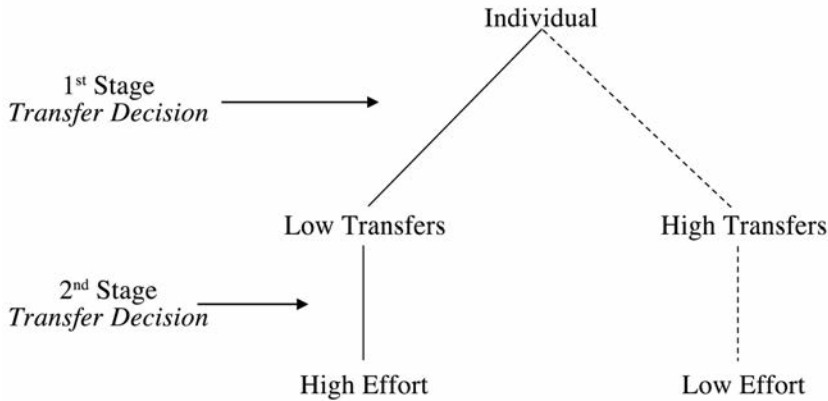
GRAPH 1A

NESTED STRUCTURE OF THE INDIVIDUAL'S DECISION PROCESS (MODEL 1)



GRAPH 1B

NESTED STRUCTURE OF THE INDIVIDUAL'S DECISION PROCESS (MODEL 2)



We use two different econometric procedures to estimate the two models depicted in Graphs 1A and 1B: the conditional logit (CL) and the nested logit model (NL). These two procedures are quite similar, with the CL being a special case of a NL. In the first stage of the decision process the individual chooses a transfer level $t \in n$, where n is the number of available transfer alternatives. Similarly, in the second stage, she chooses the level of effort $e \in m$, where m is the available number of effort alternatives. Let us now define two vectors of variables, respectively x_{te} and z_t , which are specific to the transfers and effort (t, e) and only transfers (t) categories, respectively. The nested logit estimator allows us to calculate the probability that an individual chooses a precise combination of transfer (t) and effort (e) on the basis of what stated in the theoretical hypotheses. Therefore, in Graph 1A (1B) we estimate the probability that an individual chooses the combination of high (low) transfers and low (high) effort on the basis of the effect carried out on her decision process by the extent of free choice and control over life outcomes she enjoys.

More specifically, in our empirical analysis we estimate the following probability:

$$(1) \quad Pr_{te} = Pr_{e|t} \times Pr_t$$

Note that the first term in equation (1), which is the conditional probability, is calculated using the vector of regressors x_{te} only, according to the following expression:

$$(2) \quad Pr_{te} = \exp(x_{te} \beta) / \sum_n \exp(x_{tn} \beta)$$

where β is a vector of parameters.

The second term in equation (1) is calculated by using the vector of regressors z_t as well as the inclusive values of category (e) which links the choice operated by the individual at the second stage to the choice she took in the first stage of the decision process. The inclusive values of category (t, e) are defined as follows:

$$(3) \quad I_t = I_n [\sum_n \exp(x_{tn} \beta)]$$

which can be used to obtain the following probability for the first stage of the decision process:

$$(4) \quad Pr_t = \exp(z_t \gamma + \eta_t I_t) / \sum_m \exp(z_{tm} \gamma + \eta_{tm} I_{tm})$$

It is interesting to note that if the inclusive parameters are all equal to unity the equation (1) reduces to the probability of the conditional logit (CL):

$$(5) \quad Pr_{te} = U_{te} / \sum_n \sum_m \exp U_{te}$$

with $U_{te} = X_{te} \beta + z_t \gamma$.

Clearly, the inclusive values indicate whether a choice is nested into another. Therefore, they are crucial for checking whether the NL procedure fits the estimation of our model better than CL. We check whether the NL procedure is preferable to the CL procedure by running a test for the independence of irrelevant alternatives (Hausman and McFadden, 1984). This test is based on the idea that if a subset of the choice set is irrelevant with respect to the other alternative, then its inclusion among the regressors does not lead to inconsistency in the estimation of the parameters of the model. The NL model (1) is fitted using a full-information-maximum-likelihood estimation.

3.2 *The Data*

We use data from the World Value Survey (WVS) database produced by the Institute for Social Research of the University of Michigan. This collection of data provides microdata obtained from face-to-face interviews carried out to representative samples of the population across more than 60 independent countries around the world. This empirical source is designed to enable a cross-national comparison of values and norms and to monitor changes in individual attitudes across the globe. On the basis of the availability of data required by our theory, we selected 45 countries over a time span that goes between 1981 to 2004.

The Dependent Variable

The individual decision process is structured in two stages. We then have two dependent choice variables in our study. The first is the individual's transfer decision which reveals his preferences for redistribution. We proxy these preferences by answers to the following WVS question (E035 - Income inequality):

How would you place your views on this scale? 1 means that you agree completely with the statement that we need large income differences as incentives; 10 means that you completely agree with the statement that incomes should be made more equal; if your views fall somewhere in between, you can choose any number in between.

From the construction of the question, each individual's taste for income transfer is ordered in a descending fashion: high values indicate high preferences for transfers and *vice versa*.

A more careful look at the question of the WVS may lead someone to argue that, while the statement "income should be made more equal" clearly reflects a dislike for redistribution, the expression "we need larger income differences as incentives" might not always lead to the individual's support for lower transfers. This observation, if correct, would clearly compromise the validity of the question to proxy the individual's preferences for redistribution. However, it is clear that the two statements above do not have to be considered in isolation, but rather as different components of the same question designed to pinpoint divergent tastes of the respondents on the same issue. Since the denomination of the WVS question indicates that the issue at stake is income inequality, it is apparent that the two contrasting views whose survey's respondents are asked to pay attention to are more equal incomes and larger income differences. We believe that these two

opposing preferences for the distribution of income reasonably lead to two opposing tastes for income redistribution. Therefore, we argue for the use of the WVS question E035 - Income inequality, as an appropriate variable to measure the individual tastes for income transfer in our empirical analysis.

The second choice variable concerns with the individual's effort decision and, therefore, indicates the amount of effort that she is willing to carry out in her work activities. We proxy the individual's effort decision with a four point scale variable whose values are ordered in descending fashion as follows: high values indicate that work is not important and low values indicate that work is important. Therefore, in terms of our model high values of the variable show high preferences for effort and *vice versa*.

The Independent Variables

The theoretical hypotheses developed in previous section of this study indicate that both the individual's chosen levels of transfers and effort are affected by the individual's level of freedom of choice and control over life outcomes. The WVS offers a variable that we consider appropriate to measure the extent of freedom of choice and control over life outcomes enjoyed by the individuals. This measure is given by the individuals' answers to the following question (A173 - How much freedom of choice and control):

Some people feel they have completely free choice and control over their lives, while other people feel that what they do has no real effect on what happens to them. Please use this ten-point scale in which 1 means none at all, and 10 means a great deal to indicate how much freedom of choice and control you feel you have over the way your life turns out.

In terms of our measure of freedom and control over life outcomes the variable is coded in ascending order: high values indicate a high degree of freedom and control and *vice versa*.

Several control variables are included in our dataset. They range from socio-demographic variables to individuals' opinions about politics, religion and society. As far as the socio-economic variables are concerned, to capture the effect of income on the individuals' preference for redistribution we consider the level of self-reported income. Respondents were asked to express the level of their income on a ten-point scale with low and high values indicating low and high levels of income, respectively. In line with similar variables used in other studies (Alesina

and Glaeser, 2004), the individual's perception of income mobility is measured by a binary dummy which takes the value of 1 if the respondent believes that there is little chance for poor people to escape out of poverty and the value of 0 if the respondent believes that people do have a chance to escape from poverty. If the individual believes that people do have chances to work their way out of poverty, it is reasonable to hypothesize that she thinks that in the community where she lives there is income mobility. The opposite applies in the case she believes that the poor have little chances escape from poverty. Another binary dummy variable is used to indicate the gender of respondents. It takes the value 1 if respondent is male and 2 if female. Age is expressed in years. The education level is computed on a eight-point scale in ascending order going from low to high levels of education.

The variables measuring people attitudes towards politics, religion and society are the following. The respondent's political opinion is measured over a ten-point scale whose values 1 and 10 indicate extreme left and extreme right political orientation, respectively. The respondent's religiosity is indicated by her assessment of how religion is important in life. The variable is coded in descending order over a four point range with the two extreme values, 1 and 4, indicating that religion is very important and not at all important, respectively.

In Table 1 we show some summary statistics and the correlation matrix of the variables used in the empirical investigation. In the Data Appendix we report the definition of all the variables together with their respective source.

TABLE 1

SUMMARY STATISTICS AND CORRELATION MATRIX^(a)

Variable	Obs	Mean	SD	Min	Max	1	2	3	4	5	6	7	8
1 Preferences for Transfers	107793	6,010771	2,974672	1	10	1							
2 Preferences for Effort	107793	4,358892	2,898011	1	10	-0.0185*	1						
3 Autonomy Freedom	107793	6,667288	2,416495	1	10	0.0820*	-0.0598*	1					
4 Gender	107670	0,510421	0,4998937	0	2	-0.0390*	0.0424*	-0.0411*	1				
5 Political Orientation	84221	5,53286	2,21545	1	10	0.1322*	-0.0697*	0.0677*	-0.0022	1			
6 Income Mobility	53057	1,588989	0,4920218	1	2	-0.0326*	0.1376*	-0.1380*	0.0502*	-0.1116*	1		
7 Religiosity	99366	4,799952	2,53062	1	8	0.0200*	0.0916*	-0.0282*	-0.0937*	-0.1641*	0.0972*	1	
8 Education	64431	1,948596	0,7274976	1	3	0.1567*	-0.0029	0.0870*	-0.0313*	-0.0177*	-0.0237*	0.0442*	1

^(a)* indicates statistical significance of the correlation coefficient.

4. - Estimation and Results

As pointed out in the empirical methodology section, two econometric approaches are used: CL and NL. However, since the CL specification is rejected by the generalised Hausmann test (results are reported but available upon request from the authors), we present the NL findings only. In Table 2 we show the findings of our estimation. First of all, it is important to highlight that the likelihood *ratio* (LR) test in both Model 1 and Model 2 rejects the null hypothesis that the inclusive values for categories (t , e) are equal to 1. As pointed out in the empirical methodology section, this implies that CL and NL give rise to different probability estimates. Further, note that the results displayed in the table include year as well as country dummies in the estimation.

We start describing the regression results for Model (1) where the nested structure of the two-stage decision process is the one shown in Graph 1*a*. In the top part of the first column we show regression results for the transfers decision (first stage). In the bottom part of the column we display our findings for the effort decision (second stage). The outcome of the empirical analysis allows us to evaluate whether and to what extent an individual's free choice and control over life outcomes (*i.e.*, the level of an individual's autonomy freedom) affects the probability that she chooses the combination high transfers and low effort.

As far as the effect of our primary independent variable is concerned, we can say from the outset that the theory fits very well the data. The extent of an individual's free choice and control over life outcomes determines both transfer and effort decisions as predicted by the theoretical hypotheses. The higher the individual's level of free choice and control over life outcomes (*i.e.*, her level of autonomy freedom), the lower the probability that she chooses high income transfers in the first stage and the higher her optimal level of effort in the second stage. This effect is statistically significant and consistent across the two stages of the model. Thus Hypotheses 1 and 3 are empirically confirmed.

The degree of income mobility as perceived by the individual determines both the transfer and the effort decision. The lower the perceived degree of income mobility, the higher the probability that the individual chooses a high level of income transfers in the first stage and a low level of effort in the second stage. Again the effect of the individual's perception of the degree of income mobility is statistically significant and consistent across the two stages of the model. Regarding the other independent variables and controls, we note that the individual's preferences for high transfers increase if female, the lower her income and the more

left-wing she is politically oriented. Further, we observe that individuals are more likely to unfold lower levels of work effort the older and the less educated they are and the weaker their religious beliefs.

TABLE 2

NESTED LOGIT ESTIMATION ^(a)		
	Model (1) <i>High Transfer and Low Effort</i> TRANSFER DECISION	Model (2) <i>Low Transfer and High Effort</i> TRANSFER DECISION
Autonomy Freedom	-0.024** (-2.39)	0.248*** (2.90)
Gender	0.242*** (5.94)	-1.536*** (-4.03)
Income Mobility	0,058 (1.30)	-0.994*** (-2.66)
Income	-0.017** (-2.08)	0.177*** (6.98)
Political Orientation	-0.019** (-2.01)	-0,024 (-0.85)
Country Dummies	YES	YES
Year Dummies	YES	YES
Weights	YES	YES
	EFFORT DECISION	EFFORT DECISION
Autonomy Freedom	-0,024 (-0.89)	0.026*** (2.73)
Income Mobility	0.210* (1.78)	-0.087** (-2.18)
Gender	-0,156 (1.40)	-0.193*** (-4.73)
Age	0.042*** (10.86)	0.004*** (5.65)
Education	-0.208*** (-2.83)	-0.009 (-1.00)
Religiosity	0.081*** (3.47)	0.009*** (3.20)
Country Dummies	YES	YES
Year Dummies	YES	YES
Weights	YES	YES
LR Test	2.72*	91.97***

^(a) *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Let us now move on to the estimation of Model (2). Regression results for transfers and effort decisions are shown in the second column of Table 2. In this model we empirically evaluate whether and to what extent, in the two-stage decision process described Graph 1*b*, an individual's level of freedom of choice and

control over life outcomes affects the probability that she chooses the combination low transfers and high effort. As before, in the top part of the column we show the estimates concerning the determinants of the transfer decision (first stage). In the bottom part of the column we show the regression results related to the determinants of the effort decision. Again, country and year dummies are included in the estimation.

The effect of the primary independent variable is again very much in line with the predictions of our theory. This implies that the estimates shown in the first column of Table 2 are specular to those appearing in the second column of the same table. An individual's level of free choice and control over life outcomes determines both transfer and effort decisions as anticipated by our theoretical hypotheses. The higher the individual's extent of free choice and control over life outcomes, the higher the probability that she chooses low income transfers in the first stage and the higher her optimal level of effort in the second stage. This effect is statistically significant and consistent across the two stages of the model. Therefore, hypotheses 2 and 4 are empirically confirmed.

The degree of income mobility as perceived by the individual determines both the transfer and effort decisions. The higher the perceived degree of income mobility, the higher the probability that the individual chooses a low level of transfer in the first stage and a high level of effort in the second stage. Again the effect of the individual's perception of the degree of income mobility is statistically significant and consistent across the two stages of the model. Regarding the other independent variables and controls, we note that the individual's preferences for low transfers increase if male and the higher his income. Further, we observe that individuals are more likely to carry out higher levels of work effort if male, the younger and more educated they are and the stronger their religious beliefs.

To sum up, the main result of the empirical analysis shows that individuals enjoying high levels of free choice and control over life outcomes (*i.e.*, high autonomy freedom) do not only ask for lower levels of income transfers, but are also willing to work and produce more. Differently, low autonomy free individuals, ask for higher transfers and are likely to work less.

5. - Concluding Remarks

In this study we argued that the more autonomous an individual, the higher her perception of being in control over her own choices and the greater she feels

responsible of the way her life turns out. We empirically demonstrated that the greater the exercise of a person's autonomous behaviour the more she believes that the working of the process of social mobility is just and, therefore, the lesser her demand for redistribution.

The results of the empirical analysis indicate that individuals enjoying high levels of free choice and control over life outcomes ask for lower levels of income transfers and are more likely to work and produce more. On the other hand, individuals characterized by low levels of free choice and control over life outcomes ask for higher transfers and are likely to work less. While in the first case individuals are more likely to generate a productive society, in the second case they are more likely to generate a rent seeking society. It is important to note that the support of low transfers by an autonomous individual emerges out of a sense of procedural fairness whereby she believes that her economic conditions are deserved, no matter whether she is rich or poor.

The implications of our study are important with regard to the classical problem of the trade-off between freedom and income inequality in liberal democracies. The political debate over income inequality has been traditionally characterized by two opposing views. On the one hand, liberals consider economic inequality unjust and socially destructive. On the other, conservatives generally feel that riches are the best way to reward those who contribute the most to prosperity and/or that a generous welfare state encourages idleness and folly amongst the poor. These two apparently divergent views may be reconciled in the light of the results obtained in this study. Income inequality can be in fact considered as a fair outcome according to the extent of autonomy freedom people enjoy.

BIBLIOGRAPHY

- ALESINA A. - ANGELETOS G.M., «Fairness and Redistribution», *American Economic Review*, vol. 85, 2005, pages 960-980.
- ALESINA A. - GLAESER E.L., *Fighting Poverty in the US and Europe: A World of Difference*, Oxford University Press, Oxford, 2004.
- ALESINA A. - GLAESER E.L. - SACERDOTE B., «Why Doesn't the US Have a European Style Welfare State», *Brookings Papers on Economic Activity*, Fall, 2001, pages 187-278.
- BAVETTA S. - GUALA F., «Autonomy Freedom and Deliberation», *Journal of Theoretical Politics*, vol. 15, 2003, pages 423-443.
- BAVETTA S. - NAVARRA P., *The Economics of Freedom: Theory, Measurement and Policy Implications*, New York, Cambridge University Press, 2012.
- BAVETTA S. - PERAGINE V., «Measuring Autonomy Freedom», *Social Choice and Welfare*, vol. 26, 2006, pages 31-45.
- BÉNABOU R. - TIROLE J., «Belief in a Just World and Redistributive Politics», *Quarterly Journal of Economics*, vol. 121, 2006, pages 699-746.
- CORNEO G. - GRÜNER H.P., «Social Limits to Redistribution», *American Economic Review*, vol. 90, 2000, pages 149-157.
- FONG C., «Social Preferences, Self-Interest, and the Demand for Redistribution», *Journal of Public Economics*, vol. 82, 2001, pages 225-246.
- .-, *Prospective Mobility, Fairness and the Demand for Redistribution*, mimeo, SDS, Carnegie Mellon University, 2006.
- FREY B.S. - BENZ M. - STUTZER A., «Introducing Procedural Utility: Not Only What, but Also How Matters», *Journal of Institutional and Theoretical Economics*, vol. 160, 2004, pages 377-401.
- PIKETTY T., «Social Mobility and Redistributive Politics», *Quarterly Journal of Economics*, vol. 60, 1995, pages 551-583.
- PUTTERMAN L., «Why Have the Rabble not Redistributed the Wealth? On the Stability of Democracy and Unequal Property», in ROEMER J. (ed.), *Property Relations, Incentives and Welfare*, London, MacMillan, 1996.
- ROTTER J.B., *Social Learning and Clinical Psychology*, Prentice Hall, New York, 1954.
- .-, «Internal versus External Locus of Control of Reinforcement: A Case History of a Variable», *American Psychologist*, vol. 45, 1990, pages 489-493.
- WEINER B. - HECKHAUSEN H. - MEYER W.U., «Causal Ascriptions and Achievement Behaviour: Conceptual Analysis of Effort and Reanalysis of Locus of Control», *Journal of Personality and Social Psychology*, vol. 21, 1972, pages 239-248.

Regolazione decentrata e servizio concentrato: le ferrovie regionali viaggiano su un binario stretto?

Chiara Bentivogli*
Banca d'Italia, Bologna

Eugenia Panicara*
Banca d'Italia, Bologna

This paper focuses on the recent evolution of structure and regulation of local railways transportation in Italy. The regulation changes aimed at improving the efficiency and the quality of service of regional trains produced very modest results. Regions have not exploited the tools available, encouraging the dominance of Trenitalia (the most important provider of transportation services under public control) in local markets too. Service improvement requires more “neutral rules” and greater ability of Regions in selecting the operator and in defining the content of service contracts. The existing network is sufficient, given its present use, also in international comparison.

[JEL Classification: H54; H77; L92; R41].

Keywords: trasporto ferroviario locale; regolamentazione dei trasporti; federalismo.

* <chiara.bentivogli@bancaditalia.it>; <eugenia.panicara@bancaditalia.it>; Divisione Analisi e Ricerca Economica Territoriale. Desideriamo ringraziare le Regioni e le imprese ferroviarie per la preziosa collaborazione, Antonio Manuguerra e Monica Roselli del MIT, Eleonora Pieralice di Isfort e Pietro Marra per i dati messi a disposizione, Gianluigi Annibaldi e Mirko Madau per la collaborazione nell'elaborazione dei dati, Fabrizio Balassone e Piero Casadio della Banca d'Italia, Romeo Danielis dell'Università di Trieste, Alfredo Macchiati, Andrea Pezzoli dell'AGCM, Giorgio Stagni della Regione Lombardia per gli utili commenti. Le opinioni contenute nel lavoro sono da attribuirsi interamente alle autrici e non impegnano in alcun modo l'istituzione di appartenenza.

1. - Introduzione

Negli ultimi vent'anni nella UE gli orientamenti delle politiche dei trasporti si sono diretti verso obiettivi di maggiore efficienza, ecosostenibilità e completamento del mercato unico europeo (aumento della concorrenza fra le imprese di trasporto e interoperabilità delle reti). Al trasporto ferroviario regionale è stato attribuito il compito di agevolare la mobilità dei pendolari in maniera più efficiente che in passato e con un minore impatto ambientale in termini di inquinamento e congestione rispetto a quello su gomma.

In Italia questo mutamento di prospettiva è stato introdotto nella normativa e nella politica economica parallelamente al decentramento alle Regioni delle competenze nei servizi pubblici locali, replicandone spesso lentezze, esitazioni e ripensamenti. Questo lavoro fa il punto sullo stato dell'infrastruttura e del servizio ferroviario italiano, con particolare riferimento a quello destinato ai viaggiatori regionali. Si valuta la trasformazione dell'assetto complessivo del comparto in relazione sia all'evoluzione normativa sia alla sua maggiore rispondenza alle finalità pubbliche dichiarate. Si analizzano i cambiamenti nella rete e nell'offerta di trasporto rispetto alle esigenze, anche qualitative, espresse dalla domanda dei pendolari. Si discutono inoltre i principali aspetti critici del nuovo assetto, le caratteristiche dei possibili assetti futuri e le indicazioni di politica economica che ne discendono.

2. - La rete e il trasporto ferroviario di passeggeri: aspetti teorici

L'infrastruttura ferroviaria, costituita principalmente dai binari e dalle stazioni, è l'*input* fondamentale del servizio di trasporto ferroviario. Rispetto ad altre infrastrutture di trasporto, la ferrovia è caratterizzata da un basso consumo di territorio, da costi fissi elevati e da tempi lunghi di ammortamento degli investimenti. Il servizio offerto su rotaia è molto meno inquinante di quello su gomma: viaggiare su rotaia produce da 3 a 10 volte meno CO₂ rispetto al trasporto aereo o su strada; il trasporto ferroviario assorbe meno del 3% dell'energia consumata dalla totalità dei trasporti, con una quota di mercato del 6% (passeggeri; UIC, 2008).

La lunghezza della rete non misura pienamente la dotazione di infrastrutture ferroviarie di un territorio; è necessario tenere anche conto del tipo di tecnologia utilizzato (binario unico o doppio, rete elettrificata o meno, strumenti di segnalazione, tecnologia di regolazione degli accessi alle stazioni, che influenzano signifi-

cativamente l'efficienza del servizio e la possibilità di ampliarne l'offerta) e del grado di interoperabilità della rete (ossia del grado di omogeneità tecnica e normativa della stessa), che può segmentare il mercato e ridurre la qualità del servizio.

I principali operatori del comparto sono il proprietario della rete, il gestore della rete (che assegna le tracce orarie e riscuote i relativi pedaggi dai gestori del servizio) e le imprese ferroviarie (che gestiscono il servizio). Accanto al pedaggio, i principali costi del servizio ferroviario sono quelli relativi al materiale rotabile (i treni) e al personale. In alcuni paesi agli operatori sopra indicati si aggiungono anche società proprietarie del materiale rotabile, che lo affittano ai gestori del servizio.

Nei servizi ferroviari viaggiatori, l'*output* è misurato dai passeggeri-km. Il servizio fornito dalle imprese ferroviarie si concretizza tuttavia in una determinata disponibilità di treni-km o di posti-km (cfr. glossario). I costi variabili sono perciò legati più al numero di treni utilizzati, ai posti disponibili, ai km complessivamente percorsi e ai tempi di percorrenza che ai passeggeri trasportati, mentre questi ultimi incidono principalmente sui ricavi.

Il settore è caratterizzato da un'elevata intensità di capitale (soprattutto la rete e il materiale rotabile, con vita economica però molto elevata) a cui si accompagna una limitata sostituibilità di capitale e lavoro. Innovazioni tecnologiche possono però di tanto in tanto determinare cambiamenti rilevanti nel rapporto tra mezzi in circolazione e *input* di lavoro (ad es. nel numero di macchinisti per treno, oppure nel personale dedicato alla vendita di biglietti, grazie alle vendite via internet o presso biglietterie automatiche). Gli elevati investimenti richiesti per la rete e per il materiale rotabile (la cui vita economica può attualmente superare i 30 anni), determinano significative economie di densità per tratta. La ferrovia risulta economicamente sostenibile per intensità di utilizzo piuttosto elevate (in termini di coefficiente di riempimento o *load factor* dei singoli treni e di frequenza d'uso della rete) e quindi in aree densamente popolate e/o in cui l'attività produttiva è concentrata nel territorio (Ponti e Beria, 2007). Infine, alcuni costi sono più sensibili ai tempi di percorrenza (ad esempio il costo del lavoro) che alla distanza percorsa (ad esempio il costo della rete e dei meccanismi di controllo).

La domanda di trasporto ferroviario passeggeri è segmentabile in due gruppi: quella del trasporto locale, all'interno di un sistema locale del lavoro o comunque che soddisfa i pendolari, e quella a lunga percorrenza, in cui il trasporto ferroviario ha come sostituti, oltre al trasporto individuale o collettivo su gomma, anche il trasporto aereo. Soprattutto per il primo segmento vi è una forte dipendenza della domanda da fattori di contesto e di *policy* locale. Fattori quali il tipo di urbanizzazione di una regione (la concentrazione della popolazione lungo la rete, ad esem-

pio) e l'intensità del fenomeno del pendolarismo, nonché l'ampiezza e la coerenza delle misure restrittive al traffico privato (presenza di zone a traffico limitato, targhe alterne, diffusione della tariffazione della sosta e del *road pricing*, ecc.) e la disponibilità di servizi accessori quali i parcheggi di interscambio, influiscono sulle condizioni del traffico, sui tempi complessivi del viaggio, sulla puntualità e più in generale sulla qualità del servizio offerto. Questa, a sua volta, è tra le determinanti della domanda di trasporto ferroviario al pari, se non più, del costo del servizio (Hensher e Brewer, 2001). In particolare, poiché il trasporto ferroviario richiede quasi sempre l'utilizzo di altre infrastrutture per completare lo spostamento richiesto (strada, rete metropolitana, ecc.), i tempi di attesa per lo scambio intermodale possono scoraggiare l'utilizzo del treno per distanze brevi, anche se quest'ultimo ha una velocità commerciale relativamente elevata (Litman, 2007). Le peculiari caratteristiche della domanda di trasporto hanno determinato l'introduzione nell'analisi economica del concetto di costo generalizzato, che include oltre ai costi monetari, anche il costo-opportunità legato al tempo di trasporto di ciascuna modalità.

La domanda presenta notevoli fluttuazioni nel corso della giornata, della settimana e nei diversi mesi dell'anno. Quella pendolare è di tipo radiale (verso e dal centro urbano più importante dell'area) e concentrata nelle ore di punta. L'alta variabilità infragiornaliera della domanda e l'elevato investimento minimo necessario per aumentare l'offerta tendono a determinare capacità in eccesso o affollamento dei passeggeri. Negli anni più recenti la domanda di trasporto è diventata sempre più individualizzata a causa della minore uniformità degli orari di lavoro e del tempo libero, tendendo a spostarsi verso il più flessibile trasporto su auto privata; questa tendenza è stata anche favorita dalla riduzione relativa dei costi di acquisto dell'auto e dalla dislocazione dei centri commerciali in aree extraurbane (a sua volta in parte risultato dei vincoli all'accesso delle auto nei centri urbani; Ponti e Beria, 2007; Ramella, 2010).

Le politiche pubbliche del trasporto ferroviario cercano di incentivare il trasporto ferroviario locale con due principali obiettivi: garantire il diritto alla mobilità dei cittadini e favorire il riequilibrio modale a scapito dell'automobile per il minor inquinamento, congestione, incidentalità e costi di usura delle strade. Il maggior utilizzo dell'auto deriverebbe dalla sua comodità relativa e dal fatto che le esternalità negative generate non si riflettono correttamente nei suoi costi d'uso. Per conseguire questi obiettivi si utilizzano sussidi tariffari in favore del trasporto ferroviario e misure di disincentivo dell'utilizzo dell'auto, quali imposte e tasse sul carburante, vincoli alla circolazione, ecc. Quest'ultimo tipo di misure, assieme

a un miglioramento qualitativo del servizio, risulterebbe più efficace perché potrebbe orientare verso il treno anche i cittadini più abbienti, meno sensibili al costo relativo del viaggio. Il riconoscimento di un diritto alla mobilità si traduce nell'attribuzione al trasporto ferroviario della qualifica di "servizio universale", per il cui svolgimento sono imposti e compensati specifici obblighi di servizio pubblico (OSP). Nonostante questi obiettivi generalmente dichiarati, nella gran parte dei paesi industriali il trasporto su gomma è più sussidiato delle ferrovie¹.

Un'altra importante peculiarità del trasporto ferroviario, che può giustificare politiche pubbliche (in questo caso rivolte all'offerta), sta nel fatto che la rete per molti aspetti risulta un monopolio naturale il cui utilizzo comporta una struttura dei costi (economie di scala e subadditività) tale da suggerire la presenza sul mercato di un solo operatore. Qualora la rete abbia le caratteristiche di un monopolio naturale, la teoria economica suggerisce la separazione proprietaria fra gestore della rete e gestore del servizio e l'affidamento di quest'ultimo a operatori in concorrenza fra loro. La separazione verticale e la regolazione del gestore della rete (criteri di accesso e relativi pedaggi; Campos e Cantos, 1999) avrebbe il vantaggio di garantire in maniera equa e non discriminatoria l'accesso alla rete, stimolando la concorrenza. La regolamentazione è auspicata dalla teoria nel caso della rete, generalmente considerata una *essential facility*, ma non del materiale rotabile, in quanto facilmente riproducibile; tuttavia i costi elevati del materiale rotabile nuovo e i tempi per una sua acquisizione (da 2 a 4 anni; AGCM, 2003a), la pressoché totale assenza di un mercato dell'usato (anche a causa di differenze di standard) e la necessità da parte di un potenziale entrante nella gestione del servizio di possedere il materiale rotabile (o acquisirlo a vario titolo) possono costituire una rilevante barriera all'entrata (Cambini, 2009). Fra le possibili soluzioni vi è l'introduzione sul mercato di *rolling stock companies* (ROSCOs, come è avvenuto

¹ L'Agenzia europea per l'ambiente stimava che nel 2005 nella UE-15 i sussidi ai trasporti ammontassero a circa 270-290 miliardi di euro all'anno, di cui il 45% al trasporto su gomma e il 26 a quello su rotaia (EUROPEAN ENVIRONMENT AGENCY, 2007). Sono presenti forme di sussidio in pressoché tutte le tipologie di trasporto, da quello aereo (ad esempio sotto forma di sussidi alle compagnie *low cost* da parte di enti locali), a quello su gomma (sgravi fiscali e sconti sui pedaggi agli autotrasportatori e soprattutto sussidi per la realizzazione e la manutenzione delle infrastrutture), alle ferrovie, accompagnate talvolta da imposte e tasse specifiche mirate a scoraggiare l'utilizzo di una determinata modalità (ad esempio l'accisa sul gasolio per autotrazione). Non stupisce dunque come le politiche dei trasporti mirate alla sostenibilità ambientale non siano state molto efficaci: il trasporto causa circa un quarto di tutte le emissioni di CO₂; fra il 1990 e il 2005 le emissioni di gas serra da trasporto interno alla UE-15 sono aumentate del 26%; oltre il 90% delle emissioni è dovuto al trasporto su strada (UIC, 2008).

nel Regno Unito), che possono offrire il materiale rotabile in *leasing* alle compagnie ferroviarie.

Se la separazione proprietaria ha vantaggi concorrenziali, perché favorisce una maggiore trasparenza e terzietà del gestore della rete, l'integrazione ha alcuni vantaggi industriali, in termini di possibilità di sfruttare economie di scopo e vantaggi informativi, di maggiori incentivi alla manutenzione/investimento della rete e di riduzione dei costi di transazione (Macchiati *et al.*, 2007; Merkert *et al.*, 2010). Le analisi empiriche non sarebbero tuttavia conclusive né sugli effetti in termini di efficienza legati all'integrazione verticale, né sulla presenza di economie di scopo fra il trasporto merci e quello passeggeri. Merkert *et al.* (2010) mostrano che empiricamente la separazione verticale non determinerebbe significative differenze nell'efficienza tecnica complessiva, mentre i maggiori costi di transazione che ne derivano avrebbero un significativo effetto negativo².

Qualora la presenza di più operatori nella gestione del servizio generi congestione e/o se il regolatore ha scelto di sussidiare anche parzialmente il servizio (per uno qualsiasi dei motivi sopra detti), è preferibile l'introduzione di meccanismi di concorrenza per il mercato (scelta del gestore con meccanismi d'asta; Boitani e Petretto, 2002) e, nel caso di un'impresa che effettui sia servizi concorrenziali sia servizi pubblici, la separazione proprietaria fra le due attività, per evitare comportamenti e strategie volti a falsare il gioco della concorrenza sul primo mercato (AGCM, 2008).

La definizione di meccanismi di concorrenza per il mercato richiede di affrontare numerose scelte di dettaglio (gli assetti istituzionali della relazione tra autorità pubblica e gestori del servizio, il meccanismo d'asta, il numero di tratte da affidare a ciascun gestore, le caratteristiche del contratto di servizio, la definizione del livello delle tariffe e dei sussidi). Vi è in generale consenso sulla necessità di separazione fra l'autorità di indirizzo, programmazione e controllo e i gestori, per ridurre le inefficienze legate sia alla minore coerenza dei vincoli di bilancio qualora vi sia identità fra controllore e gestore, sia più in generale ai fenomeni di cattura del regolatore. La scelta della dimensione del mercato da mettere a gara pone invece alcuni *trade-off* in termini di facilità di monitoraggio, numero di potenziali partecipanti alla gara e rischio di collusione. Un ulteriore strumento, complementare alle gare, è rappresentato dal contratto di servizio (CS) che, in presenza

² L'effetto dei costi di transazione sull'efficienza non è scontato; alti costi di transazione potrebbero accrescere l'efficienza tecnica se determinassero migliori decisioni o se riducessero il fabbisogno di altri *input* (MERKERT R. *et al.*, 2010).

di asimmetrie informative, dovrebbe definire meccanismi di incentivo affinché il gestore offra la quantità e gli standard richiesti del servizio, evitando di attribuire un vantaggio informativo all'*incumbent* anche mediante una durata eccessivamente lunga del contratto. La letteratura evidenzia come siano più incentivanti per il gestore contratti che prevedono un metodo di remunerazione *net cost*, in cui il gestore, pur ricevendo un corrispettivo, si fa carico del rischio industriale (costi e gestione) e di quello commerciale (ricavi), rispetto al metodo *gross cost*, in cui i ricavi sono incassati dall'ente concedente. Se lo stesso gestore ha oneri di servizio su alcune tratte ma non su altre (potenzialmente remunerative), l'ente locale contraente può consentire all'impresa ferroviaria l'utilizzo di sussidi incrociati fra i due tipi di servizio; tuttavia ciò non dovrebbe essere ammesso qualora sulle tratte remunerative il gestore sia in concorrenza con altre imprese.

3. - L'evoluzione della regolamentazione in Italia

L'assetto generale

Gli anni novanta rappresentano uno spartiacque importante per il settore ferroviario in Italia che, soprattutto per effetto del recepimento di diverse direttive europee, è stato interessato da rilevanti cambiamenti regolatori. La diffusa convinzione della necessità di cambiamenti nel settore era legata al declino nell'utilizzo delle ferrovie sia per il trasporto dei passeggeri sia per quello di merci (a fronte di un costante aumento nel trasporto privato su gomma) e alle pesanti perdite accumulate negli anni dalle imprese ferroviarie. La necessità di riforma del settore, in declino anche in molti altri paesi europei, era stata evidenziata dal Libro Bianco sulle ferrovie del 1996³. I principi ispiratori della nuova regolamentazione europea delle ferrovie (la realizzazione del mercato unico, la riduzione dei sussidi pubblici e la promozione della concorrenza) sono stati introdotti gradualmente e con alcune eccezioni, legate soprattutto al riconoscimento della funzione di servizio universale ad alcune tipologie di trasporto ferroviario.

Gli interventi europei nella regolamentazione delle ferrovie sono stati avviati con la Direttiva CE 29 luglio 1991, n. 440, che prescriveva in particolare misure per:

- i) il risanamento delle imprese ferroviarie (attraverso la creazione di aziende separate dallo Stato con criteri di gestione commerciale);
- ii) la liberalizzazione dell'accesso all'infrastruttura ferroviaria per il trasporto ferroviario di merci (con la separazione almeno contabile fra gestore dell'infra-

³ White Paper, *A Strategy for Revitalising the Community's Railways*, COM(96), no. 421, 30 luglio 1996.

struttura e gestore del servizio, il divieto di sussidi incrociati fra le due attività e l'applicazione di principi non discriminatori nell'allocazione delle tracce orarie alle imprese ferroviarie).

La Direttiva escludeva infine dal suo campo di applicazione i trasporti regionali. In Italia il relativo regolamento di attuazione, successivo di 7 anni alla Direttiva (decreto del Presidente della Repubblica 8 luglio 1998, n. 277), prevedeva una separazione solo contabile fra gestione dell'infrastruttura e servizio e liberalizzava il servizio solo per alcuni casi particolari⁴. La riluttanza dell'Italia ad aderire pienamente al nuovo approccio europeo è stata segnalata anche dall'AGCM (1998), che indicava la preferibilità sia di una separazione effettiva sia della terzietà del soggetto che assegna delle tracce anche rispetto al gestore dell'infrastruttura.

I successivi passi della regolamentazione italiana delle ferrovie nazionali sono costituiti principalmente dall'attuazione dei tre pacchetti ferroviari della UE. Il primo, effettivo dal 2004, introduce regole trasparenti e uniformi di accesso alla rete e alcune prime misure per la liberalizzazione del trasporto ferroviario di merci transeuropeo. Il secondo, effettivo dal 2005, prevede la piena apertura del mercato UE del trasporto ferroviario merci dal 2007, poi anticipata al 1° gennaio 2006. Il terzo prevede l'apertura del mercato dei servizi ferroviari dei passeggeri dopo il primo gennaio 2010⁵. In generale, il processo di liberalizzazione definito a livello europeo non ha dato risultati soddisfacenti. In molti paesi la separazione verticale fra rete e gestore *incumbent* è ancora solo formale e il libero accesso di concorrenti

⁴ Per le associazioni internazionali di imprese ferroviarie di cui almeno una avesse la sede principale in Italia per il trasporto internazionale con altri paesi UE in cui sono stabilite le imprese che costituiscono l'associazione e per le imprese ferroviarie con sede nel territorio UE, a condizioni eque, per l'espletamento di servizi internazionali di trasporto merci. Per il trasporto merci è stata introdotta dalla legge 23 dicembre 2000, n. 388 (legge finanziaria per il 2001), la condizione di reciprocità.

⁵ Il primo pacchetto ferroviario (Direttive CE 26 febbraio 2001, nn. 12, 13 e 14) riprende e modifica i contenuti della Direttiva CE 1991/440 e delle Direttive CE 19 giugno 1995, nn. 18 e 19, recepite con il decreto del Presidente della Repubblica 16 marzo 1999, n. 146. È seguita la Direttiva CE 19 marzo 2001, n. 16, che prevede una serie di misure di rimozione degli ostacoli all'interoperabilità delle reti europee (recepita con il D.lgs. 30 settembre 2004, n. 268). Il secondo pacchetto include la Direttiva CE 29 aprile 2004, n. 49 che modifica in parte la Direttiva CE 2001/14, la Direttiva CE 29 aprile 2004, n. 50 (sistema ferroviario transeuropeo), la Direttiva CE 29 aprile 2004, n. 51 (sviluppo delle ferrovie comunitarie) e il Regolamento CE 29 aprile 2004, n. 881 che istituisce un'agenzia ferroviaria europea. La liberalizzazione del trasporto internazionale di merci era prevista per il 2008 dal primo pacchetto ferroviario. Il terzo pacchetto include la Direttiva 23 ottobre 2007, n. 58 (liberalizzazione del trasporto passeggeri), la Direttiva 23 ottobre 2007, n. 59 (certificazione dei macchinisti) e il Regolamento CE 23 ottobre 2007, n. 1371 (diritti e obblighi dei passeggeri).

incontra notevoli ostacoli (fra cui quello della mancanza della reciprocità). In Italia il recepimento della normativa europea è a uno stadio più avanzato rispetto ad altri paesi europei, ma non sempre anche dal punto di vista sostanziale.

In Italia il “primo pacchetto ferroviario” è stato recepito con il decreto legislativo 8 luglio 2003, n. 188, che introduce:

- i)* il principio di libertà di accesso a nuove imprese ferroviarie (che devono ottenere una licenza e un certificato di sicurezza);
- ii)* la separazione (giuridica, organizzativa o decisionale) tra gestore dell’infrastruttura e del servizio; al primo è affidato anche il compito di riscossione dei canoni, di assegnazione della capacità e di rilascio dei certificati di sicurezza;
- iii)* la separazione nel bilancio tra la gestione del servizio passeggeri e quella di trasporto merci;
- iv)* una chiara attribuzione nel bilancio degli eventuali sussidi, con divieto di effettuare sussidi incrociati;
- v)* l’attribuzione al Ministero delle infrastrutture e dei trasporti (MIT) delle funzioni di regolatore del mercato. Sono escluse dall’applicazione del decreto le reti ferroviarie locali e regionali isolate (non interconnesse con la rete nazionale)⁶ adibite al trasporto passeggeri. Per le ferrovie regionali e locali rientranti nel decreto è richiesta la separazione aziendale e contabile fra gestione dell’infrastruttura e gestione del servizio⁷.

Il quadro normativo nazionale si è ulteriormente allineato alle prescrizioni comunitarie attraverso una serie di norme⁸, fra le quali la legge 23 luglio 2009, n. 99, che recepisce il terzo pacchetto e liberalizza il trasporto ferroviario di passeggeri dal 1° gennaio 2010, introducendo tuttavia la possibilità che il servizio passeggeri possa essere soggetto a limitazioni.

L’insieme di queste norme avvia una liberalizzazione del trasporto merci e di parte del trasporto nazionale a lunga percorrenza di passeggeri ma mantiene alcuni segmenti del trasporto passeggeri nell’ambito dei “servizi essenziali”, per i quali non è attuabile una concorrenza nel mercato. Già dalla fine degli anni sessanta la normativa comunitaria aveva introdotto la possibilità per le autorità pubbliche di stipulare un CS pubblico con imprese di trasporto al fine di garantire servizi di

⁶ In genere tutte le ferrovie a scartamento ridotto, come ad esempio la Circumetnea o le ferrovie suburbane della Campania (Circumvesuviana, Sepsa).

⁷ Il D.lgs. n. 188 del 2003 è stato adeguato al terzo pacchetto comunitario con il decreto legislativo 25 gennaio 2010, n. 15.

⁸ Fra queste, il D.lgs. 10 agosto 2007, n. 162 (relativo ad aspetti di sicurezza) e il D.lgs. 10 agosto 2007, n. 163 (che recepisce il secondo pacchetto europeo).

trasporto ritenuti sufficienti (servizi minimi) o tariffe agevolate per particolari categorie di utenti. Nella misura in cui tali servizi siano in contrasto con l'interesse commerciale dell'impresa, l'autorità pubblica, a fronte dell'obbligo di offrire detti servizi, eroga una compensazione corrispondente che non costituisce aiuto di stato⁹.

In Italia sono stati sottoposti a obblighi di servizio pubblico alcuni servizi nazionali e il trasporto regionale. Per il trasporto notturno di passeggeri, la legge 1 agosto 2002, n. 166 prevedeva l'affidamento con gara entro la fine del 2003, prorogato poi alla fine del 2005¹⁰ e nel frattempo la stipula di un CS con Trenitalia¹¹. La legge n. 222 del 2007 ha però sostituito l'obbligo di procedure concorsuali con la previsione di contratti almeno quinquennali (rivedibili annualmente). Se fino al 2006 tutte le tariffe passeggeri erano stabilite dal Cipe, all'inizio del 2007 Trenitalia ha introdotto per la prima volta, unilateralmente, adeguamenti tariffari su Eurostar (ES), intercity (IC; allora non sottoposti a obblighi di servizio) e alta velocità (AV), introducendo di fatto un principio di autonomia tariffaria sul servizio in concorrenza.

A livello locale il cambiamento fondamentale è avvenuto con il decreto legislativo 19 novembre 1997, n. 422 (decreto Burlando)¹², che ha previsto il trasferimento alle Regioni, entro il 1° gennaio 2000 e sulla base di accordi di programma con lo Stato, delle competenze di programmazione, amministrazione e controllo e delle risorse finanziarie relative al trasporto ferroviario locale dei passeggeri. Nel caso delle ferrovie in gestione commissariale governativa affidate per la ristrutturazione a FS e di quelle di interesse regionale e locale in concessione a soggetti diversi da FS era previsto anche il trasferimento a titolo gratuito dei beni, delle

⁹ Regolamento CE 26 giugno 1969, n. 1191 come modificato dal Regolamento CE 20 giugno 1991, n. 1893 e sostituito dal Regolamento CE 23 ottobre 2007, n. 1370. La base normativa di riferimento in materia di obblighi di servizio pubblico è l'art. 73 del Trattato CE. Le condizioni affinché il sussidio non sia da considerare aiuto di stato sono state precisate nella sentenza Altmark, 2003.

¹⁰ Art. 6 del D.l. 24 giugno 2003, n. 147, convertito con modificazioni nella legge 1° agosto 2003, n. 200.

¹¹ Art. 38 della legge n. 166 del 2002 e modifiche (art. 9 della legge 29 novembre 2007, n. 222 di conversione del D.l. 1 ottobre 2007, n. 159); sono sussidiate anche alcune agevolazioni tariffarie per determinate categorie di passeggeri.

¹² E le successive modifiche, in particolare il D.lgs. 20 settembre 1999, n. 400.

¹³ Per le Regioni a statuto speciale (RSS) e le province autonome di Trento e Bolzano il trasferimento di funzioni e risorse doveva avvenire mediante norme di attuazione. Solo nel 2009 la legge n. 99 del 2009 ha attribuito alle due province autonome la competenza dei servizi ferroviari locali svolti su quei territori, utilizzando le risorse già destinate a tale titolo al pagamento dei corrispettivi in favore di Trenitalia.

infrastrutture e degli impianti¹³. Sono state trasferite alle Regioni anche le competenze relative ai treni interregionali (ora ricompresi nella categoria dei regionali), assegnandole di volta in volta a una delle amministrazioni interessate (Stagni, 2004).

Altre importanti innovazioni erano costituite dalla separazione contabile (o costituzione di imprese separate) fra gestore della rete e impresa ferroviaria, dall'introduzione di un CS con il gestore di durata non superiore a 9 anni (modificata in 6 anni rinnovabili di altri 6)¹⁴, che preveda compensazioni economiche per OSP, dalla scelta del gestore mediante gare entro il 2003. L'applicazione di quest'ultimo principio è stata graduale e incerta.

Una delle difficoltà più importanti incontrate dalle Regioni nell'avviare le gare è stata quella del reperimento del materiale rotabile, che Trenitalia non ha loro trasferito, nonostante fosse stato acquistato con contributi statali e regionali (Corte dei conti, 2004). In base al D.lgs. n. 422 del 1997 (come modificato dalla legge n. 166 del 2002), il bando di gara per il servizio regionale doveva garantire che la disponibilità delle reti, degli impianti e delle altre dotazioni patrimoniali necessarie non costituissero un elemento discriminante nella valutazione delle offerte e che questi beni fossero messi a disposizione dell'aggiudicatario. Tuttavia, come segnalato anche dall'Antitrust, la quasi totalità del materiale rotabile era, ancora nel 2003, di proprietà di Trenitalia, gestore di gran parte del servizio regionale in regime di affidamento diretto (AGCM, 2003a)¹⁵. L'AGCM aveva suggerito l'attuazione in tempi rapidi delle gare, permettendo che l'attività dell'aggiudicatario avesse inizio solo dopo un tempo sufficiente per reperire il materiale rotabile e aveva sconsigliato la soluzione di costituire imprese ferroviarie dove confluissero come soci Trenitalia, le Regioni e gli enti locali e le società di gestione delle ferrovie in concessione. Tuttavia quest'ultimo approccio ha prevalso in quasi tutte le regioni in cui il servizio è stato affidato con gara (cfr. oltre). Non è stata invece presa in considerazione la possibilità di trasferire il materiale rotabile a separate *rolling stock companies*.

¹⁴ Legge 9 aprile 2009, n. 33. I CS devono assicurare un rapporto tra ricavi da traffico e costi operativi almeno pari a 0,35 entro il 1° gennaio 2000, al netto dei costi dell'infrastruttura (cioè del "pedaggio" a RFI nel caso dei servizi ferroviari), cfr. D.lgs. n. 422 del 1997, art. 19, comma 5.

¹⁵ Solo il materiale delle ferrovie regionali (*ex concesses*) è di norma di proprietà del demanio della Regione. Alcune ROSCOs europee sono attive anche in Italia, ma prevalentemente nel noleggio di locomotive per i servizi merci "liberalizzati" (cioè esercitati da imprese diverse da Trenitalia).

Già alcuni mesi prima dell'eliminazione dell'obbligo di gara nel trasporto regionale (legge n. 99 del 2009), che avrebbe potuto svolgere una funzione di selezione fra i potenziali gestori, è stata addirittura emanata una normativa che, fra questi, privilegiava Trenitalia. Il decreto legge 29 novembre 2008, n. 185 (e la legge 28 gennaio 2009, n. 2 di conversione) autorizzava infatti una erogazione di risorse statali di 480 milioni di euro per ciascuno degli anni 2009, 2010 e 2011 alle Regioni qualora avessero stipulato i nuovi CS con Trenitalia¹⁶.

4. - L'assetto attuale del settore

4.1 *Il trasporto nazionale*

Nel 1992 l'ente Ferrovie dello Stato (FS), costituito nel 1986, è stato trasformato in società per azioni, di proprietà interamente pubblica; ad esso è stata successivamente attribuita la concessione per l'esercizio del servizio ferroviario pubblico¹⁷. La separazione tra gestore dell'infrastruttura e dei servizi di trasporto è stata completata nel luglio 2001 con la creazione di Rete ferroviaria italiana (RFI). Nel 2002 il gruppo FS è stato completamente ridisegnato, creando una *holding* Ferrovie dello Stato spa, di proprietà del Ministero dell'Economia e delle finanze (MEF) con al suo interno numerose società autonome, tra cui RFI e Trenitalia.

A RFI è stata trasferita la titolarità della concessione prima in capo a FS, la cui durata è stata ridotta nel 2000 da 70 a 60 anni¹⁸ e, sulla base di un contratto di programma tra FS e il MIT, la gestione dell'infrastruttura (inclusi gli investimenti), di proprietà dello Stato. È compito di RFI amministrare il funzionamento della rete assicurando la circolazione in sicurezza dei treni, nel rispetto del criterio

¹⁶ Le risorse sono state erogate anche con la giustificazione formale di garantire che nel 2009 non vi fossero aumenti tariffari e si aggiungono a quelle stanziare con i due decreti del Presidente del Consiglio dei Ministri del 2000, attuativi del decreto Burlando, che trasferivano alle Regioni le competenze e le risorse dei servizi ferroviari regionali (complessivamente 1,2 mld di euro per Trenitalia, più circa 670 milioni per le *ex* concesse (cfr. la legge 24 dicembre 2007, n. 244, tabella 1, colonna 2), e con il decreto 1 agosto 2001 del MEF (di ripartizione del finanziamento di 80 mld. di lire di cui all'art. 52, co. 11, l. n. 388 del 2000, relativo ai CS stipulati dalle RSO con FS, per far fronte a servizi aggiuntivi). Il MIT ha poi disposto un riparto di tali risorse in proporzione a quelle storicamente riconosciute (50 mln alle RSS e 430 mln alle RSO).

¹⁷ Il 1 gennaio 1986 l'Azienda autonoma venne trasformata in ente pubblico. L'attribuzione è avvenuta nel 1993 in base al D.m. 26 novembre 1993, n. 225/T, ma è stata perfezionata sotto il profilo giuridico soltanto nel 1995.

¹⁸ D.m. 31 ottobre 2000, n. 138/T.

di tendenziale equilibrio tra le entrate e le uscite¹⁹. Sotto il profilo commerciale, l'attività di RFI consiste nella vendita delle tracce orarie sulla rete nazionale alle imprese ferroviarie e nell'attuazione degli interventi di manutenzione e ampliamento programmati, a fronte dei quali RFI riceve dallo Stato contributi in conto esercizio²⁰. L'allocazione delle tracce avviene in contropartita di un pedaggio e sulla base del Prospetto informativo della rete (PIR), approvato dal MIT con la supervisione del CIPE e pubblicato annualmente da RFI. Anche le ferrovie regionali devono richiedere le tracce a RFI per la parte di trasporto regionale che transita sulla rete nazionale.

A Trenitalia, nata nel giugno 2000 e articolata in tre divisioni (Cargo, Passeggeri e Trasporto regionale), è stata trasferita la quasi totalità del materiale rotabile circolante in Italia, ad eccezione di quello relativo alle *ex* ferrovie in concessione, attribuito alle Regioni (cfr. sopra). Trenitalia gestisce, sulla base di una licenza, un servizio di trasporto passeggeri nazionale e sulla base di un CS il trasporto nazionale passeggeri sottoposto a obblighi di servizio pubblico. Nonostante le modifiche societarie, il gruppo FS continua a rappresentare un'unica entità economica. Questa situazione è altamente critica sotto il profilo concorrenziale sia per il legame verticale fra segmenti di attività sia per la possibilità di sussidi incrociati fra i diversi servizi (merci e passeggeri, passeggeri sul mercato concorrenziale e su quello con obblighi di servizio pubblico), che ostacolano l'entrata o la crescita di imprese concorrenti; anche l'Autorità antitrust lo ha più volte segnalato (AGCM, 2003*b*, 2008). Inoltre Trenitalia continua a svolgere, da sola o come socio del gestore, in quasi tutte le regioni, il servizio di trasporto regionale su tratte che riguardano sia la rete di competenza delle Regioni sia quella nazionale. Questa sovrapposizione sulle stesse direttrici di servizi a mercato e servizi sussidiati offre forti incentivi a Trenitalia (e a RFI nell'assegnazione delle tracce orarie ai treni regionali operanti su parti della rete nazionale) a privilegiare i primi anche a scapito dei secondi (con effetti in termini di soppressione di corse, tempi di percorrenza, puntualità, ecc.; AGCM, 2009).

La situazione particolarmente favorevole all'*incumbent* non ha stimolato una significativa entrata di nuove imprese nel servizio nazionale. Oltre a Trenitalia

¹⁹ RFI è inoltre tenuta a presentare un Piano di priorità degli investimenti (PPI), da elaborare in conformità al piano d'impresa e sulla base degli indirizzi stabiliti dal Piano generale dei Trasporti e della Logistica, da sottoporre alla valutazione del CIPE.

²⁰ L'ammontare effettivo annuo dei contributi statali è di fatto definito ogni anno dalla legge finanziaria (PONTI M., 2008).

nel novembre 2010 era stato avviato un nuovo servizio fra Torino e Milano da parte di Arenaways, società a capitale privato con vocazione nel trasporto regionale. In seguito la società ha esteso l'offerta di trasporto anche sulla linea Torino-Livorno e ha gestito sul territorio italiano treni turistici per il trasporto con auto al seguito che collegavano Alessandria e Trieste con la Germania e l'Austria. L'Ufficio per la Regolazione dei Servizi Ferroviari ha ordinato da subito la cancellazione delle fermate intermedie tra Torino e Milano, perché in contrasto con la regolamentazione del trasporto regionale²¹. La società, fallita nell'estate del 2011, è stata riacquistata nel 2012 da un nuovo gruppo di imprenditori ed è prevista la ripresa dell'attività nel corso dell'anno. NTV, il primo operatore privato italiano sulla rete ferroviaria ad alta velocità, ha ottenuto la licenza nel 2007 e nello stesso anno ha sottoscritto il protocollo d'intesa sui "servizi universali", con il quale si è impegnata a partecipare alla copertura degli oneri dei CS per le tratte non remunerative. Ad aprile 2012 NTV ha avviato il trasporto viaggiatori. A regime saranno servite 9 città italiane e 12 stazioni, con circa 50 corse giornalieri.

Nel trasporto internazionale di passeggeri sono attive le ferrovie svizzere (FFS), SNCF (ferrovie francesi), Trenitalia e la Società Trenitalia Veolia Transdev; da dicembre 2009 DB e OBB (ferrovie tedesche e austriache) operano congiuntamente sull'asse del Brennero, con la trazione a cura della società LeNORD. Nel dicembre 2010 l'Ufficio per la regolazione dei servizi ferroviari ha vietato le fermate intermedie in territorio italiano dei treni internazionali DB-OBB, sospendendo successivamente il divieto per tre mesi; successivamente la sospensione è stata prorogata a tempo indeterminato, escludendo tuttavia la fermata di Vicenza. Fra il 2009 e il 2010 ha operato nell'area barese il Consorzio Acquario (costituito da Trenitalia e da Ferrotramviaria) senza usufruire direttamente di contributi pubblici.

I contratti di programma dello Stato con RFI. - I rapporti fra gestore dell'infrastruttura e Stato sono disciplinati da una concessione e da un contratto di programma di durata minima triennale, aggiornabile anche annualmente, in cui sono definiti gli obiettivi in termini di manutenzione ordinaria, straordinaria e sviluppo della rete, nonché le modalità del finanziamento statale²². Il contratto di pro-

²¹ Sulla vicenda è stato anche interessato il Tar Lazio; l'Antitrust ha aperto un'istruttoria nei confronti del gruppo FS e della controllata Trenitalia per abuso di posizione dominante sul mercato dell'accesso all'infrastruttura ferroviaria nazionale nei confronti di Arenaways.

²² D.lgs. n. 188 del 2003, che ha anche ridotto la durata minima del contratto, stabilita dal D.p.r. n. 277 del 1998 in cinque anni.

gramma può inoltre prevedere un indennizzo al gestore dell'infrastruttura qualora l'assegnazione di tracce per lo sviluppo del trasporto merci o per obblighi di servizio pubblico di passeggeri determini perdite finanziarie.

Il primo contratto di programma, valevole sino al 2000, venne introdotto con la finanziaria del 1994 e fu sottoscritto nel 1996 con FS che era, all'epoca, anche titolare della concessione per la gestione della rete. Il successivo (2001-05), nel quale subentrò RFI, nel frattempo divenuta entità giuridicamente distinta da FS, venne prorogato per il 2006 con delibera del CIPE (2 dicembre 2005, n. 159) e fu oggetto di varie modifiche, l'ultima delle quali apportata nel 2007, che prorogava la validità dell'accordo fino alla stipula del nuovo contratto 2007-2011. L'aggiornamento apportato in febbraio 2008 ha ripartito le maggiori risorse stanziare con il DL 159/2007 (art. 2) e da altre fonti (FESR, finanziaria 2007, ecc.), per complessivi 3.158 milioni di euro²³. L'aggiornamento 2010-11 è stato esaminato dal CIPE nel gennaio 2012.

I CS dello Stato con Trenitalia. - Dal 1994 i rapporti tra Stato e FS, poi Trenitalia, per gli obblighi di trasporto e tariffari, sono stati disciplinati da un CS sottoscritto tra Trenitalia e il MIT, di durata dapprima triennale, quindi di durata minima quinquennale, rivedibile annualmente²⁴.

I vari contratti succedutisi nel tempo sono stati stipulati sempre successivamente alla prevista scadenza. Con il contratto 2000-2001, sottoscritto dopo il periodo di riferimento, fu introdotta una clausola di continuità, in base alla quale l'erogazione del servizio era garantita fino alla stipula del successivo contratto; in questo modo, ad esempio, alcuni servizi merci vennero prorogati sulla base del contratto 2000-2001 fino all'applicazione di quello specifico 2004-06, sottoscritto nel 2007.

Il CS 2004-05, prorogato più volte fino al 2009 riguardava una parte del trasporto merci (cioè quello disciplinato con il contratto 2004-2006 specifico per i servizi merci sottoposti al OSP), una parte del servizio nazionale passeggeri a

²³ In base all'art. 3 del contratto, a decorrere dal 2008 eventuali modifiche possono essere definite per tener conto di ulteriori risorse stanziare dalla legge finanziaria, entro il mese di gennaio di ogni anno, oppure in conseguenza di sopravvenute variazioni delle risorse disponibili o di sopravvenuti obblighi di legge, su istanza di ciascuna parte. Tali modifiche devono essere sottoposte al parere del CIPE e delle competenti Commissioni parlamentari.

²⁴ Legge n. 166 del 2002 così come modificata dal D.l. n. 159 del 2007, art. 9 co. 2-*bis* (convertito con legge n. 222 del 2007), che ha tra l'altro cancellato l'obbligatorietà delle gare per questo tipo di servizi.

media e lunga percorrenza (MLP) e – in attesa della completa attuazione del decreto Burlando – anche il servizio di trasporto nelle RSS e nelle province autonome di Trento e Bolzano. Alla fine del 2008 è scaduto il CS relativo al servizio IC giorno e il servizio notte è prestato sulla base di una clausola di continuità del contratto 2004-2005. Il CS 2009-2014 è stato sottoscritto a marzo 2010.

Il perimetro dei servizi inclusi negli OSP è variato più volte nell'ultimo triennio: così, mentre nel 2007 erano previsti i soli servizi notturni di MLP, nel 2008 venivano inclusi e pertanto sussidiati anche i servizi IC e IC notte; a dicembre 2009 il CIPE ha ampliato la tipologia di servizi di utilità sociale relativi al trasporto ferroviario passeggeri sulla MLP per i quali non è possibile raggiungere l'equilibrio economico includendo, oltre ai treni notturni, gli IC e, almeno per il 2010, alcuni ES di collegamento fra Roma e Umbria, Puglia e Calabria, assimilabili a IC²⁵. Allo stesso tempo ha escluso i servizi internazionali e i servizi MLP erogati in modo totale o prevalente sulla cosiddetta rete forte (tratta Milano/Napoli, Torino/Venezia, Bologna/Venezia, Bologna/Verona). Oltre a un nuovo perimetro dei servizi, il nuovo contratto (2009-2014) prevede anche una riduzione dell'offerta, in particolare del servizio IC. Alla fine del 2010 Trenitalia ha manifestato l'intenzione di sopprimere alcuni servizi a MLP (circa 154 treni) non redditizi per l'azienda e non sussidiati dal CS.

Come per il trasporto regionale, anche per i servizi universali di MLP, FS ha perseguito l'obiettivo di rafforzare l'equilibrio economico proponendo nuovi CS basati su uno schema dettagliato di servizi e relativi corrispettivi richiesti, date le tariffe fissate dal MIT (il cosiddetto Catalogo; Corte dei conti, 2009). Va però detto che, al pari dei servizi regionali (di cui si tratterà al punto seguente), anche per i servizi MLP Trenitalia ha beneficiato di risorse aggiuntive, con carattere di *una tantum* ricorrente: ad esempio 104 milioni di euro nel 2008 (l. n. 244 del 2007, art. 2, comma 252) e poi 110 milioni per ciascuno degli anni 2009, 2010 e 2011²⁶.

²⁵ La finanziaria per il 2008 aveva attribuito al CIPE il compito di individuare il perimetro dei servizi di trasporto da assoggettare all'obbligo di servizio pubblico, disponendo il regime di liberalizzazione per le tratte sulle quali sia possibile il mantenimento o il raggiungimento dell'equilibrio economico. Nel marzo 2008 era stato raggiunto un accordo ponte che aveva aggiunto alcuni servizi al vigente contratto sugli espressi notte. Rientra negli obblighi di servizio anche parte del trasporto merci, in particolare i collegamenti con le isole maggiori, i trasporti superiori ai 1.000 km, i collegamenti fra il porto di Trieste e l'Ungheria, il trasporto internazionale via porto di Trieste.

²⁶ Deliberazione CIPE 8 maggio 2009, n. 23, pubblicata su G.U. 29 dicembre 2009.

4.2 *Il trasporto regionale*

L'attuazione del decreto Burlando risulta ancora parziale e incompleta. Nel caso delle RSS esso è stato attuato (seppure molto di recente) per le Province di Trento e Bolzano e per il Friuli-Venezia Giulia²⁷. Per la Sardegna e la Valle d'Aosta sono già state deliberate le norme attuative (decreto legislativo 21 febbraio 2008, n. 46 e decreto legislativo 26 ottobre 2010, n. 194). Per la Valle d'Aosta sono state trasferite le funzioni di programmazione, mentre la rete è interamente di proprietà statale e di competenza RFI. Per la Sardegna la proprietà delle reti a scartamento ridotto è stata trasferita alla Regione, ma non è stata ancora definita quale parte della rete RFI sia da considerarsi a carattere regionale (e quindi eventualmente da trasferire alla Regione) né quale parte del servizio ancora svolto da Trenitalia e finanziato dallo Stato sia di competenza della Regione²⁸. Fra le regioni a statuto ordinario (RSO), in Puglia la rete non è ancora stata incorporata nel demanio regionale. La Sicilia non ha ancora avviato il processo di decentramento.

Riguardo alla fornitura del servizio, il trasporto regionale è fornito principalmente da Trenitalia cui si affiancano una ventina di gestori di proprietà prevalentemente regionale e provinciale, in alcuni casi consorziati al gestore nazionale (Tavole 1 e 2). Trenitalia esercita in esclusiva il servizio nelle Marche, in Molise e in Valle d'Aosta, e in quasi esclusiva (con la presenza di un solo altro gestore di dimensioni molto piccole) in Sardegna e in Sicilia.

Il trasferimento alle Regioni del materiale rotabile è avvenuto soltanto parzialmente in quanto ha riguardato solo le ferrovie regionali (*ex* concesse) nonostante le previsioni normative originarie. La quasi totalità del materiale rotabile infatti è di proprietà di Trenitalia, anche nei casi in cui Trenitalia lo ha acquistato o rinnovato utilizzando risorse finanziarie fornite dagli Enti (come largamente avvenuto intorno al 2002-04 ad esempio per i treni *Minuetto*). Questo aspetto ha reso complessa la specificazione, nei bandi di gara e/o nei CS, delle modalità di fornitura e reperimento del materiale rotabile da parte del gestore affidatario e la disciplina della cessione o meno dello stesso materiale in caso di mancato rinnovo

²⁷ Per la Regione Friuli-Venezia Giulia il trasferimento dal demanio statale a quello regionale degli impianti e delle infrastrutture relative alla linea Udine-Cividale è avvenuto nell'ottobre 2010.

²⁸ La definizione della parte dei servizi attualmente erogati da Trenitalia da trasferire sotto le competenze della Regione deve avvenire mediante un accordo di programma fra MIT, MEF e Regione Sardegna. Nel caso della Valle d'Aosta il decreto legislativo prevede la possibilità, su richiesta della Regione e mediante un accordo di programma, di trasferire a quest'ultima le infrastrutture ricadenti nel territorio regionale non ritenute di rilevanza nazionale.

o subentro di un nuovo gestore. I nuovi CS (a Catalogo) con Trenitalia recepiscono lo stato di fatto e prevedono che Trenitalia fornisca il materiale necessario e provveda, in parte con l'autofinanziamento e in parte con il sostegno delle Regioni contraenti, all'acquisto e al rinnovo del parco rotabile sulla base di uno specifico programma definito per ciascuna regione. La durata dei nuovi CS è funzionale alla realizzazione di tale programma che prevede un ammortamento del costo su un arco temporale molto ampio (24 anni), consentendo inoltre a Trenitalia, in caso di mancato rinnovo del contratto o di subentro di un nuovo gestore, di rientrare delle spese sostenute e di trasferire alla Regione o all'impresa ferroviaria subentrante i treni e la quota di oneri non ancora ammortizzati²⁹. Indipendentemente dall'"autofinanziamento", negli ultimi anni alcuni provvedimenti legislativi hanno previsto lo stanziamento di risorse dedicate al finanziamento dell'acquisto di veicoli destinati al trasporto pubblico locale, incluso quello ferroviario, da parte delle regioni³⁰.

I bandi di gara e/o i CS definiscono in alcuni casi parametri di misurazione della qualità del servizio e forme di monitoraggio (compiti trasferiti alle Regioni dalla riforma), che consentono l'applicazione di sistemi di penalità-premio nei confronti del gestore. In generale i CS prevedono la rilevazione da parte del gestore e la comunicazione periodica alla Regione di alcuni dati riguardanti la puntualità e la pulizia, che vengono poi pubblicati da quest'ultima. Nel caso dei CS con Trenitalia, a tali report si aggiungono periodiche indagini sulla *customer satisfaction*, curate dal gestore. Soltanto in alcuni casi le indagini effettuate dal gestore sono affiancate da analoghe iniziative periodiche promosse dalla Regione (ad esempio Toscana, Marche, Lombardia, Liguria), così come costituiscono una minoranza gli Enti che provvedono a pubblicare regolarmente i risultati dell'attività di monitoraggio (ad esempio Piemonte, Liguria, Emilia-Romagna, Sardegna). Ne consegue una situazione in cui di fatto il controllo risulta delegato al gestore stesso.

Nel periodo 2001-2010 in numerose regioni sono state applicate penalità per un totale di oltre 102 milioni di euro (Legambiente, 2009, 2010). Il 75,6% riguarda 4 regioni (Lombardia, Lazio, Emilia-Romagna e Veneto). Le risorse re-

²⁹ Complessivamente l'investimento ammonta a 1.500 mln di euro oltre a 500 mln di contributi FAS. I CS prevedono penalità in caso di ritardo nella consegna del materiale rotabile.

³⁰ Si tratta in particolare della legge 27 dicembre 2006, n. 296, della legge n. 244 del 2007 e della legge n. 133 del 2008. Gli importi previsti per l'acquisto del materiale rotabile risultano tuttavia modesti, in quanto a tale finalità è riservata soltanto una quota della dotazione dei due fondi così istituiti, il cui finanziamento copre l'arco temporale compreso tra il 2007 e il 2011 (complessivamente circa 760 milioni di euro per tutto il periodo).

cuperate nel 2010 sono state destinate alla concessione di *bonus* e agevolazioni tariffarie agli abbonati in Lombardia, Liguria, Emilia-Romagna, Veneto e Toscana, mentre in Campania, Lazio, Piemonte e Basilicata sono state reinvestite in iniziative finalizzate al miglioramento della qualità del servizio e del materiale rotabile e a servizi aggiuntivi, mentre nelle Marche e in provincia di Bolzano non era stata decisa la destinazione delle risorse rivenienti dalle penali. Negli altri casi le penali o non risultano applicate o non erano previste prima del 2009, come nel caso delle RSS. In Piemonte a fine 2010 le penali sono state sospese.

Le gare per il servizio regionale. - La proroga reiterata della scadenza entro la quale le Regioni avrebbero dovuto affidare il servizio con gara ha rallentato la realizzazione della concorrenza “per il mercato” nel trasporto ferroviario regionale, così come il mancato trasferimento da Trenitalia alle Regioni del materiale rotabile (Cambini e Margari, 2005; Parolin e Klarman, 2007). Fino al 2010 soltanto in Lombardia, Emilia-Romagna e Veneto si è provveduto ad aggiudicare il servizio con gara, stipulando un nuovo contratto. In Liguria è stata vinta da AMT la gara per il servizio sulla Genova-Casella, linea *ex* concessa di proprietà della Regione (unica gara effettuata su una *ex* concessa isolata, cioè non interconnessa alla rete RFI); AMT è subentrata nella gestione del servizio il 16 aprile 2010 (Tavola 2). Nel 2009 anche una larga fetta del trasporto locale piemontese, suddiviso in tre lotti, era stato messo a gara; le procedura è stata successivamente sospesa³¹ e, nello scorcio del 2010, la Regione ha stabilito che ne sarà indetta una nuova relativa all’intero trasporto ferroviario regionale. Nel frattempo il servizio è stato affidato a Trenitalia fino al 2016 sulla base di un contratto ponte che usufruisce dei fondi stanziati dal D.l. n. 185 del 2008. A fine 2011 la Giunta ha deliberato di assegnare a SCR Piemonte S.p.A. l’incarico di predisporre il relativo bando di gara.

La gara, poi sospesa, relativa ai primi tre (dei cinque) lotti piemontesi (nord-est, sud-est e la Torino-Milano) per la gestione del servizio a partire dal 2013, per circa il 60% degli attuali treni-km annui complessivi per oltre 100 mln annui di corrispettivo, presentava alcune caratteristiche innovative che sembravano tese a limitare il vantaggio dell’*incumbent*. La partecipazione alla gara per la tratta Torino-Milano (quella commercialmente più ricca) veniva vincolata alla presentazione di un’offerta per almeno uno degli altri lotti; il corrispettivo offerto dalla Regione variava in funzione inversa al potenziale di redditività della tratta; come in altri

³¹ La nuova amministrazione ha motivato la decisione sulla base delle difficoltà di accesso ai depositi-officine attualmente esistenti.

casi (ad es. Veneto ed Emilia-Romagna) il materiale rotabile doveva essere appor-
tato dall'affidatario, seppure entro tre anni dall'aggiudicazione, ed erano previsti:

- a) il mantenimento del vincolo di destinazione d'uso della parte del parco rotabile finanziata dalla Regione e la sua restituzione all'ente da parte del gestore ces-
sante per cederlo a quello subentrante se diverso, oppure la restituzione della
quota dei finanziamenti ricevuti per il periodo relativo al mancato utilizzo³²;
- b) il diritto di prelazione dell'aggiudicatario sugli altri beni senza vincolo di destina-
zione d'uso. Alla scadenza per la presentazione delle domande (15 febbraio 2010)
avevano manifestato interesse 4 concorrenti.

Oltre a Trenitalia, in associazione con LeNORD (TLN), vi erano importanti ge-
stori esteri: la multinazionale inglese Arriva, presente in 12 paesi europei, la francese
Veolia in ben 28, ma anche negli USA e in Nuova Zelanda, e le Ferrovie Federali
Svizzere (che sembravano interessate anche per la contiguità territoriale con il Pie-
monte nord-orientale)³³.

In tutti e tre i casi in cui la procedura ha avuto seguito, l'aggiudicatario è ri-
sultato un gruppo di imprese costituito essenzialmente dagli *incumbent*, inclusa
Trenitalia. Soltanto in Emilia-Romagna l'affidamento ha riguardato l'intero ser-
vizio, mentre nelle altre due regioni è stato parziale (circa $\frac{3}{4}$ del servizio per il Ve-
neto a Trenitalia insieme a Sistemi Territoriali S.p.A.). Anche la Liguria aveva
avviato nel 2004 procedure di gara per l'intero servizio regionale, ma entrambi
gli offerenti sono stati esclusi: Trenitalia per aver formulato un'offerta condizio-
nata e LeNORD per aver avanzato un'offerta non conforme al capitolato di gara.
In Lombardia, due dei tre lotti posti a gara nel 2004 non sono stati aggiudicati
in quanto l'emissione del bando non è stata seguita dalla pubblicazione del capi-
tolato e quindi non si è mai proceduto a ricevere le offerte (Stagni, 2010). Tra le
principali determinanti dell'insuccesso della maggior parte delle gare vi sono le
condizioni relative al materiale rotabile (gli oneri connessi con il reperimento, il
rinnovo o la cessione all'ente appaltante al termine del contratto), probabilmente
ritenute troppo gravose rispetto al corrispettivo offerto dalla Regione e sostan-
zialmente impraticabili se la gara è riferita a un'intera Regione (Stagni, 2004).

³² Non è stata invece data attuazione al progetto della Regione di creare una società mista che
gestisse in proprio anche il parco rotabile.

³³ Con riguardo alla questione del reperimento del materiale rotabile da parte del vincitore, alcune
dichiarazioni rilasciate alla stampa dall'A.D. di Arriva Italia indicherebbero che eventuali pro-
blemi per il rispetto dei tempi previsti nel bando potrebbero sorgere soltanto in relazione a ri-
tardi da parte dei fornitori dei treni, comunque superabili proprio in forza delle dimensioni e
della solidità economica della società e della sua esperienza nel settore.

I contratti per il trasporto regionale. - Al primo CS, relativo al 2000 e siglato fra Stato e FS, è seguito un “contratto sperimentale” (2001) e un “contratto transitorio” 2002-03 tra Trenitalia e le singole RSO (in attesa delle gare). Dopo la scadenza di quest’ultimo contratto, molte Regioni hanno prorogato il contratto (non la Lombardia ad esempio, che ha predisposto un suo CS 2004-2005); tutte hanno fissato le proprie tariffe, in quanto la legge assegna alle Regioni la potestà tariffaria per i servizi ferroviari regionali, e in genere hanno introdotto penali per il mancato rispetto dei parametri quantitativi e qualitativi. Dal 2001 al 2007 le risorse trasferite dallo Stato sono rimaste immutate, anche se il *trend* dei servizi è stato moderatamente in crescita: questo ha contribuito a un progressivo deterioramento dei bilanci di Trenitalia. Solo alcune Regioni (soprattutto la Toscana e la Lombardia) hanno aggiunto risorse proprie che, eventualmente insieme agli aumenti tariffari, hanno finanziato i servizi di nuova istituzione (Stagni, 2004).

Con il D.l. n. 159 del 2007, il MIT ha erogato direttamente a Trenitalia 311 milioni di euro per il 2007, inaugurando una stagione di finanziamenti diretti all’azienda, che scavalcava le Regioni titolari dei contratti. In attesa della rideterminazione dei criteri di ripartizione dei trasferimenti statali alle regioni in materia di trasporti, si è proseguito con un’assegnazione di 80 milioni di euro all’inizio del 2008 (D.l. 8 aprile 2008, n. 60) e di ulteriori 300 milioni pochi mesi più tardi (D.l. n. 112 del 2008).

In generale, dove non sono state effettuate le gare per l’intero servizio, il trasporto ferroviario regionale risulta oggi fornito da Trenitalia e uno o più gestori “regionali” (le *ex* concesse), sulla base di un affidamento diretto, disciplinato da un apposito e distinto CS. Per la maggior parte di queste regioni i rapporti con Trenitalia sono stati recentemente regolati tramite la sottoscrizione di un nuovo contratto (in genere 2009-14), per poter garantire, da un lato, almeno i livelli di servizio preesistenti e, dall’altro, per assicurarsi il sostegno finanziario statale previsto dai provvedimenti normativi emanati tra la fine del 2008 e l’inizio del 2009. Tra i primi a sottoscrivere il nuovo CS con Trenitalia è stata la Provincia Autonoma di Trento (novembre 2008), per la gestione del servizio locale per il periodo 2008-2013. In Valle d’Aosta, Sicilia e Sardegna, l’attuale contratto di servizio con Trenitalia risulta sottoscritto dal MIT anziché dalla Regione.

Per la stipula dei nuovi CS Trenitalia ha presentato ai contraenti un “Catalogo” che introduce un nuovo criterio di determinazione del corrispettivo annuo effettivo, basato su un “listino prezzi” dei vari servizi (sia di trasporto sia accessori, come biglietterie, informazioni e assistenza), che ha sostituito il parametro dei treni-km (come base di calcolo dei corrispettivi) con il costo orario del servizio. Legare una

quota considerevole del costo aziendale al costo orario rende possibili aumenti di fatturato dovuti anche soltanto a una revisione della durata di percorrenza delle singole tratte; l'allungamento dei tempi di percorrenza tra l'altro è stato in tempi recenti un modo per contenere i ritardi, al verificarsi dei quali erano sovente dovute penali da riconoscere alle Regioni, proporzionali alla loro frequenza ed entità. Sono inoltre previste maggiorazioni percentuali sul costo orario per i servizi notturni, di sabato e nei festivi, a bassa domanda, o con materiale nuovo. Poiché il corrispettivo richiesto alla Regione è pari alla differenza tra il costo complessivo come sopra determinato e gli introiti dell'azienda (biglietti), si realizza di fatto un sistema *gross cost* che trasferisce il rischio di domanda dal gestore alla Regione.

5. - Caratteristiche del settore

5.1 *L'infrastruttura*

La proprietà pubblica nel settore del trasporto ferroviario non è sempre stata una caratteristica della rete italiana. Per tutto il XIX secolo sul territorio italiano accanto all'operatore pubblico hanno operato numerosi privati. La ferrovia nasce in Italia nel 1839 con l'inaugurazione della Napoli-Portici (7,640 km). L'esempio del Regno delle Due Sicilie fu seguito da altri Stati italiani e nel 1861 la rete era lunga poco più di 2.000 km (il 18% era di proprietà dello Stato ed il 25% in sua gestione diretta). Alla sua costruzione contribuirono in maniera significativa capitali privati, anche esteri (in particolare il banchiere livornese Pietro Bastogi e James Rothschild). Nello stesso periodo la rete ferroviaria era già molto più sviluppata in altri paesi europei (misurava circa 11.700 km in Germania, 14.600 nel Regno Unito e 9.400 in Francia; Berengo Gardin, 1988). Nel 1865 una legge di riordino concentrò rete e gestione del servizio in 5 società concessionarie³⁴. Nel 1872 la rete italiana misurava circa 7.000 km, la cui gran parte era gestita da 4 società. Fra il 1878 e il 1880 lo Stato rilevò due società ferroviarie, colpite da gravi dissesti finanziari. Nel 1884-1885 si decise di riaffidare ai privati la gestione delle ferrovie, lungo la rete di 8.510 km, stipulando convenzioni sessantennali

³⁴ Il riordino è legato al fatto che le garanzie statali sugli investimenti gravavano pesantemente sulle finanze pubbliche e che l'eccessivo frazionamento ostacolava lo sviluppo delle ferrovie. Le 22 società furono ridotte a 5: l'Alta Italia (Rothschild) alla quale si affittarono le ferrovie statali, le Romane (a capitale prevalentemente francese), le SFM (Bastogi), le Calabro-Sicule (a capitale francese) e le Sarde (a capitale prevalentemente inglese).

con 3 società private. La situazione economica delle ferrovie rimase disastrosa e richiese elevati contributi statali.

Nel 1905 sotto il governo Giolitti lo Stato creò l'Azienda autonoma delle Ferrovie dello Stato, che assunse la gestione diretta di 10.557 km di rete (di cui 9.868 già di proprietà statale). Nel 1906, con la confluenza della rete della Società italiana per le strade ferrate meridionali (SFM), la rete statale raggiunse i 13.075 km, di cui 1.917 a doppio binario. Molte altre linee preesistenti, di interesse locale, rimasero tuttavia attive sulla base di concessioni: sono quelle comunemente chiamate "ferrovie concesse", che oggi, in maggioranza trasferite alle Regioni, costituiscono le cosiddette *reti regionali*, cioè le infrastrutture ferroviarie non in capo a RFI. Nel 1912 fu introdotto l'istituto della gestione commissariale governativa (specifico per le aziende di trasporto concessionarie in crisi), modificato più volte nel dopoguerra, attribuendo la gestione commissariale dapprima a un incaricato governativo e infine a FS, prima del loro recente passaggio alle Regioni.

Nel 1922 con l'avvento del fascismo e in presenza di una grave situazione economico-finanziaria, all'Ente fu imposto un commissario in sostituzione del consiglio di amministrazione e due anni dopo fu messo a capo dell'Ente il ministro delle comunicazioni, che ripristinò il consiglio di amministrazione (dal 1944 le competenze passarono al ministero dei trasporti); fu avviata una politica di taglio del personale, che passò da 241.000 a 138.000 unità fra il 1922 e il 1939. La costruzione di nuove reti fu affidata al ministero dei lavori pubblici. In questo periodo le ferrovie conobbero un periodo di notevole sviluppo: nel 1939 la rete aveva raggiunto i 22.800 km (di cui oltre 5.000 elettrificati). La seconda guerra mondiale danneggiò notevolmente il sistema ferroviario (nel giugno 1945 il 40% della rete e gran parte del materiale rotabile erano danneggiati), ma il recupero fu rapido: nel 1949 la rete ferroviaria in funzione aveva raggiunto i 21.369 km e rimase sostanzialmente invariata fino all'inizio degli anni sessanta.

Negli ultimi 40 anni del XX secolo si sviluppano alcuni servizi di punta per i viaggiatori a lungo raggio, alcune reti soprattutto in ambito urbano e suburbano e viene avviata l'alta velocità, con la costruzione della Direttissima Firenze-Roma, la prima del genere realizzata in Europa, iniziata nel 1970 e conclusa nel 1992 (la prima tratta è stata inaugurata nel 1977). A partire dagli anni cinquanta viene avviata, in misura crescente con la scelta politica di privilegiare le infrastrutture stradali, e in particolare nel 1985 con il decreto Signorile, la dismissione di alcune linee e l'abbandono di altre in seguito alla realizzazione di varianti di tracciato, per circa 2.550 km per quelle di pertinenza FS e 3.150 km per le *ex* in conces-

sione. A queste si aggiungono circa 520 km di linee incompiute³⁵. Di recente sono state riattivate alcune linee, di lunghezza molto ridotta.

La rete attiva è andata diminuendo dal 1990 al 2000 (Tavola 3), per poi riprendere una crescita modesta, concentrata nella parte nazionale nell'alta velocità (oltre a circa 250 km nella rete regionale; Tavola 4).

Alla fine di questo decennio la rete italiana supera di poco i 20.000 km, 3.500 dei quali (circa il 17%) di competenza regionale. Le due reti si differenziano qualitativamente, con il 71% circa di quella nazionale elettrificata (a fronte del 37% di quella regionale) e una quota ancora limitata di rete a doppio binario (44% e 12% circa).

Le reti regionali del Mezzogiorno rappresentano oltre il 60% del totale delle reti locali, con un'estensione particolarmente ampia in Puglia, Sardegna e Campania (807, 614 e 345 km; Tavola 4). Seguono al Nord la Lombardia e l'Emilia-Romagna (321 e 302 km, rispettivamente). In molte regioni la rete regionale è completamente non elettrificata e solo in sette regioni (di cui due nel Sud) la quota di rete elettrificata è superiore ai tre quarti del totale. L'estensione complessiva della ferrovia sul territorio regionale riflette in modo significativo le caratteristiche del territorio: da una media di 6,6 km di rete ogni 100 km² di superficie, si scende ai minimi in alcune regioni montane (Valle d'Aosta, Trentino-Alto Adige, Basilicata), ma anche nelle Marche.

I recenti progetti di ampliamento della rete hanno riguardato, a livello nazionale, l'AV, avviata nel 1994 e che affianca la linea storica, cui si aggiungono il programma di riqualificazione di 13 grandi stazioni (inserito nella legge obiettivo) e la riqualificazione e il potenziamento di alcune tratte. La quota di investimenti infrastrutturali nella rete storica si è ridotta dal 46,3% nel 1999 al 28,3% nel 2003 in favore dell'AV per poi risalire al 58,2% nel 2009 (Grafico 1). Gli investimenti complessivi sono in calo dal 2006, sia in valore assoluto sia rispetto al PIL corrente.

Raramente sono state sperimentate forme di finanziamento innovative rispetto ai tradizionali contributi dello Stato. Nel caso dell'AV, l'obiettivo iniziale di un *project finance* con 60% di capitale privato non è stato conseguito. L'investimento non risultava infatti sufficientemente profittevole per gli investitori privati e risultava pertanto una forma mascherata di finanziamento pubblico³⁶. Un caso in-

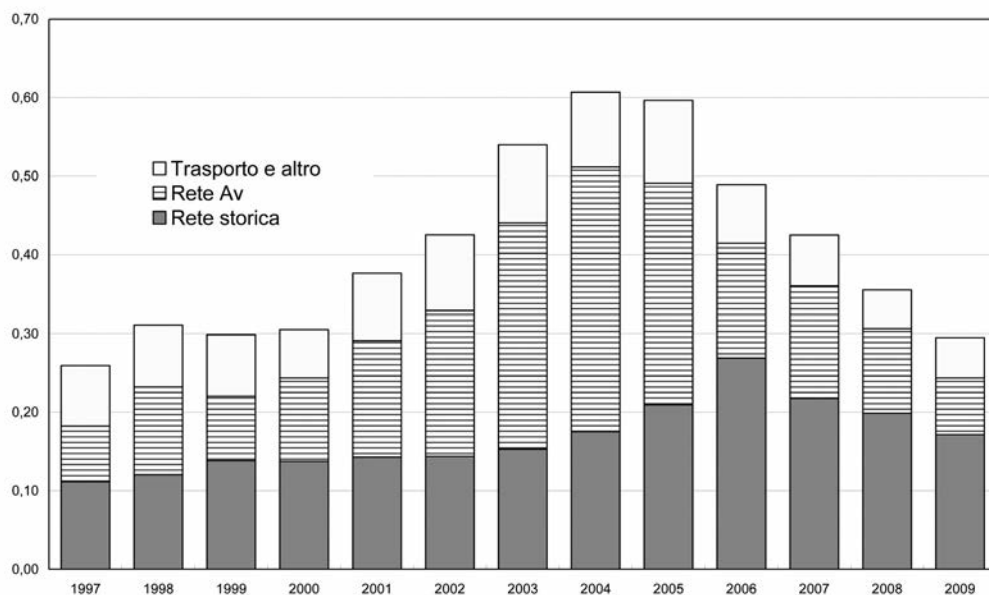
³⁵ Cfr. il sito www.ferrovieabbandonate.it. Ne è un esempio la ferrovia Roma Nomentana-Roma San Pietro (FS), lunga 12,8 km, i cui lavori sono iniziati nel 1913 e interrotti più volte.

³⁶ La società Infrastrutture S.p.A., costituita per finanziare grandi opere (fra cui l'AV) con emissioni obbligazionarie, rientrando del debito attraverso il canone di pedaggio pagato per l'uso della rete (e con lo Stato che si doveva far carico dello sbilancio tra ricavi e servizio del debito), è stata assorbita dalla Cassa depositi e prestiti dopo che l'Eurostat nel 2005 aveva deciso che l'operazione non rispettava i presupposti finanziari per considerarla come *project financing* e per trattare i debiti ISPA come debiti privati (non inclusi cioè nel debito dello Stato).

teressante riguarda il Trentino-Alto Adige, dove l'Autostrada del Brennero, società partecipata dalle amministrazioni locali, ai sensi della legge finanziaria 1998, accantona parte dei proventi da pedaggi in un fondo destinato al rinnovo dell'infrastruttura ferroviaria; l'utilizzo del fondo dovrebbe avvenire attraverso una partecipazione della stessa Autostrada del Brennero Spa alla società che costruirà e/o gestirà il nuovo tunnel ferroviario del Brennero³⁷.

GRAF. 1

INVESTIMENTI DEL GRUPPO FS IN PERCENTUALE DEL PIL
A PREZZI CORRENTI (%)



Fonte: ISTAT e bilanci del gruppo FS.

³⁷ La Direttiva CE 17 giugno 1999, n. 62 e successive modifiche e integrazioni (Eurovignette) consente una forma simile di finanziamento prevedendo la possibilità di aumentare (secondo fasce di incremento predeterminate) i pedaggi per gli autoveicoli pesanti adibiti al trasporto merci su strade congestionate e situate in regioni montane. VAGHI C. e MILOTTI A. (2009) mostrano tuttavia che un aumento dei pedaggi sull'Autobrennero nei limiti indicati dalla Eurovignette per finanziare il nuovo tunnel ferroviario del Brennero permetterebbe di finanziare solo un quarto dell'opera. All'inizio del 2009 erano stati accantonati fondi per circa 385 milioni di euro. L'impresa italo austriaca incaricata della costruzione, e partecipata da FS, non sembrava intenzionata ad aprire il proprio capitale sociale ad Autobrennero S.p.A. (RAGAZZI G., 2009).

5.2 *Il mercato nazionale*

Sul mercato nazionale (MLP), che include sia i servizi di mercato sia quelli sussidiati e soggetti a OSP, fra il 2000 e il 2009 si è avuta una riduzione dei passeggeri del 4,3% (Tavola 8); la riduzione è stata ancora più intensa in termini di passeggeri-km (-19,2%).

Il materiale rotabile utilizzato sulla media e lunga percorrenza è relativamente vecchio, risalendo in maggioranza agli anni settanta e ottanta, con la rilevante eccezione di quello per l'alta velocità (che ha circa 10 anni). Il *load factor* ha mostrato un lieve calo fra il 2000 e il 2008 (arrivando intorno al 50,3% nel 2008).

Nel 2009 i ricavi sul mercato nazionale hanno raggiunto i 1.986 milioni di euro, di cui l'88,7% dal mercato (biglietti e abbonamenti) e la parte restante derivante dal CS per la MLP. Fra il 2001 e il 2009 sia i ricavi da mercato sia quelli derivanti dal CS con lo Stato sono aumentati di circa il 18%. I trasferimenti dallo Stato per il CS MLP sono passati da 190 a 224 milioni di euro (Tavola 9). Secondo alcune stime, i sussidi per la MLP sarebbero troppo elevati. Brenna (2009) stima un costo del trasporto ferroviario MLP parte dell'attuale CS pari a 8,80 euro per treno-km in un contesto concorrenziale, a fronte di un costo medio complessivo dichiarato da Trenitalia di 15,03 per il 2008. L'attuale ricavo medio per la stessa tipologia di treni sarebbe pari a 0,07 € pass.-km, con un *break-even* di circa 126 passeggeri a treno, a fronte di oltre 700 posti offerti per treno.

A fine 2009 è stato completato il sistema di alta velocità sulla tratta Torino-Napoli-Salerno. Il nuovo servizio è stato accompagnato da aumenti del prezzo dei biglietti e da maggiori limiti nei rimborsi per ritardi.

Alcune critiche sono state mosse ai costi dell'infrastruttura relativa all'AV, più alti che in altri paesi (Ponti e Beria, 2007); in questo caso l'investimento resterebbe in gran parte a carico dello Stato, nonostante i fruitori non siano cittadini a basso reddito, e i costi di produzione risulterebbero comunque molto più elevati del corrispondente trasporto aereo. Negli ultimi anni il gruppo FS ha avviato alcuni sforzi di razionalizzazione delle risorse e di contenimento dei costi: il personale complessivo è passato da 111.621 a 89.431 unità da fine 2000 a fine 2008 (da 56.771 a 42.718 dal 2001 al 2008 nel caso di Trenitalia).

5.3 *Il mercato regionale*³⁸

Dal lato della domanda il mercato regionale si caratterizza, rispetto ad altri servizi di trasporto, per una qualità del servizio bassa e in diminuzione e una quota di utilizzatori in lieve calo ma con percorrenze medie e numero di viaggiatori in aumento.

La domanda di mobilità non è mutata in maniera rilevante nel decennio in corso, almeno in termini quantitativi. I dati Isfort indicano che la quota di popolazione che ha effettuato spostamenti in un giorno feriale tipo è passata dall'84,6% nel 2000 all'83,5% nel 2009 (con un massimo dell'85,8% nel Nord Est). Le abitudini di mobilità hanno invece avuto alcuni mutamenti fra il 2000 e il 2008: a fronte di un numero di spostamenti medi giornalieri piuttosto costante (circa 3) si è avuto un calo dell'incidenza di quelli per motivi di studio o di lavoro, caratterizzati da orari standard e da picchi prevedibili (da quasi il 44% nel 2004 al 35,9%), un aumento della durata media e della distanza media dello spostamento (da 9,6 a 12,7 km dal 2000 al 2008) e uno più significativo della velocità media (Tavola 5).

Il servizio ferroviario regionale soddisfa una domanda di trasporto pendolare legata prevalentemente a motivi di studio o di lavoro. L'utilizzo giornaliero del treno regionale resta appannaggio del 4,6% della popolazione, una quota che ha oscillato lievemente negli ultimi anni ed è in diminuzione rispetto al 1995 (4,8% nel 2001; Tavola 6). Sul territorio nazionale l'incidenza dei pendolari ferroviari è particolarmente elevata nel Nord Ovest e nel Centro (5,6% e 5%, rispettivamente); nel Mezzogiorno la ferrovia locale ha un utilizzo significativo solo in Campania (7%). In Sicilia, Umbria e Basilicata si rilevano le quote di utilizzo più basse (fra l'1,6% e il 2,1%). In termini assoluti si tratta di quasi un milione e seicentomila persone di 14 anni e più che utilizzano il treno tutti i giorni o qualche volta a settimana. Il 64,1% dei pendolari ferroviari è concentrato in 5 regioni (Lombardia, Campania, Lazio, Piemonte e Veneto).

L'aspetto più eclatante relativo all'utenza è il forte peggioramento delle valutazioni sulla qualità del servizio rispetto al 1995; la quota di utenti che si dichiaravano soddisfatti è scesa da quasi il 60% al 47,2% nel 2009. Fra il 2007 e il 2009

³⁸ In quanto segue non si commentano i dati relativi ai due operatori attivi nel trasporto passeggeri che operano esclusivamente in regime di licenza (la SAD, che opera nella provincia autonoma di Bolzano e SERFER, che effettua un trasporto *charter* passeggeri fra Roma e Civitavecchia, ma si occupa prevalentemente di progettazione, manutenzione e cargo ed è integralmente di proprietà di Trenitalia).

si rileva un modesto miglioramento nella media, più accentuato nel Nord Est, a fronte di una sostanziale stabilità della qualità percepita del servizio in Sicilia e Campania e di un progressivo peggioramento in Calabria (Tavola 6)³⁹. Questi dati si riferiscono tuttavia all'intero servizio ferroviario nazionale e regionale e non consentono pertanto di valutare la qualità dello specifico segmento regionale. Un'indicazione su quest'ultimo aspetto si può derivare dall'indagine Isfort sulla mobilità, che segnala per il 2009 una valutazione media del trasporto ferroviario locale di 5,91 punti (Tavola 7), valore minimo e inferiore anche a quello dell'autobus extraurbano (6,63; la moto ha un punteggio di 8,26 e l'auto di 7,89). La valutazione dei treni locali è migliore sia nel Nord Est sia nel Centro (6,18 e 6,06, rispettivamente). L'indicatore di puntualità del servizio regionale Trenitalia segnala che, fra il 2000 e il 2008, la percentuale di treni arrivati fra 0 e 5' di ritardo è aumentata dall'85,5% al 90,3% (Tavola 9).

Dal lato dell'offerta, sul mercato regionale opera Trenitalia, in molte regioni come monopolista e in altre affiancato da alcuni altri gestori locali ("ferrovie regionali") generalmente controllati da Regioni ed enti locali. Trenitalia ha un'ampia quota di mercato sia in termini di viaggiatori sia di viaggiatori-km; l'insieme degli altri gestori regionali ha tuttavia registrato una maggiore crescita.

Il segmento di mercato regionale rappresenta, in termini di viaggiatori, la gran parte del trasporto ferroviario, con una quota di passeggeri del 91,2%, pari nel 2008 a 725,3 milioni. Nel 2008 le ferrovie regionali trasportavano il 29,2% dei viaggiatori locali (dal 28,2% nel 2000). I dati del Conto nazionale delle infrastrutture e dei trasporti (CNIT) mostrano un costante calo dei passeggeri delle ferrovie regionali da circa 192 milioni nel 1980 fino a un minimo di 145 milioni nel 1993. Dal 2000 si è avuta una ripresa significativa, con un aumento dei passeggeri del 31,9% fino al 2008, sfiorando i 212 milioni di passeggeri nel 2008 (Tavola 8). Nel caso dei treni locali gestiti da Trenitalia (tipicamente presente su reti più lunghe e meglio interconnesse) l'aumento dei passeggeri dal 2000 al 2008 è stato più modesto (25,8%), ma con una significativa ripresa nel 2009.

Se si considerano anche le percorrenze relative, la quota di viaggiatori-km delle ferrovie regionali sul totale del trasporto regionale si dimezza al 14,3%, a causa della minore percorrenza media rispetto a Trenitalia. Tuttavia fra il 2000 e il

³⁹ Questo dato si riferisce al 5,9% degli studenti e al 2,9% dei lavoratori che escono di casa per svolgere le loro attività; le percentuali salgono all'8,5% e 3,4% se si considerano solo coloro che utilizzano mezzi di trasporto a motore.

2008 le ferrovie regionali hanno avuto una crescita notevolmente superiore (52,2% contro 13,2%, rispettivamente). Per le ferrovie regionali il percorso medio di un viaggiatore è aumentato da 15,2 km a 17,5 km nel 2008; nello stesso periodo si è registrata una diminuzione per la percorrenza media degli utenti locali di Trenitalia, da 48 km a 43 km (-10,4%).

Anche utilizzando l'indicatore di offerta in termini di treni-km, la quota di trasporto regionale di Trenitalia risulta sensibilmente superiore (quella delle ferrovie regionali è del 17,1%), ma segna, rispetto al 2000, una crescita molto più modesta rispetto a quella delle ferrovie regionali (10,8% contro 19,6%). A fronte di un *load factor* sostanzialmente costante per Trenitalia fra il 2000 e il 2008 (circa 30%), quello delle ferrovie regionali si è ridotto drasticamente dal 32,1% al 21,7% (Tavola 8), più che dimezzandosi nel Mezzogiorno.

Lo *stock* di materiale rotabile delle ferrovie regionali è lievemente aumentato fra il 1990 e il 2000; negli anni successivi si è registrato un modesto aumento dei mezzi di trazione e un calo del numero di carrozze. Per Trenitalia fra il 2000 e il 2008 si rileva un notevole calo di tutto il materiale rotabile (incluso quello utilizzato sulle tratte nazionali). Il contestuale aumento dei treni-km sia per le ferrovie regionali sia per il trasporto regionale di Trenitalia sembra indicare un aumento della frequenza di utilizzo del materiale rotabile. Anche in questo caso diverse fonti ne segnalano anche un'elevata età media, nonostante una politica di rinnovo avviata all'inizio degli anni 2000, che riguarderebbe in particolare il trasporto regionale. Nel 2007, secondo dati del MEF, l'età media del parco rotabile era superiore a 20 anni (con picchi di 30-40 anni in alcune regioni; Coordinamento dei Pendolari *et al.*, 2008). Nonostante la riduzione dei costi di acquisto del materiale rotabile collegata all'utilizzo delle gare comunitarie a partire dagli anni novanta (circa il 30%; Gardina, 2006), le notevoli differenze nelle reti regionali (scartamento, alimentazione e altri vincoli tecnici legati alle singole infrastrutture) limitano la possibilità di forme di acquisti in gruppo per ottenere i maggiori sconti applicati su lotti di dimensione elevata.

Le informazioni sui risultati di gestione delle imprese ferroviarie sono scarse. Per le ferrovie regionali i dati del CNIT mostrano come il rapporto fra ricavi (al netto delle sovvenzioni in conto esercizio) e costi della produzione sia mediamente al di sotto della soglia indicata dalla normativa (35%), con una media del 24,5% (Tavola 9). Fra il 2000 e il 2008, dopo un iniziale peggioramento nel Meridione, dove si sono toccati valori minimi intorno al 14%, a partire dal 2003 si è avuto un pro-

gressivo recupero. Nelle altre aree il miglioramento è risultato più debole e parziale⁴⁰. I ricavi netti per viaggiatori-km si aggiravano nel 2008 intorno ai 12 centesimi di euro (dai 9% del 1998), con minimi di 4-5 centesimi nel Centro Nord.

Il trasporto locale ha un'incidenza significativa sul bilancio di Trenitalia: la quota dei ricavi del trasporto passeggeri regionale (inclusa la parte relativa alle RSS) sul totale del trasporto ferroviario passeggeri è passata dal 54,9% al 58% fra il 2001 e il 2008 (Tavola 9). Questo dato è significativo se si considera che il mercato regionale rappresenta l'88% dei viaggiatori Trenitalia, il 69,8% dei suoi treni-km e il 60,3% dei suoi posti-km complessivamente offerti.

I corrispettivi per CS con le Regioni sono passati da 1.214 mln di euro nel 2001 a 1.756 nel 2009; i contributi dallo Stato per il servizio svolto nelle RSS da 218 a 183 mln di euro, soprattutto per effetto della sottoscrizione diretta negli anni recenti da parte di alcuni enti dei CS (Friuli-Venezia Giulia e Province autonome di Trento e Bolzano). I dati di Legambiente (2011) mostrano una elevata variabilità dei corrispettivi dei CS in rapporto ai treni-km annui; i confronti interregionali basati sui treni-km erogati sono tuttavia poco significativi, dato l'utilizzo nei contratti con Trenitalia di parametri che tengono anche conto di una serie di altre caratteristiche del servizio richiesto.

Le valutazioni sulla congruità di questi corrispettivi sono discordi. Gli incassi complessivi (inclusi quelli tariffari) sarebbero passati da 9,2 a 11,0 centesimi di euro a passeggero-km fra il 2001 e il 2008 fino a 11,8 nel periodo precedente ai nuovi CS. Secondo FS, in Francia e Germania i corrispettivi per il trasporto ferroviario regionale sarebbero più elevati: 22,1 centesimi di euro a passeggeri-km in Francia e 19,2 in Germania; questo indice, usato largamente da Trenitalia, rischia di essere fuorviante non solo per motivi di confrontabilità legati alle differenze nel potere d'acquisto, ma anche perché un corrispettivo per passeggeri-km più alto significa, di norma, che è riferito a treni con meno viaggiatori ed è pertanto un sussidio meno efficiente.

Secondo la Corte dei conti (2009) il ricavo medio unitario sarebbe mediamente 12,5 centesimi di euro a passeggeri-km, inferiore al prezzo di equilibrio individuato in 14,5 centesimi di euro. Arenaways, l'impresa ferroviaria che aveva avviato il servizio sulla tratta Torino-Milano, ha dichiarato che, applicando i contratti collettivi di lavoro e senza contributi pubblici, è in grado di coprire i costi con un *load factor* del 30%⁴¹.

⁴⁰ I dati includono anche il trasporto merci da parte dei gestori locali e risentono a metà degli anni 2009 di operazioni di scorporo.

⁴¹ «L'ex capostazione che sfida le FS», *Il Sole 24Ore*, 26 settembre 2010, p. 16.

6. - Il trasporto ferroviario passeggeri in alcuni paesi europei

Nonostante l'organizzazione del settore sia stata oggetto di tre ondate di direttive europee, permangono notevoli differenze fra i paesi della UE. La proprietà delle rete è pubblica nei principali paesi (Italia, Regno Unito, Germania, Francia, Austria, Paesi Bassi e Svezia), ma la separazione fra gestore della rete e del servizio è completa solo negli ultimi due e nel Regno Unito, mentre Italia, Austria e Germania accorpano le due attività sotto una stessa *holding*. Solo il Regno Unito ha sperimentato un breve periodo di privatizzazione della rete. Fra i paesi extraeuropei, negli Stati Uniti e in Giappone la rete è di proprietà delle imprese ferroviarie.

Sulla lunga distanza l'*incumbent* statale detiene il 90-100% del mercato, con l'eccezione di Olanda (dove operano due principali gestori), Svezia e Regno Unito (dove lo Stato non ha partecipazioni nei gestori)⁴².

Il Regno Unito ha avviato la ristrutturazione del settore nel 1993, con una frammentazione della British Rail in un centinaio di imprese indipendenti e separate verticalmente. Mentre per le merci è stata introdotta la concorrenza nel mercato, per quasi tutto il trasporto passeggeri sono state introdotte gare per le tracce basate su *standard* di qualità e livello di sussidi richiesto, con una durata dei contratti fra i 7 e i 15 anni (Affuso, 2003). Sono attive 3 ROSCOs che detengono il 97% del mercato di noleggio del materiale rotabile per il servizio passeggeri alle imprese ferroviarie. Inizialmente la proprietà della rete e di gran parte delle stazioni è stata assegnata a un'impresa privata ma, dopo una serie di incidenti alla fine degli anni novanta, collegati sia alla scarsa manutenzione della rete sia al mancato coordinamento fra gli operatori del settore, essa è stata acquisita nel 2002 da Network Rail (impresa a capitale pubblico), che ha avviato un intenso programma di ristrutturazione e di miglioramento della rete e della tecnologia collegata. Il settore è in espansione: sono previsti ulteriori rilevanti investimenti per migliorare la sicurezza, il livello tecnologico e la capacità di trasporto della rete. Attualmente sono presenti circa 19 imprese di trasporto passeggeri, e nel 2009 il numero dei viaggiatori ha raggiunto il picco massimo in 60 anni. Dal 1995 sono state aperte 27 nuove linee (per un totale di 199 miglia di binari) e 68 stazioni. Nel giugno 2009 l'associazione delle imprese ferroviarie ha pubblicato un rapporto in cui identificava come redditizio un progetto di riapertura o costruzione di 40 stazioni e di alcune linee per i servizi passeggeri. Si stanno valutando alcuni progetti per l'alta velocità (ad esempio la tratta Londra-Edimburgo),

⁴² Alcune imprese ferroviarie britanniche sono partecipate da imprese di proprietà di altri Stati.

sulla spinta della congestione dei principali aeroporti. Fra i problemi aperti vi è il ruolo delle ROSCOs, che secondo la *Competition Commission* tenderebbe a distorcere la concorrenza. Sono stati identificati alcuni aspetti problematici:

- i) le ROSCOs non offrono un'ampia varietà di tipologie di materiale rotabile e ciò limita le possibilità di scelta delle imprese ferroviarie;
- ii) la necessità per le imprese ferroviarie di avviare accordi preliminari di *leasing* con le ROSCOs prima di fare un'offerta per il servizio (in modo da avere la disponibilità del materiale rotabile) ne riduce il potere contrattuale e le sottopone a rischi di comportamenti opportunistici delle ROSCOs);
- iii) le ROSCOs hanno scarsi incentivi per farsi concorrenza, anche a causa del loro codice di condotta in cui si impegnano a offrire le stesse condizioni a tutte le imprese ferroviarie.

La *Competition Commission* ha raccomandato fra l'altro che le concessioni durino almeno 12-15 anni, per aumentare le opzioni a disposizione delle imprese ferroviarie.

In Francia il trasporto regionale con obblighi di servizio è affidato direttamente dalle Regioni alla SNCF con CS quinquennali (con possibilità di estensione a 10 anni; Corte dei conti, 2009) e i corrispettivi sono indicizzati sia all'inflazione programmata sia alla crescita del PIL. In Germania i Länder assegnano con gara i servizi nelle aree rurali (meno redditizi) e a DB quelli più redditizi (aree metropolitane); i corrispettivi annui sono indicizzati a una crescita media dell'1,5% e i CS hanno durata 10-12 anni. La rete tedesca risulta congestionata in alcune aree e vi sono problemi di allocazione delle tracce fra MLP, trasporto locale e trasporto merci. In Spagna i servizi sono affidati direttamente a RENFE-Operadora dal ministero dei trasporti; il CS pluriennale definisce il corrispettivo sia a copertura delle spese correnti sia degli investimenti in nuovo materiale rotabile. Sono stati avviati ingenti investimenti in AV; dopo l'apertura della tratta ferroviaria Barcellona-Madrid all'inizio del 2008, che ha sottratto significative quote di mercato al trasporto aereo, è in progetto la costruzione di altri 9.000 km di rete nel prossimo decennio, con l'obiettivo di creare la rete AV più estesa d'Europa, con il 90% dei potenziali utenti che vivono a meno di 50 km da una stazione ferroviaria⁴³.

Le iniziative di costruzione di nuove ferrovie con finanziamenti almeno in parte privati sono limitate, ma stanno sviluppandosi quelle in *project financing*. Fra queste, l'Eurotunnel sotto la Manica è stato completato nel 1994 con un rilevante aggravio di costi (legato soprattutto a una non corretta allocazione dei rischi); il

⁴³ «A Surprising Conversion», *The Economist*, 30 dicembre 2008.

progetto ha generato il suo primo utile nel 2007 (la società è stata creata nel 1986). Una seconda iniziativa in finanza di progetto riguarda la realizzazione della prima tratta AV tra il confine belga ed Amsterdam HSL Zuid (98 km)⁴⁴. Un caso interessante riguarda il progetto per la linea ferroviaria ad alta velocità Perpignan-Figueras (44,4 km) in cui i rischi di costruzione, disponibilità e mercato sono stati allocati al privato⁴⁵. Il progetto per la ferrovia Stoccolma-aeroporto (42 km, costo € 700 milioni, concessione fino al 2040) ha ricevuto un contributo pubblico, non a fondo perduto, del 25%, a fronte dell'integrale allocazione del rischio ai privati. Infine nel 2001 è stata introdotta in Svizzera una "tassa sul traffico pesante commisurata alle prestazioni" (TTPCP) per i mezzi su gomma⁴⁶; la quota degli introiti spettante alla Confederazione è destinata alle ferrovie. Recentemente è stata affidata in Francia una nuova linea AV (Bordeaux-Tours, ed eventuali future estensioni verso la Spagna) con la formula del partenariato pubblico privato, in cui il concessionario è un'impresa di costruzione (Vinci) che coprirà il 50% del finanziamento e si ripagherà, per l'intera durata della concessione (35 anni a partire dal 2016), con i ricavi da pedaggio pagati dalle società ferroviarie che utilizzeranno la tratta per il passaggio dei loro treni (303 km, costo presunto € 7,4 miliardi, € 3,6 miliardi finanziati dallo stato tramite Rff, ricavi annui previsti € 200-250 milioni).

Se si considerano i primi 4 paesi indicati oltre alla Spagna, la Germania presenta la rete ferroviaria più estesa, anche rispetto alla sua superficie (Tavola 10). La lunghezza totale della rete si è ridotta ovunque rispetto al 1970; in Italia e Spagna si è avuto un modesto aumento rispetto al 1990 legato soprattutto allo sviluppo dell'AV. Nel 2008 la densità della rete italiana era simile a quella della Francia e del Regno Unito e quasi la metà di quella tedesca⁴⁷. Nel 2008 Francia

⁴⁴ Si tratta di una concessione trentennale (2001-2031, di cui 5 anni per la costruzione ed i restanti per la gestione) ad una società di scopo costituita da investitori istituzionali (tra cui lo Stato olandese) ed operatori industriali. Il costo complessivo è di circa 3,9 miliardi di € ed è finanziato per il 28% da privati, che ricevono un canone sulla base di alcuni indicatori di *performance*.

⁴⁵ Oltre al 57% fra contributi pubblici e comunitari, l'opera è stata finanziata da privati e da prestiti bancari, per un costo complessivo di circa 1,1 mld €. La linea è stata completata, con l'eccezione di una stazione, nel febbraio 2009 e la concessione ha una durata di 50 anni.

⁴⁶ La tassa ha ridotto significativamente gli autocarri transitati attraverso le Alpi, aumentato il coefficiente di carico e accelerato i tempi di rinnovo dei mezzi. Le ferrovie rappresentano circa i due terzi del trasporto merci transalpino. Tasse simili sono state introdotte successivamente in altri paesi europei.

⁴⁷ Nel confronto con i principali paesi extraeuropei, la rete della UE-27 è lunga 215,6 migliaia di km, contro i 203,6 degli USA e i 27,6 del Giappone. Le tre reti trasportano, rispettivamente, 394,6, 23,7 e 396 milioni di passeggeri-km (dati Eurostat 2006/2007).

e Germania trasportavano il numero più elevato di passeggeri-km; in termini di dinamica, il Regno Unito e la Spagna sono i paesi in cui si sono avuti i maggiori aumenti di utenza (oltre il 50% fra il 1990 e il 2008), mentre l'Italia registra il risultato peggiore.

7. - Alcune valutazioni

Federalismo ferroviario e liberalizzazione del mercato nazionale. - Il trasferimento delle competenze alle Regioni è stato lento, incompleto e spesso solo formale. Alle Regioni sono state trasferite le piccole imprese ferroviarie locali, spesso in condizioni economiche dissestate, e reti secondarie, spesso scollegate fra loro, di inferiore qualità rispetto alla rete nazionale e non interoperabili. Nella scelta del gestore del servizio, in molti casi le difficoltà di reperire e mettere a disposizione il materiale rotabile, una certa inerzia o la carenza delle competenze necessarie per realizzare bandi di gara migliorativi in termini di servizio e di costi, hanno indotto le Regioni a scegliere di affidare il servizio a Trenitalia, talvolta in cogestione con imprese pubbliche locali, rafforzandone il potere di mercato.

La natura pubblica di Trenitalia (e, per altri versi, l'essere parte dello stesso gruppo di RFI) le attribuisce vantaggi di posizione, oltre a quelli classici di un *incumbent*, che non solo discriminano i potenziali entranti sui mercati regionali, ma limitano fortemente la capacità contrattuale delle Regioni. La scelta di introdurre per legge sussidi aggiuntivi per chi sceglie come gestore Trenitalia è anti-concorrenziale e riduce le opzioni per le Regioni, anche in termini di selezione dell'operatore più efficiente. Inoltre nell'attuale situazione di federalismo incompleto, i fondi trasferiti alle Regioni sono frutto più di una contrattazione politica che di valutazioni sulle effettive necessità nei diversi territori.

I dati mostrano un significativo divario territoriale nel grado di utilizzo degli strumenti messi a disposizione dal D.lgs. n. 422 del 1997. Le Regioni a statuto ordinario del Nord hanno mostrato un maggiore attivismo, sia introducendo le gare per la gestione del servizio, sia includendo condizioni specifiche di monitoraggio nella qualità del servizio e penalità collegate. Sinora la Regione Lombardia è stata l'unica a realizzare soluzioni per superare la barriera all'entrata costituita dalla disponibilità del materiale rotabile (15 treni forniti gratuitamente nel caso della gara della Linea S15). Inoltre, laddove le gare sono state almeno bandite, si sono avute manifestazioni di interesse da parte di alcune imprese estere, segnalando che il mercato regionale italiano, almeno in alcuni suoi segmenti, risulta appetibile.

Ulteriori vantaggi di posizione derivano a Trenitalia dal suo monopolio sul mercato nazionale. La liberalizzazione del trasporto nazionale passeggeri è avvenuta di recente e l'adeguamento delle norme italiane è stato, soprattutto all'inizio, lento e parziale e volto a tutelare l'*incumbent*. I notevoli investimenti pubblici sostenuti per l'AV e la scarsa concorrenza intermodale del trasporto aereo hanno consentito a Trenitalia di aumentare tariffe e passeggeri sulle tratte nazionali più redditizie (aumento che ha raggiunto ben il 90% sulla Milano-Roma tra il 2005 e il 2010). La contestuale attuazione di una politica di contenimento dei costi (in particolare attraverso la riduzione del personale) ha migliorato i conti economici di Trenitalia, rafforzando ulteriormente la posizione contrattuale nei confronti delle Regioni.

Si è parallelamente ridotta l'offerta dei servizi pubblici nazionali, che continuano a essere affidati senza nessuna procedura di selezione a Trenitalia. Non sono stati resi noti in modo esplicito i fattori alla base della scelta del nuovo perimetro dei servizi pubblici nazionali; i relativi contributi restano significativi e, anche per la presenza di un monopolio, non basati su valutazioni trasparenti dei costi.

Molti sono dunque i fattori che costringono le ferrovie regionali a viaggiare su un binario stretto. L'attuazione di un effettivo federalismo ferroviario e di un servizio regionale più efficiente e di migliore qualità richiede in primo luogo un passo indietro dello Stato verso una maggiore "neutralità normativa", che non privilegi il proprio operatore (peraltro su mercati regolati da altri operatori pubblici quali le Regioni), con la creazione di un'Autorità indipendente di settore.⁴⁸ Un percorso alternativo potrebbe essere la privatizzazione di Trenitalia; in questo caso tuttavia bisognerebbe evitare di attribuire una posizione dominante *ex ante* al nuovo gestore e allo stesso tempo di frazionare eccessivamente l'*ex* monopolista.

In secondo luogo è necessario che le Regioni acquisiscano maggiori competenze nell'attività di scelta del gestore e nella definizione dei contenuti dei CS. In questo ambito potrebbero risultare particolarmente utili forme di collaborazione e consulenza fra le Regioni e la diffusione di informazioni sulle *best practices* attuate nel settore. Forme analoghe di collaborazione potrebbero essere attuate anche nel campo del materiale rotabile, ad esempio nella creazione di gruppi d'acquisto per effettuare ordini di maggiore importo (soprattutto quando la domanda sia qualitativamente omogenea).

⁴⁸ L'Autorità di regolazione dei Trasporti, istituita all'inizio del 2012 non era ancora operativa a metà luglio.

Domanda effettiva e potenziale. - La quota di trasporto ferroviario regionale rispetto al totale del trasporto locale, oltre a essere molto contenuta, risulta in lieve diminuzione negli ultimi 15 anni, evidenziando l'insuccesso delle politiche di incentivo del trasporto pubblico su rotaia nonostante la crescente congestione stradale. La domanda potenziale resta ancora elevata, ma in buona parte insoddisfatta, e catturata da altre modalità di trasporto, a causa della bassa qualità del servizio (in termini di *comfort* e puntualità, ad esempio). Le valutazioni sulla qualità del servizio sono le più basse fra tutte le modalità di trasporto e inferiori perfino a quelle sul trasporto extraurbano su gomma, oltre che peggiorate nel tempo.

Politiche di stimolo del trasporto ferroviario dovrebbero tener conto di queste caratteristiche della domanda, migliorando gli aspetti qualitativi del servizio e quelli accessori, come quelli che rendono agevole il cambio modale (offerta di aree di sosta nei pressi delle stazioni e di servizi su gomma collegati con orari compatibili con quelli dei treni). Date le caratteristiche della funzione di produzione del servizio, che lo rendono economico solo in aree dense, sarebbe opportuno anche concentrare maggiormente gli aumenti di offerta sulle linee a maggiore domanda. Queste politiche richiedono una parallela attività di coordinamento con la pianificazione delle nuove aree abitative e industriali.

Nelle aree a maggiore densità di popolazione alcune stime, oltre alla presenza di potenziali entranti, indicherebbero una significativa redditività del servizio. In questi casi sarebbe forse possibile eliminare o ridurre notevolmente i sussidi, soprattutto se si stimolasse la domanda aumentando la qualità intesa nel suo senso più ampio.

Dimensione dell'infrastruttura. - Se si escludono gli ampliamenti relativi all'alta velocità, la dimensione della rete (in termini di lunghezza) è rimasta pressoché invariata da decenni. Tale aspetto tuttavia non sembra condizionare il livello di servizio, dato il medio/basso utilizzo dell'infrastruttura esistente, nel confronto internazionale. A livello locale, più che linee nuove, sarebbe utile potenziare l'uso di quelle esistenti soprattutto nelle cinture dei maggiori centri urbani, e migliorare la qualità dell'esistente (raddoppio binari, tecnologia, organizzazione). Lo sviluppo dell'AV sta ponendo problemi di accesso alle stazioni e di allocazione delle tracce alle imprese ferroviarie regionali e ai concorrenti di Trenitalia, che potrebbe trovare una soluzione nella previsione di una maggiore terzietà del regolatore.

Altre esperienze europee (ad esempio il Regno Unito) mostrano che, in presenza di vincoli di bilancio stringenti, spesso gli investimenti per l'AV sono alternativi a quelli per la rete locale/destinata al trasporto dei pendolari. Finora questo

non sembra essere avvenuto in Italia. Tuttavia i vincoli di finanza pubblica richiedono che ai diversi livelli di governo sia fatta una scelta esplicita del segmento cui dare la priorità. Nel trasporto a lunga percorrenza il treno sta diventando molto più costoso del trasporto su gomma e sembra destinato sempre più ai viaggiatori *business* e sempre meno alle famiglie. Nella scelta del tipo di infrastruttura in cui investire sarebbe quindi opportuna un'attenta valutazione dei costi e dei benefici del potenziamento dell'AV, qualora la costruzione della rete fosse sussidiata anche in futuro. La partecipazione dei privati al finanziamento di nuove iniziative potrebbe essere una valida alternativa ai sussidi pubblici qualora sia possibile ripartire i rischi del progetto in maniera efficiente e allo stesso tempo garantire un'adeguata redditività degli investimenti.

TAV. 1

GESTORI DEL SERVIZIO FERROVIARIO LOCALE NELLE REGIONI ITALIANE
OPERATIVI SULLA BASE DI UN CONTRATTO DI SERVIZIO
(SITUAZIONE A LUGLIO 2012)

Regione o macroarea	Società ^(a)	Annotazioni
Nord Ovest		
Piemonte	<i>Gruppo Torinese Trasporti</i> , Trenitalia	Gara per l'affidam. del servizio su 3 lotti annullata a fine 2010 (Tavola 2)
Val d'Aosta	Trenitalia	Trasf. comp. e funzioni alla Regione nell'ottobre 2010
Lombardia	1) ATI composto da Trenitalia, Ferrovie Nord Milano e ATM Trenitalia-LeNORD ^(b)	Parzialm. affidato con gara (tratta 1); Tavola 2); 2009: accordo tra Trenitalia e LeNORD per la costituzione della nuova impresa ferroviaria Trenitalia-LeNord srl che gestisce le restanti tratte (con 2 distinti CS). A maggio 2011 Trenitalia-LeNORD diventa Trenord (partecipata in modo paritario)
Liguria	AMT, Trenitalia	Linea Genova-Casella affidata con gara a AMT; subentro all'ex gestore Ferrovia Genova-Casella dall'1 aprile 2010 (Tavola 2)
Nord Est		
Trentino-Alto Adige	<i>Sad Trasporto Locale</i> , Trentino trasporti esercizio, Trenitalia	Trentino trasporti esercizio: affid. in house
Veneto	<i>Sistemi Territoriali</i> , Trenitalia	<i>Sistemi Territoriali</i> gestisce la linea Adria-Mestre; dal 01/01/2012 i servizi affidati con gara (Tavola 2), sono confluiti nel CS con Trenitalia
Friuli-Venezia Giulia	<i>Società Ferrovie Udine Cividale</i> , Trenitalia	
Emilia-Romagna	Consorzio Trasporti integrati (<i>FER</i> e Trenitalia)	Interamente affidato con gara (Tavola 2)
Centro		
Toscana	<i>Trasporto Ferroviario Toscano</i> ^(b) , Trenitalia	
Umbria	<i>Ferrovia Centrale Umbra</i> , Trenitalia	
Marche	Trenitalia	
Lazio	ATAC, Trenitalia	

Sud e Isole		
Abruzzo	<i>Ferr. Adriatico Sangritana</i> , Trenitalia	
Molise	Trenitalia	
Campania	Ferr. Circumvesuviana, <i>Metrocampania Nordest</i> , SEPSA, Trenitalia	
Puglia	<i>Ferrovie del Gargano</i> , <i>Soc. Ferrotramviaria</i> , <i>Ferr. del Sud Est</i> , <i>Ferrovie Appulo Lucane</i> , Trenitalia	Ferrovie del Gargano: fam. Germano; Ferrotramviaria: azionisti privati; Ferr. del Sud Est e Ferrovie Appulo Lucane: MIT
Basilicata	Ferr. Appulo Lucane, Trenitalia	Ferrovie Appulo Lucane: MIT
Calabria	Ferr. della Calabria, Trenitalia	
Sicilia	GCG Ferr. Circumetnea, Trenitalia	Trasf. competenze e funzioni in materia non ancora attuato. La Circumetnea è in gestione commis. governativa (MIT)
Sardegna	<i>ARST</i> , Trenitalia	In corso di definizione quota rete RFI da trasferire alla Regione e linee su rete RFI da considerarsi regionali

Fonte: MINISTERO DELLE INFRASTRUTTURE E DEI TRASPORTI, *Conto nazionale delle infrastrutture e dei trasporti*; LEGAMBIENTE, *Pendolaria*, 2009.

⁽⁶⁾ Dove non diversamente specificato, i gestori diversi da Trenitalia sono sotto il controllo degli Enti locali. Sono in corsivo le reti “non isolate”, cioè che sono interconnesse alla rete RFI ai sensi del decreto ministeriale 5 agosto 2005, n. 28T. ⁽⁶⁾ Impresa ferroviaria che ha attuato la separazione societaria tra rete e servizio (cioè in cui è presente anche uno specifico gestore dell’infrastruttura ferroviaria regionale, analogo a RFI).

TAV. 2

GARE PER IL TRASPORTO FERROVIARIO REGIONALE (SITUAZIONE A LUGLIO 2012;
VALORI ASSOLUTI E IMPORTI IN MILIONI DI EURO)

Regione	Data bando	Durata contratto (anni) e percorrenza minima richiesta (treni-km annui)	Tratte e base d'asta (mln €, netto IVA)	Materiale rotabile	Impresa aggiudicataria/partecipanti
Piemonte	Bando: 11.11.2009 scad. domande: 15.2.2010 offerte: entro 90gg 1.6.2010: sospensione procedure Nuovo termine present. offerte: 31.10.2010 1.12.2010: annullamento gara	Durata : 6+6 Treni-km/anno minimi: 1) 4.022.310 2) 5.438.229 3) 2.108.393	1) Lotto 1 (N/E): 43,4/anno 2) Lotto 2 (S/E): 60,3/anno 3) Lotto 3 (To-Mi): 1/anno (da pres. con un'altra)	a) a carico aggiudicatario; b) 3 anni per reperimento; c) obbligo di cessione al gest. subentrante in caso di mancato rinn. se finanz. a qualsiasi titolo dalla Regione (LR 1/2000 e agg.ti), salvo restituz. quota contrib. erogati commisurata al periodo di non utilizzo; d) diritto di prelaz. del gestore subentr. su quello privo di vincolo di dest. d'uso	Trenitalia- LeNORD Arriva Veolia DB Regio Ferr. Fed. Svizzere (le ultime tre hanno successivamente ritirato la manifestazione di interesse)
Lombardia	bando: 2004 1) aggiudicato: 2005	1) 9; 1.660.000 2) 9; 1.040.000 3) 7; 1.040.000 + 2.824 mila bus-km	1) Pioltello-MI Passante-Gallarate-VA (S5): 69,3 2) MI-Molteno-LC e CO.LC: 93,6 3) BS-Iseo-Edolo: 84,8	1) messo a disp. da Regione in uso gratuito 2) a carico aggiudicatario (valutato all'aggiud.ne) 3) messo a disp. da Regione (proprietaria) e possibilità di integrazione a carico agg. (valutato all'aggiud.ne)	1) ATI composto da Trenitalia, Ferrovie Nord Milano e ATM 2), 3) capitolato non pubblicato, procedura di gara interrotta

Liguria	1) 2004 2) bando del 6.8.2009 agg.to il 21.1.2010 (decorrenza 16.4.2010)	1) 9; 6.900.000 2) 9+6;168.499	1) Tutto il servizio: 589,8 (per i 9 anni) 2) Servizio su Linea ex concessa Genova-Casella: 15 ,1 per 9 anni e compreso gestione e manutenzione ordinaria dell'impianto ferroviario	1) A carico aggiudicatario con impegno di progressivo rinnovo della flotta e obbligo di cessione alla Regione al termine dell'affidamento 2) Concesso in uso all'aggiudicatario dalla Regione (propr.) per l'intera dur.del contratto	1) Esclusione offerte pervenute: Trenitalia (condizionata); LeNord (non conforme al capitolato). 2) Unico offerente: AMT (gestore trasp. su gomma)
Veneto	bando: 2003 agg.to: 2004	6; 10.555.760	Servizi "Lotto 2" (prov. di VI, VR, BL, RO): 422,5	A carico aggiudicatario con impegno di progressivo rinnovo della flotta	Unico offerente e aggiudicatario: Trenitalia+Sistemi territoriali
Emilia-Romagna	bando: aprile 2005 agg.to: 31-3-2008	3+3; 15.666.350	Tutto il servizio: 90,8/anno	Apportato dall'aggiudicatario a proprio carico	Unico offerente: Consorzio trasporti integrati (Trenitalia, FER, ATCM MO, ACT RE, ATC BO)

Fonte: CAMBINI C. (2009); CORTE DEI CONTI (2009) e siti internet delle Regioni.

TAV. 3

ESTENSIONE DELLA RETE FERROVIARIA ITALIANA (KM)

Tipo di rete	1990	1995	1997	2000	2007	2008
Nazionale (RFI)	16.066	16.005	16.030	15.974	16.335	16.529
di cui: <i>elettrificata</i>	9.512	10.205	10.358	10.714	11.531	11.727
<i>a doppio binario</i>	5.771	6.023	6.106	6.156	7.050	7.306
Regionale	3.510	3.527	3.392	3.443	3.700	3.655
di cui: <i>elettrificata</i>	1.281	1.321	1.300	1.165	1.372	1.335
<i>a doppio binario</i>	188	230	218	217	442	429
Totale	19.576	19.532	19.422	19.417	20.035	20.184

Fonte: MINISTERO DELLE INFRASTRUTTURE E DEI TRASPORTI, *Conto nazionale delle infrastrutture e dei trasporti*.

TAV. 4

ESTENSIONE DELLA RETE FERROVIARIA PER REGIONE

(KM e %)

Regione o macroarea	Rete regionale (2000) ^(a)		Rete regionale (2009) ^(a)		Estensione complessiva della rete	
	km	% elettrificata	km	% elettrificata	2000 ^(b)	2009 ^(b)
<i>Nord Ovest</i>	465	24,7	457	71,8	4.465	4.601
Piemonte	121	62,0	111	86,0	1.962	1.971
Valle d'Aosta	0	0,0	0	0,0	118	81
Lombardia	320	61,0	321	67,0	1.837	2.020
Liguria	24	100,0	25	100,0	548	529
<i>Nord Est</i>	478	20,9	440	35,2	3.593	3.769
Trentino-Alto Adige	56	100,0	65	100,0	421	421
Veneto	117	0,0	73	0,0	1.349	1.263
Friuli-Venezia Giulia					471	469
Emilia-Romagna	305	14,0	302	16,0	1.352	1.616
<i>Centro</i>	379	61,2	375	61,9	3.702	3.813
Toscana	84	100,0	84	100,0	1.482	1.526
Umbria	152	0,0	152	0,0	496	563
Marche	0	0,0	0	0,0	383	386
Lazio	143	100,0	139	100,0	1.341	1.338
<i>Sud e Isole</i>	2.139	33,0	2.238	24,7	7.687	8.076
Abruzzo	116	89,0	142	79,0	639	677
Molise	0	0,0	0	0,0	245	263
Campania	275	85,0	345	88,0	1.278	1.562
Puglia	806	18,0	807	18,0	1.657	1.652
Basilicata					322	318
Calabria	220	0,0	220	0,0	1.042	1.072
Sicilia	114	0,0	110	3,0	1.458	1.489
Sardegna	608	0,0	614	0,0	1.046	1.043
<i>Italia</i>	3.461	38,0	3.510	37,0	19.447	20.259

Fonte: MINISTERO DELLE INFRASTRUTTURE E DEI TRASPORTI, *Conto nazionale delle infrastrutture e dei trasporti*, Regione Lombardia e RFI. I dati nazionali possono non coincidere con la somma di quelli regionali a causa di arrotondamenti.

^(a) Dati al 1° gennaio. Esclusa la rete gestita da RFI. I dati relativi al Friuli-Venezia Giulia sono inclusi nel Veneto, quelli relativi alla Basilicata nella Puglia. ^(b) I dati relativi alla rete RFI si riferiscono al 28 maggio 2000 e al 12 dicembre 2010.

TAV. 5

INDICATORI DI MOBILITÀ GIORNALIERA (%)

Regione o macroarea	Numero medio di spostamenti giornalieri			% di spostamenti per lavoro/studio			Durata media dello spostamento (minuti)			Velocità media (km/h)		
	2000	2004	2008	2000	2004	2008	2000	2004	2008	2000	2004	2008
<i>Nord Ovest</i>	3,0	2,9	3,1	42,4	45,4	37,0	19,5	19,0	22,1	24,4	21,2	30,5
Piemonte	3,0	3,0	3,1	42,8	44,7	37,3	19,6	18,1	21,3	24,5	21,3	34,2
Valle d'Aosta	3,0	3,0	3,4	44,1	46,9	37,0	20,8	16,4	17,4	28,4	21,1	35,4
Lombardia	3,1	2,9	3,1	43,0	46,2	38,6	19,5	19,7	22,4	24,4	21,2	30,2
Liguria	3,0	2,8	2,9	38,2	42,9	26,7	19,6	17,6	22,7	23,5	20,7	21,6
<i>Nord Est</i>	3,1	3,0	3,2	42,1	44,0	38,2	18,2	17,5	21,3	26,3	22,2	32,0
Trentino-A. Adige	3,2	3,1	3,3	36,1	43,5	36,8	16,2	16,9	23,0	22,3	23,2	32,9
Veneto	3,1	3,0	3,3	43,2	44,7	36,8	18,1	17,0	20,5	26,9	22,2	32,1
Friuli-Ven. Giulia	3,1	3,1	3,0	39,6	42,1	35,7	17,4	17,1	22,9	27,3	23,0	31,5
Emilia-Romagna	3,2	2,9	3,1	43,0	43,9	41,1	19,2	18,3	21,3	26,3	21,8	31,7
<i>Centro</i>	3,1	3,0	3,1	40,4	46,0	36,6	20,5	20,6	22,7	25,9	22,2	31,5
Toscana	3,1	3,1	3,2	39,3	46,7	35,5	19,1	17,7	20,0	25,7	22,3	30,0
Umbria	3,0	3,0	3,3	39,0	40,1	38,7	17,0	15,3	19,1	26,8	22,1	58,0
Marche	3,4	3,0	3,3	39,7	47,4	36,5	16,4	17,7	17,5	25,8	23,9	36,0
Lazio	3,1	2,9	3,0	41,5	46,0	37,1	23,3	24,3	26,7	26,0	21,6	26,8
<i>Sud e Isole</i>	3,1	3,0	3,2	36,1	41,6	33,4	18,6	18,0	19,6	26,3	20,8	34,2
Abruzzo	3,3	2,9	3,2	35,3	39,2	31,5	18,0	17,2	18,2	27,0	23,7	36,4
Molise	3,1	3,0	3,2	40,8	43,3	34,0	16,6	16,0	18,5	25,8	21,6	37,3
Campania	3,0	3,0	3,2	38,2	42,2	36,0	20,2	19,3	21,6	30,5	19,9	28,9
Puglia	3,3	3,0	3,4	31,8	40,8	33,2	17,5	17,2	17,3	23,9	20,8	26,5
Basilicata	3,1	3,1	3,0	42,9	42,6	34,4	17,5	16,9	20,4	25,3	22,9	33,6
Calabria	3,1	3,0	3,1	36,4	42,5	31,7	17,5	16,2	19,8	25,3	20,8	49,0
Sicilia	3,2	3,0	3,1	36,1	41,6	34,2	18,6	18,0	19,0	24,0	20,3	33,4
Sardegna	3,3	3,1	3,4	37,4	41,7	26,8	18,3	18,6	21,3	26,1	21,6	53,0
<i>Italia</i>	3,1	3,0	3,2	39,8	43,9	35,9	19,1	18,7	21,2	25,7	21,4	32,2

Fonte: ISFORT, *Indagine audimob della mobilità*, anni vari.

TAV. 6

INDICE DI UTILIZZAZIONE ABITUALE DEL TRASPORTO FERROVIARIO
E GRADO DI SODDISFAZIONE DEL SERVIZIO (%)

Regione o macroarea	Indice di utilizzazione abituale ^(a)				Grado di soddisfazione del servizio ^(b)				
	1995	2000	2005	2007	1995	2000	2005	2007	2009
<i>Nord Ovest</i>	6,1	5,4	4,9	5,6	58,3	53,2	45,8	40,9	44,5
Piemonte	3,7	3,0	5,3	4,1	58,5	52,6	42,4	40,9	43,8
Valle d'Aosta	3,8	1,7	3,3	3,5	64,1	55,1	39,3	41,1	50,2
Lombardia	6,2	5,6	4,5	5,7	58,9	52,5	46,6	42,0	45,5
Liguria	12,5	12,5	7,4	9,2	55,4	57,7	49,6	35,3	41,1
<i>Nord Est</i>	3,7	3,4	3,5	3,9	66,0	57,7	52,2	47,2	51,1
Trentino-Alto Adige	4,0	3,3	2,8	4,1	69,1	64,5	59,1	59,2	63,5
<i>Bolzano</i>	5,1	3,9	3,7	5,9	68,0	64,5	59,1	63,1	68,1
<i>Trento</i>	2,9	2,7	2,0	2,6	70,4	64,4	59,2	55,0	58,2
Veneto	3,4	4,2	3,9	4,1	68,7	59,2	50,5	47,1	48,5
Friuli-Venezia Giulia	4,1	4,0	4,5	3,3	63,1	57,1	56,0	46,6	52,8
Emilia-Romagna	4,0	2,4	3,0	3,7	63,1	54,7	51,7	44,7	50,6
<i>Centro</i>	5,0	4,8	5,1	5,0	60,0	56,0	52,0	48,9	50,3
Toscana	6,6	5,0	6,1	4,8	61,6	53,2	47,1	46,0	49,8
Umbria	5,4	5,3	3,6	2,0	56,1	52,7	50,2	44,7	47,9
Marche	2,0	2,5	1,8	3,0	56,1	52,2	52,4	44,0	47,9
Lazio	4,7	5,2	5,6	6,2	60,1	59,1	55,4	52,6	51,7
<i>Sud e Isole</i>	4,7	4,9	4,3	4,1	53,1	49,5	49,7	42,7	44,5
Abruzzo	1,8	2,4	2,2	2,7	58,5	54,0	56,5	38,1	47,7
Molise	0,8	0,8	1,5	3,0	54,9	47,7	49,7	38,3	42,2
Campania	9,0	10,2	7,7	7,0	57,9	56,6	58,7	51,1	51,5
Puglia	5,5	5,4	6,1	4,5	56,0	49,7	45,3	40,9	47,4
Basilicata	3,4	2,3	2,6	2,1	47,3	36,8	51,9	36,4	43,6
Calabria	3,8	2,6	2,4	3,6	40,1	37,3	37,4	33,4	28,8
Sicilia	2,2	2,1	1,3	1,6	50,2	45,2	42,7	34,2	33,9
Sardegna	2,5	2,4	2,3	3,4	56,2	50,2	48,2	42,8	52,0
<i>Italia</i>	5,0	4,7	4,5	4,6	58,6	53,6	49,6	44,3	47,2

Fonte: ISTAT, *Indagine multiscopo*, anni vari.

^(a) Lavoratori, scolari e studenti di 3 anni e più che utilizzano il treno abitualmente per recarsi a lavoro, asilo o scuola sul totale degli intervistati, inclusi coloro che vanno a piedi o in bicicletta. ^(b) Media delle persone che si dichiarano soddisfatte delle 7 caratteristiche del servizio rilevate (frequenza corse, puntualità, possibilità di trovare posto a sedere, pulizia delle vetture, comodità degli orari, costo del biglietto, informazioni sul servizio) sul totale degli utenti del servizio.

TAV. 7

INDICI DI SODDISFAZIONE PER ALCUNI MEZZI DI TRASPORTO
(VALORI MEDI - VOTO DA 1 A 10)

Anno e area		Moto	Auto privata	Pullman/bus extraurbano	Treno IC/ES	Treno locale/ regionale
2004	Nord-Ovest	8,04	7,54	6,59	7,35	6,73
	Nord-Est	7,86	7,63	6,66	7,40	6,92
	Centro	8,22	7,52	6,49	7,35	6,80
	Sud e Isole	7,90	7,48	6,54	7,19	6,88
	<i>Italia</i>	<i>8,00</i>	<i>7,53</i>	<i>6,56</i>	<i>7,31</i>	<i>6,82</i>
2005	Nord-Ovest	8,45	7,99	6,42	6,78	6,01
	Nord-Est	8,22	8,08	6,93	6,58	6,25
	Centro	8,41	7,76	6,52	6,68	6,04
	Sud e Isole	8,38	7,92	6,44	6,52	5,96
	<i>Italia</i>	<i>8,38</i>	<i>7,94</i>	<i>6,53</i>	<i>6,64</i>	<i>6,05</i>
2006	Nord-Ovest	8,42	8,09	6,48	6,56	5,86
	Nord-Est	8,37	8,26	6,88	6,66	5,99
	Centro	8,46	8,02	6,50	6,91	6,05
	Sud e Isole	8,34	8,20	6,37	6,66	5,91
	<i>Italia</i>	<i>8,39</i>	<i>8,15</i>	<i>6,50</i>	<i>6,69</i>	<i>5,94</i>
2007	Nord-Ovest	8,68	7,97	6,55	6,53	5,79
	Nord-Est	8,49	8,15	6,89	6,44	6,19
	Centro	8,55	7,77	6,45	6,59	6,17
	Sud e Isole	8,25	7,96	6,38	6,67	5,92
	<i>Italia</i>	<i>8,46</i>	<i>7,96</i>	<i>6,53</i>	<i>6,57</i>	<i>6,00</i>
2008	Nord-Ovest	8,47	7,85	6,53	6,51	5,87
	Nord-Est	8,45	8,06	7,10	6,90	6,34
	Centro	8,55	7,66	6,38	6,59	6,10
	Sud e Isole	8,34	7,80	6,38	6,27	5,93
	<i>Italia</i>	<i>8,44</i>	<i>7,83</i>	<i>6,54</i>	<i>6,54</i>	<i>6,03</i>
2009	Nord-Ovest	8,35	7,91	6,58	6,45	5,67
	Nord-Est	8,44	8,23	6,93	6,91	6,18
	Centro	8,27	7,69	6,51	6,81	6,06
	Sud e Isole	8,11	7,80	6,58	6,48	5,86
	<i>Italia</i>	<i>8,26</i>	<i>7,89</i>	<i>6,63</i>	<i>6,63</i>	<i>5,91</i>

Fonte: ISFORT, Indagine audimob della mobilità, anni vari.

TAV. 8

INDICATORI DEL SERVIZIO FERROVIARIO VIAGGIATORI^(a)

Voci	1990	1995	1997	2000	2007	2008	2009
Trenitalia - totale							
mezzi di trazione	5.405	5.269	5.106	5.272	3.140	3.070
carrozze e rimorchi	12.280	11.736	10.927	10.437	8.038	7.614
di cui: <i>a media/lunga perc.</i>							
viaggiatori trasportati	70.000	72.146	70.282	66.994
viaggiatori-km	27.537	24.166	23.587	22.237
percor. media di un viaggiat.	393	335	336
posti-km	48.441	47.779	46.915
treni-km	82.638	83.276	80.957	80.056
di cui: <i>trasporto regionale</i>							
viaggiatori trasportati	408.200	479.725	513.672	516.055
viaggiatori-km	19.596	21.819	22.180	22.168
percor. media di un viaggiat.	48	45	43
posti-km	64.998	71.862	71.351
treni-km	169.193	185.838	187.485	187.140
Ferrovie regionali ^(b)							
viaggiatori trasportati	149.351	148.721	151.103	160.424	214.786	211.665
viaggiatori-km	2.780	2.792	3.001	2.439	3.749	3.712
percor. media di un viaggiat.	18,6	18,77	21,31	15,2	17,45	17,54
treni-km ^(c)	32.421	39.061	38.773
mezzi di trazione	938	1.089	1.057	1.187	1.188	1.189
carrozze	706	746	702	728	571	580

Fonte: MINISTERO DELLE INFRASTRUTTURE E DEI TRASPORTI, *Conto nazionale delle infrastrutture e dei trasporti*, anni vari.

^(a) Viaggiatori e treni-km in migliaia, viaggiatori-km e posti-km in milioni, percorrenza media di un viaggiatore in km. Dal 2000 la serie relativa ai viaggiatori e ai viaggiatori-km Trenitalia è calcolata con modalità differenti; dal 2005 al 2009 i dati relativi a viaggiatori-km, posti-km e treni-km includono nella lunga percorrenza anche i volumi nazionali relativi a Cisalpino. ^(b) Esclusi i due gestori attivi nel trasporto passeggeri operanti esclusivamente in regime di licenza (SAD, che opera nella provincia autonoma di Bolzano e SERFER, di proprietà Trenitalia, che opera sulla tratta Roma-Civitavecchia. ^(c) Il dato del 2000 è relativo al 2002.

TAV. 9

INDICATORI ECONOMICI DELLE IMPRESE FERROVIARIE TRASPORTO PASSEGGERI^(a)

Voci	2000	2001	2007	2008	2009
Trenitalia - MLP					
Ricavi	1.684	1.804	1.927	1.986
di cui: <i>da CS con lo Stato</i> ^(b)	190	189	233	224
Ricavo medio a viaggiatore-km	5,5	6,7	7,2
<i>Load factor</i>	57	56	51	50
% treni arrivati fra 0 e 5' di ritardo	87,0	87,0	88,4	90,0
Trenitalia - Trasporto regionale					
Ricavi	2.049	2.584	2.662	2.735
di cui: <i>da CS</i>	1.432	1.800	1.857	1.939
<i>da CS con le Regioni</i> ^(c)	1.214	1.560	1.631	1756
<i>da CS con lo Stato (RSS)</i>	218	240	226	183
Ricavo medio a viaggiatore-km	3,2	3,6	3,6
Ricavo medio a viaggiatore-km (CS)	6,2	7,2	7,4
<i>Load factor</i>	30	29	30	31
% treni arrivati fra 0 e 5' di ritardo	85,5	88,2	90,2	90,3
Ferrovie regionali ^(d)					
<i>Load Factor</i>	32,1	33,1	21,2	21,7
Nord	35,5	35,0	33,2	35,5
Centro	29,7	29,8	21,9	21,0
Sud e Isole	29,5	34,6	13,9	14,6
Rapporto ricavi/costi	20,3	19,4	24,6	24,5
Nord	31,7	30,3	24,5	23,5
Centro	20,7	16,8	23,5	24,0
Sud e Isole	13,8	13,7	25,2	25,9

Fonte: MINISTERO DELLE INFRASTRUTTURE E DEI TRASPORTI, *Conto nazionale delle infrastrutture e dei trasporti, anni vari e bilanci Trenitalia e FS*, anni vari.

^(a) I ricavi da traffico sono in milioni di euro, il ricavo medio a viaggiatore-km in cent./€, il *load factor* è calcolato come viaggiatori/postiX100. Sono esclusi dai ritardi quelli dovuti a scioperi o a cause esterne a FS. Il rapporto ricavi/costi è dato dai ricavi al netto delle sovvenzioni in conto esercizio rispetto ai costi della produzione. ^(b) Include il CS per la MLP, per il servizio viaggiatori notturno e per le agevolazioni e gratuità tariffarie a particolari categorie di viaggiatori. ^(c) Inclusi i corrispettivi percepiti da altri enti territoriali, dal 2008 quelli per il CS con la Regione Friuli-Venezia Giulia e dal 2009 quelli per i CS con le Province autonome di Trento e Bolzano. ^(d) Il dato esclude i due gestori attivi nel trasporto passeggeri che operano esclusivamente in regime di licenza.

TAV. 10

LE FERROVIE IN ALCUNI PAESI EUROPEI

Paesi	1970	1980	1990	2000	2006	2007	2008						
Km di rete													
	Totale	Totale	Totale	di cui: AV	Totale	di cui: AV	Totale	di cui: AV	Totale	di cui: AV	Totale	di cui: AV	Var. % rete tot. 90-08
Germania	43.777	42.765	48.981	90	36.588	636	34.122	1.291	33.890	1.285	33.855	1.285	-30,9
Spagna	15.850	15.724	14.539	0	13.868	471	14.652	1.272	15.012	1.516	15.041	1.594	3,5
Francia	37.582	34.362	34.070	699	29.272	1278	29.463	1.573	29.918	1.893	29.901	1.893	-12,2
Italia	16.073	16.138	16.066	224	16.187	248	16.627	562	16.667	562	16.862	744	5,0
UK	19.330	18.030	16.914	0	17.044	0	19.956	74	16.208	113	16.218	113	-4,1

	Densità rete ^(a)	Pass.- km ^(b)	Densità rete ^(a)	Pass.- km ^(b)	Densità rete ^(a)	Pass.- km ^(b)	Densità rete ^(a)	Pass.- km ^(b)	Densità rete ^(a)	Pass.- km ^(b)	Densità rete ^(a)	Pass.- km ^(b)	Var. % pass.- km 90-08
Germania			114,8	61,0	102,5	75,4	95,6	79,0	95,0	79,3	94,2	81,8	34,0
Spagna			28,7	15,5	27,4	20,1	28,6	22,1	29,3	21,9	30,1	24,0	54,8
Francia			62,6	63,7	53,8	69,9	54,3	78,8	55,1	80,3	56,0	85,0	33,3
Italia			53,4	44,7	53,8	49,6	55,9	50,8	56,0	49,8	56,8	49,5	11,4
UK			69,3	33,4	69,9	38,4	81,8	47,0	66,4	50,1	54,0	52,7	57,7

Fonte: EUROSTAT, anni vari e ASTRID (2008). I dati sono omogenei fra loro nel tempo, ma includono solo alcune reti e gestori di ciascun paese.

^(a) Km X 1.000 km²; ^(b) Miliardi.

GLOSSARIO

Infrastruttura ferroviaria: è costituita dalla via e dalle relative opere civili, dagli impianti tecnologici per i sistemi di trazione, segnalamento e sicurezza.

Load factor (coefficiente di riempimento): rapporto fra i posti effettivamente occupati e quelli disponibili.

Passeggeri-km: somma dei prodotti del numero dei passeggeri trasportati per le relative percorrenze in km.

Reti ferroviarie: gli insiemi di linee fra loro interconnesse.

Traccia oraria: il tempo programmato di impegno dell'infrastruttura ferroviaria da parte di un treno per viaggiare dalla stazione di origine a quella di destinazione.

Treni-km (Posti-km): sommatoria dei prodotti del numero dei treni (posti) per le relative percorrenze in km.

Interoperabilità: cfr. il sito http://europa.eu/legislation_summaries/transport/rail_transport/tr0009_it.htm.

BIBLIOGRAFIA

- AFFUSO L., «Il servizio ferroviario nell'esperienza britannica e internazionale», *Mercato Concorrenza Regole*, n. 1, 2003, pp. 99-115.
- AGCM, *Schema di regolamento di attuazione della direttiva n. 91/440/CEE relativa allo sviluppo delle ferrovie comunitarie*, Segnalazione al Parlamento (AS140) del 5 giugno 1998.
- .-, *Reperimento del materiale rotabile ferroviario necessario per l'espletamento delle gare per l'aggiudicazione dei servizi ferroviari di competenza regionale*, Segnalazione al Parlamento (AS262) dell'8 luglio 2003a.
- .-, *Separazione tra gestione delle infrastrutture e servizi di trasporto ferroviario*, Segnalazione al Parlamento (AS265) del 13 agosto 2003b.
- .-, *Considerazioni e proposte per una regolazione pro concorrenziale dei mercati a sostegno della crescita*, Segnalazione al Parlamento (AS453), giugno 2008.
- .-, *Definizione del servizio universale nel trasporto ferroviario e affidamento dei conseguenti oneri di servizio*, Segnalazione al Parlamento (AS528), giugno 2009.
- ASTRID, *Infrastrutture e servizi a rete tra regolazione e concorrenza. Spunti per una politica delle infrastrutture di trasporto*, giugno 2008.
- BERENGO GARDIN P., *Ferrovie italiane*, Roma, Editori Riuniti, 1988.
- BOITANI A. - PETRETTO A., «I servizi pubblici locali tra governance locale e regolazione economica», in ROBOTTI L. (a cura di), *I servizi pubblici locali*, Bologna, Il Mulino, 2002, pp. 25-65.
- BRENNA C., «Ferrovie zero sussidi», *IBL Focus*, n. 152, 2009.
- CAMBINI C., «Concorrenza e regolazione nel settore ferroviario: l'importanza del materiale rotabile», *Economia e Politica Industriale*, n. 2, 2009, pp. 47-72.
- CAMBINI C. - BUZZO MARGARI B., «Le gare nelle ferrovie locali», *Documento di Ricerca Hermes*, ottobre 2005.
- CAMPOS J. - CANTOS P., «Rail Transport Regulation», *World Bank Research Working Papers*, November 1999.
- COORDINAMENTO DEI PENDOLARI - FEDERCONSUMATORI, «Essere pendolari, una scelta difficile - anno 2008», www.federconsumatori.it, 2008.
- CORTE DEI CONTI, *La gestione del trasporto pubblico locale e lo stato di attuazione della riforma a livello regionale*, 2004.
- .-, *Relazione sul risultato del controllo eseguito sulla gestione finanziaria di Ferrovie dello Stato S.p.A. per gli esercizi 2007-2008*, 2009.
- EUROPEAN ENVIRONMENT AGENCY, «Size, Structure and Distribution of Transport Subsidies in Europe», *EEA Technical Report*, no. 3, 2007.
- EUROSTAT, «Energy & Transport in Figures», *Statistical Pocketbook*, vari anni.

- GARDINA M., «Alte velocità ferroviarie: problematiche concorrenziali per un uso non discriminatorio dell'infrastruttura e del rotabile», in POLIDORI G. - MUSSO E. - MARCUCCI E. (a cura di), *I trasporti e l'Europa. Politiche, infrastrutture, concorrenza*, Milano, Angeli, 2006.
- HENSHER D.A. - BREWER A.M., *Transport*, Oxford University Press, 2001.
- LEGAMBIENTE, *Pendolaria*, anni vari.
- LITMAN T., «Valuing Transit Service Quality Improvements», Victoria Transport Policy Institute, www.vtpi.org/traveltime.pdf, 2007.
- MACCHIATI A. - CESARINI A. - MALLUS A. - MASSIMIANO M., «Concorrenza e privatizzazione nel settore ferroviario in Europa. Problemi aperti e prospettive», *Mercato concorrenza regole*, n. 1, 2007, pp. 13-50.
- MERKERT R. - SMITH A.S.J. - NASH C.A., «Benchmarking of Train Operating Firms - A Transaction Cost Efficiency Analysis», *Transportation Planning and Technology*, vol. 33, no. 1, 2010, pp. 33-53.
- MINISTERO DELLE INFRASTRUTTURE E DEI TRASPORTI, *Servizi di trasporto ferroviario di passeggeri a media e lunga percorrenza: servizio universale*, Relazione illustrativa sottoposta al CIPE il 27 marzo 2008.
- PAROLIN R. - KLARMANN J., «Le gare per i servizi locali - l'approccio lombardo», *Quaderni RT*, n. 10, 2007.
- PONTI M. - BERIA P., «La rotaia arrugginita e il vagone del futuro», *il Mulino*, n. 6, novembre-dicembre 2007, pp. 1028-1041.
- RAGAZZI G., «L'Autobrennero, una concessione generosa», www.Lavoce.info, 2009.
- RAMELLA F., «Controordine: meglio il tunnel dei mezzi pubblici», www.Lavoce.info, 2010.
- REGIONE EMILIA-ROMAGNA, *Rapporto annuale di monitoraggio della mobilità e del trasporto in Emilia-Romagna*, anni vari.
- SCARPA C. (2007), «Il trasporto ferroviario: l'alba di una riforma», in CAMBINI C. - GIANNACCARI A. - PAMMOLLI F., *Politiche di liberalizzazione in Italia. Proposte di riforma e linee di intervento settoriali*, Bologna, Il Mulino, 2007.
- STAGNI G., «La riforma delle ferrovie», www.miol.it/stagniweb, dicembre 2004.
- .-, «Legislazione sulla messa a gara dei servizi ferroviari», www.miol.it/stagniweb, giugno 2010.
- .-, «Finanziamenti alle ferrovie: il 2012 e una panoramica sul decennio scorso», www.miol.it/stagniweb, marzo 2012.
- UIC, *Rail Transport and Environment - Facts & Figures*, November 2008.
- VAGHI C. - MILOTTI A., «Gli strumenti innovativi per finanziare le infrastrutture di trasporto in Italia», *Economia dei Servizi*, n. 3, 2009, pp. 377-397.

A New Textbook Approach to Macroeconomics: A Debate*

Alessia Amighini
Università del Piemonte Orientale

Emiliano Brancaccio
SEGIS Università del Sannio

Francesco Giavazzi
Università Bocconi e MIT

Marcello Messori
Università di Roma “Tor Vergata”

Messori’s paper analyzes the impact of the recent crises on the teaching of macroeconomics. In contrast with what happened during the Thirties, today we do not have a new macroeconomic paradigm. This is why the mainstream textbook of Blanchard, Amighini, Giavazzi (2010) remains irreplaceable for introductory macroeconomics. This conclusion is refused by Amighini, Giavazzi as well as by heterodox economists such as Brancaccio. The first two authors argue that the criticisms raised at the “mainstream” approach to the teaching of macroeconomics overlook the need for a strong pedagogy. On the contrary, Brancaccio criticizes the mainstream approach through modification of the functional form of a given equation system and reversal of its exogenous and endogenous variables.

[JEL Classification: A20; B22; B50].

Keywords: macroeconomics; teaching; comparative approach.

* <alessia.amighini@eco.unipmn.it>, insegna presso l’Università del Piemonte orientale; <emiliano.brancaccio@unisannio.it>, insegna presso l’Università del Sannio; <francesco.giavazzi@unibocconi.it>, insegna presso l’Università Bocconi di Milano; <messori@uniroma2.it>, insegna presso l’Università “Tor Vergata” di Roma.

DEVELOPING A NEW TEXTBOOK APPROACH TO MACROECONOMICS

Marcello Messori

1. - The State of the Art

Olivier Blanchard's textbook, adapted for publication in Europe by Alessia Amighini and Francesco Giavazzi (Blanchard, Amighini and Giavazzi, 2010; henceforth BA&G), is the best introduction to macroeconomics available today. Its short-term analysis is based on the "neoclassical synthesis", with money wages exogenously given and the money supply determined by monetary policy choices (Modigliani, 1944). For short- and medium-term analysis, this textbook employs a model of aggregate supply and demand (AS-AD) that combines the monetarist reinterpretation of the Phillips curve (Phelps, 1967) with a simplified treatment of the Walrasian microfoundations, typical of the "new classical macroeconomics" (Lucas, 1972; Sargent, 1973), and the endogenous rigidities of the particular strand of the "new Keynesian economics" founded on market imperfections (Mankiw, 1985; Blanchard and Kiyotaki, 1987; Ball and Romer, 1990). For long-term analysis, it refers to the "real business cycle" and endogenous growth models that generate optimal equilibria. Consequently, monetary policy and fiscal policy are effective in the short term but neutral in the medium term, and an expansionary fiscal policy can even have a negative "real" impact in the long term. The scope for non-distortionary policy action is limited to short-term monetary policy.

Hence BA&G offers a didactic "synthesis" between the most up-to-date versions of the traditional approach (the dynamic stochastic general equilibrium models: DSGE) and the strand of the "new Keynesian economics" based on endogenous rigidities. This synthesis, which in the theoretical literature produced the DSGE models with endogenous rigidities (DSGER), dominated the field of macroeconomics and inspired (self-)regulation and policymaking between the 1990s and the first few years of the new century (see among others: Taylor and Woodford, 1999; Clarida *et al.*, 2000; Blanchard and Galí, 2007). However, the financial and economic crisis of 2007-09 and the current European sovereign debt crisis have bared the limits of this theoretical approach, demonstrating that the conceptual constructs produced by Walrasian microfoundations and DSGER models are unable to predict or explain economic phenomena characterized by

systematic market failures, persistently high rates of involuntary unemployment, rising income inequality and structural imbalances (Quiggin, 2010; Barucci and Messori, 2012).

This state of affairs should have prompted a reflection on the weak points of the dominant economic theory and on the possibility of constructing a new paradigm to incorporate into a new approach to teaching macroeconomics. But it hasn't. In contrast with what happened in the 1920s and 1930s after the crises of 1907-1908 and 1929-1933, the present decade cannot be called a period of "high theory" (Shackle, 1983). And, in accordance with Popper's doctrine of falsifiability, the lack of alternatives is keeping alive theoretical approaches that have proven inadequate to analyze the recent crises and their macroeconomic impact. So our students still rely for their training – and most likely will continue to do so – on textbooks like BA&G, which, accurate and open in its presentation as it may be, still embodies theoretical approaches that should now be obsolete in view of the legacy of the crises.

2. - Is Something Changing?

The extensive set of macroeconomic textbooks obviously includes a number of contributions which follow neither the standard traditional approach nor its most up-to-date versions. Moreover, during and immediately after the financial and "real" crises (May 2007 - April 2009), several well-known macroeconomic textbooks were brought out in new editions, some of which tried to learn a few lessons from the recession and its determinants (for instance, Colander 2010). Finally, in the recent macroeconomic debate various criticisms have been directed towards the analytical foundations of DSGE and DSGER models (for instance, De Grauwe, 2010). However, as far as I know, few authors have pursued the objective of challenging the framework of one of the most famous textbooks by means of internal criticisms. The critique of BA&G by Emiliano Brancaccio (2012) is an interesting attempt, despite the lack of a new analytical paradigm as a frame of reference, to dent the prevailing conformism of macroeconomic theory and teaching. Beyond underscoring the major weaknesses of BA&G's approach, Brancaccio sets himself the ambitious objective of constructing an alternative macroeconomic textbook. Even if he does not meet this objective, his contribution develops analytical "building blocks" while also reinterpreting or using many of BA&G's results. This opens up new paths and perspectives of inquiry and en-

ables students to become accustomed to a diversity of representations of economic reality. Let us illustrate with two examples.

First, Brancaccio renders explicit many of the links between short, medium and long-term models or links within each of these models that students find it hard to discern in the original version of BA&G. For instance, in the AS-AD model he already introduces the variable relating to technology and productivity (designated *A*). The resulting bridge between that model and the subsequent model of growth with technical progress sheds light on the analytical incongruities underlying the limited space accorded to monetary and fiscal policies in BA&G's macroeconomic approach. Even more felicitous is the expository device of graphically connecting the equilibrium between the wage curve and the price curve in the labor market with the equilibrium between the aggregate supply and aggregate demand curves (AS and AD). This makes it immediately clear why AS is determined in the labor market and why the monetarist version of the Phillips curve and the natural rate of unemployment are crucial to the modern version of mainstream macroeconomics.

Secondly, Brancaccio (2012) correctly takes over a number of analytical blocks from BA&G's schema, thereby satisfying methodological standards and incorporating recent advances in the literature. After all, robust alternative paradigms cannot be built simply by turning back to the past (often reduced to Keynes's original contribution) and rejecting seventy-five years of theoretical debate. In particular, the separation between micro- and macroeconomics, which lasted more than three decades, cannot be restored. At the turn of the 1970s, the two main branches of theoretical economics reached a unity of method and analysis. This was achieved by means of the Walrasian microfoundations of macroeconomics, which spelled the decline of Hicks and Modigliani's neoclassical synthesis and Friedman's monetarism but which also brought out many analytical weaknesses of the *General Theory*.¹ It is entirely legitimate for a critical approach to reject traditional microfoundations, *i.e.* based on the Walrasian model of general economic equilibrium, and understandable that this approach chooses to oppose every form of "methodological individualism" and pursues more complex rela-

¹ The need for macroeconomics to rest on microfoundations was raised by LUCAS R.E. (1972) and SARGENT T.J. (1973) within the "new classical macroeconomics". The subsequent critique of the new classical macroeconomics by diverse strands of the "new Keynesian economics" did not call this need into question. However, one of these strands, based on the works of Stiglitz and others (STIGLITZ J.E., 1987; GREENWALD B.C. and STIGLITZ J.E., 1987 and 1991; STIGLITZ J.E. and WEISS A., 1992), used non-Walrasian microfoundations.

tions between micro- and macroeconomics (including macrofoundations of microeconomics). The important thing is to avoid the error of restoring the sterile separation between the two main branches of theoretical economics.

3. - Structure of My Paper

In what follows I summarize the current foundations of macroeconomic theory in order to evaluate their impact on the contents of BA&G (Section 4). This brings out a number of weaknesses, which attracted Brancaccio's criticism (Section 5). The consequent building blocks of an alternative model address some crucial problems raised by the recent crises. From the analytical standpoint, however, they present as many problematic aspects as the traditional models (Sections 6 and 7). So the question is to determine whether at least some of Brancaccio's objectives cannot be attained by a different critique of BA&G's schema (Section 8). The conclusion will show that this more modest criticism also leaves a number of problems open (Section 9).

4. - The Underpinnings of Current Macroeconomics

Before the financial crisis erupted in May 2007, the frontier of macroeconomic theory was represented by a new synthesis: that between "real business cycle" theory and the strand of new Keynesian macroeconomics founded on endogenous rigidities. The former argues that economic fluctuations are caused not by monetary but by "real" shocks, more specifically by technological shocks. Cycles are thus an essential aspect of economic growth since they permit innovations and productivity gains to be incorporated into the productive apparatus. By contrast, the above-mentioned strand of the new Keynesian economics restores short-run unemployment equilibria by introducing price stickiness justified by specific "adjustment costs" and by monopolistic markets. These two factors can in fact make it advantageous for the individual firm to reduce quantities rather than prices in the face of negative external shocks to aggregate demand. The synthesis between the two approaches produces models that admit incomplete and imperfect markets, speculative bubbles and short-term unemployment equilibria. But in these models the cyclical dynamic is imprisoned within a Walrasian general economic equilibrium. It follows that expansionary monetary and fiscal policies can have "real" effects in the short term, but their impact reverts to neutral in the medium term or, in the

case of increased public spending, becomes recessionary in the long term.

It has been remarked above that Blanchard contributed to the ascendancy of the synthesis between the new “real business cycle” approach and the new Keynesianism (Blanchard and Gali, 2007). Though in a form still influenced by the old neoclassical synthesis (short-term IS-LM model) and by the monetarism of Friedman and Phelps (expectations-adjusted Phillips curve) and though it adopts a number of simplifications,² BA&G transposes many of the results reached by this new synthesis into the AS-AD model (short and medium term) and into the growth setups, particularly with technical progress (long term).

Paraphrasing the fundamental features of BA&G, we can reduce the didactic propositions that characterize current macroeconomic theory to six points:

- (i) In the short term, it is always possible and advantageous for the set of firms to supply the amount of reproducible goods and services that is demanded at a given level of money wages. Therefore, taking into account the identity between income produced and income distributed, and the dependence of consumption demand on disposable income,³ the level of total demand determines the equilibrium level of output for a given price level and can cause abnormal unemployment rates.
- (ii) In the medium term, every adjustment of output to the volume of total demand provokes adjustments in money wages and goods prices, which in turn affect agents’ expectations and their demand behaviors. Therefore, the general equilibrium is determined by the interaction between the equilibrium in the labor market (AS equation) and the equilibrium in the goods market and the financial market (AD equation).
- (iii) In the medium-term equilibrium of the AS-AD model, the output level and the associated unemployment rate are always at their natural values; given technology, at the natural rate of unemployment the real wage is set by structural or institutional parameters and agents realize their price expectations. The market equilibria are optimal.
- (iv) In the short term, monetary policy and fiscal policy can influence the total demand for goods and can thus have “real” effects on the output level and

² The most evident simplification is the adoption of adaptive expectations for future price formation. In this way, BA&G takes an analytical shortcut to results that in more complex models would be determined by Walrasian microfoundations and by the combination between endogenous rigidities and imperfectly competitive markets.

³ In turn, disposable income is equal to the difference between distributed income and taxation.

- the unemployment rate; in the medium term instead, given technology, those policies are “neutral” with respect to the optimal “real” equilibria and only affect the price level.
- (v) In the long term, given technology, the equilibrium coincides with stationary values of capital per worker and output per worker. An expansionary fiscal policy can cause a decline in the private sector’s propensity to invest and thereby lower the stationary values of those two ratios. Because of these perverse effects of expansionary fiscal policies in the long run, the only effective policy instrument is short-term monetary policy.
 - (vi) In the long term, the pace of technical progress sets the equilibrium growth rate of capital per worker and of output per worker.

5. - Two General Criticisms

Brancaccio concentrates his fire on the first four propositions and just grazes the fifth. The following passage neatly sums up his position: «From the mainstream perspective [...] the “natural” equilibrium can be said to represent the inviolable limit of wage claims and expansionary policies. It is possible to modify the “natural” equilibrium, but only by intervening on the so-called fundamentals of technical progress and the availability of labor and capital, or by reducing the market power of firms and trade unions [...]. By contrast, the critical approach considers that there is no “natural” equilibrium levels of production, employment and real wages that are independent of trade-union pressures or expansionary economic policies», (pages XVIII-XIX).

Simplifying somewhat, Brancaccio’s critical observations regarding BA&G’s AS-AD model – observations that underpin the alternative proposition we have just cited – can be reduced to two points: (a) the critique of the demonstration that monetary policy and fiscal policy are ineffective in obtaining a different medium-term equilibrium from that fixed by the natural rate of unemployment and the natural level of output; (b) the critique of the assertion of independence between the two key parameters for the working of the labor market, *i.e.* the mark-up (μ) – which Brancaccio (2012) reinterprets as “profit margin” – and the residual indicator of workers’ bargaining power (z) – which is renamed “degree of workers’ conflict”.⁴

⁴ In Section 7, where we examine the implications of point (b), we accept the renaming of parameter z but call that of μ into question.

In what follows, I intend to demonstrate that these critical observations are empirically plausible. However, if points *(a)* and *(b)* are to be taken as “building blocks” for constructing an alternative macroeconomic model, empirical plausibility is not enough. It is also essential to prove that the two points rest on rigorous analytical propositions compatible with the remaining parts of the reference model; but on this score, neither seems to pass muster. Although point *(a)* is of crucial importance, its analytical justification is weak, and as for point *(b)*, which is more “ideological” and based on a modification of the definition of the parameters examined, its impact becomes significant only if it is supported by point *(a)*.

6. - Analytical Limits of the First Critique

With regard to point *(a)*, Brancaccio (2012) quite properly aims to demonstrate that in the medium term the labor market is unable to establish the unemployment rate compatible with stability in the changes in the general price level, that is with coincidence of market prices with expected prices. This would be tantamount to rejecting the concepts of “natural” unemployment rate and “natural” output level and, consequently, to denying that market mechanisms ensure – in the medium term – the optimality of market equilibria, the neutrality of money, and hence the ineffectiveness of monetary and fiscal policies.

To arrive at these results, Brancaccio calls into question the inverse relation between the general price level and the aggregate demand that characterizes standard AS-AD model as well as BA&G’s one and makes AD a downward-sloping curve. The point of departure of his argument is the following statement: “Only if AD is decreasing [...] can it be held that price changes will always guarantee the economic system’s spontaneous convergence to the natural level of output Y_n and thus to the natural rate of unemployment u_n ” (page 42). Therefore, Brancaccio’s objective is to demonstrate that the behavior of the AD curve is indeterminate; for example, that this curve can become infinitely rigid with respect to the general price level. If this were so, equilibria with Keynesian unemployment and expansionary economic policies with positive “real” effects would be possible even in the medium term.

From the analytical standpoint, the trouble is that the thesis of AD’s non-negative slope leads us back to the problem that Keynes dealt with unsatisfactorily in Chapter 19 of the *General Theory*: to demonstrate that, starting out from a level of economic activity with involuntary unemployment, a fall in money wages

and/or the general price level does not necessarily restore a full employment equilibrium. Not by chance, Brancaccio's proof that AD may have a non-negative slope boils down to a restatement of Keynes's argument there.

Keynes (1936, pages 263-267) recognizes that a decline in money wages and in the general price level increases the amount of money supplied in real terms; and that, all else being equal, this increase is trending downwards the monetary interest rate and upwards the investment demand, which in turn raises the equilibrium level of aggregate output by means of the income multiplier. However, Keynes adds that there are two links at which this causal chain can be interrupted: a worsening of wealth holders' expectations can make the interest rate sticky downwards even in the presence of increments in the money supply in real terms (the "liquidity trap"); and a worsening of entrepreneurs' long-term expectations can make investment demand sticky upwards even in the presence of a decline in interest rates.

These observations of Keynes are empirically plausible because they reflect, albeit in a stylized way, what actually happened during some phases of the crisis of 1929-33 and above all during the crisis of 2007-2009, and what is now happening in the sovereign debt crisis of the European Economic and Monetary Union. From the analytical standpoint, however, Chapter 19 is one of the weakest parts of Keynes's framework. As the author explicitly underscores in Chapter 18 (pages 246-247), in the first seventeen chapters expectations are treated as "ultimate independent variables", and Chapters 5 and 12 serve to justify this assumption. The transformation in Chapter 19 of these "ultimate independent variables" into dependent variables, subject to changes in money wages and prices, must be written off as an *ad hoc* assumption.

Brancaccio (2012) offers a second reason, unrelated to changes in expectations, for questioning the inverse relation between the interest rate and investment demand: the inclusion of the interest rate in production costs. This implies that a fall in the interest rate will normally bring down the price of investment goods and therefore lower the value of capital assets. However, these decreases have ambiguous effects on the sign of the relation between the two variables. The price decline lowers the general price level and thus the expected price level (see no. 2 above), causing AS to shift downward. In addition, the decline in the price of investment goods raises the marginal efficiency of capital and thereby increases investment demand,⁵ causing a rightward shift of IS (in the IS-LM model) and

⁵ In the *General Theory*, Keynes defines the marginal efficiency of capital as being «equal to that rate of discount which would make the present value of the series of annuities given by the returns expected from the capital-asset during its life just equal to its supply price», (page 135).

hence of AD; and this last effect is reinforced by the impact that the fall in the general price level has on the money supply in real terms. On the other hand, the decline in the price of investment goods can lower the value of the stock of capital goods held by firms; and the rightward shift of AD partially offsets the fall in the general price level. It is difficult to determine *a priori* which of these effects prevails, but the attention paid to aggregate demand suggests that the inverse correlation between the interest rate and investment demand is strengthened. Moreover, including the interest rate in production costs would also necessitate modifying the price function and thus the specification of AS.

7. - Analytical Limits of the Second Critique

The second of Brancaccio's critical observations on BA&G's approach (which we term "point (b)" in Section 5) refers to two key elements for the functioning of the labor market: μ and z . Following the standard definition, BA&G regards μ as the degree of monopoly enjoyed by the set of firms in their supply market and z as a set of institutional labor market arrangements considered as a residual. BA&G, that is to say, posits the absence of any link between the two elements, which are treated as parameters. At the opposite, Brancaccio (2012) argues that "the levels of z and μ " are "subject to social pressures and will therefore be determined, case by case, by the relative bargaining power of firms and workers. This means they can no longer be treated as both exogenous". (page 47). In the two limiting cases, "if relative bargaining power is favorable to the workers", z remains exogenous and μ becomes endogenous and, "if relative bargaining power is favorable to firms", the opposite happens (μ remains exogenous and z becomes endogenous). In the intermediate cases the two factors become interdependent. This creates interdependence between the wage curve and the price curve. If z is exogenous, the price curve shifts so as to intersect the given wage curve at the unemployment rate corresponding to the equilibrium level of output, set by the point where AS and AD intersect; if μ is exogenous, the opposite happens. In the intermediate cases, both curves shift.

This critique of BA&G should produce at least three results: the reversal of the causal chain that determines equilibrium in the labor market and in the AS-AD space, the elimination of the notion of the natural rate of unemployment, and the demonstration that wage bargaining does not affect only money wages but sets real wages. The problem is that these results, though significant empirically, are analytically fragile.

On the first aspect, Brancaccio's critique effectively disposes of the causal chain which – in BA&G's AS-AD model – opens with the determination of the natural rate of unemployment and the real wage rate in the labor market and closes with the determination of the natural level of output in the medium term. But this does not imply that the alternative scheme generates a causal chain from level of aggregate demand to equilibrium unemployment rate, because the transformation of at least one of the two labor market parameters into a dependent variable creates an interdependence between the labor market and the AS-AD space. For any given level of aggregate demand there is a corresponding unemployment rate, which is a factor in the relative bargaining power of workers and firms and hence in the real wage level and the general price level. The latter two levels, in turn, affect the level of aggregate demand through the determination of the equilibrium between IS and LM.

As to the second aspect, Brancaccio claims that his construction has the virtue of eliminating the notion of the natural unemployment rate. But this result can be ascribed to the end of the inverse relation between the general price level and aggregate demand – and the consequent abandonment of the natural level of output – rather than to the interdependence between the price curve and the wage curve. Accordingly, if the criticisms aimed at the downward slope of AD and the notion of natural output level are analytically fragile (as Section 6 shows), then the rejection of the natural unemployment rate notion is also ill-founded.

On the third aspect, Brancaccio is right to maintain that in his model bargaining between workers and firms sets the real wage rate, not just money wages. But in this case his critique should be directed not at BA&G but at Keynes. BA&G makes money wages a function of the general price level, the unemployment rate and the parameter z only as a first approximation. As they specify carefully from the outset and make evident in their construction of the AS-AD model and the expectations-adjusted Phillips curve, money wages depend on expected prices, not current prices. This means that workers do not fall victim to the money illusion as they do in the *General Theory*.⁶

Finally, the hypothesis of interdependence between z and μ needs some further explanation. As noted earlier, Brancaccio introduces this hypothesis by defining “mark-up” as “profit margin” rather than an indicator of firms' degree of mo-

⁶ In Chapter 2 of the *General Theory* Keynes asserts: «Whilst workers will usually resist a reduction of money-wages, it is not their practice to withdraw their labour whenever there is a rise in the price of wage-goods», (page 9).

nopoly power. This makes it plausible to postulate that as workers' bargaining power increases firms will have to forgo a part of their profit margin. It would be much less plausible to suppose that workers' increased power can affect the form of the firms' supply market, especially if one maintains the Keynesian "money illusion". Brancaccio (2012, Section 2.9) does tackle this problem, but the proposed solution rests, yet again, on the fragile analytical foundation of a non-negative slope of the aggregate demand curve.

8. - A More Modest Critique

As we have seen, Brancaccio finds various weak points in BA&G and suggests some possible alternative approaches. In many cases these alternatives are empirically plausible, but they require further analysis. They would necessitate: *(i)* making the expectations of financial investors and firms endogenous; *(ii)* clarifying the links between workers' bargaining power, the structure of the firms' supply market, and the consequent formation of profit; *(iii)* specifying the interactions between micro- and macroeconomics and the representation of the economic processes that characterize an alternative model. These are three complex, inter-related operations that can be declined in various ways.

This is not the place for a discussion of the merits of these three points, but I would suggest that the decisive one is point *(iii)*. One strand of the new Keynesian economics, which is represented mainly by Stiglitz and his co-authors⁷ and is alternative to the endogenous rigidities approach, has produced a rigorous formulation of Keynes's main results by incorporating into its models rational expectations, contractual designs determined by agency relationships, information asymmetries, quantity instead of price constraints, and a partial instead of a general equilibrium approach. It therefore differs from the prevailing macroeconomic synthesis and the related approach of BA&G, in that it adopts non-Walrasian micro-foundations that distance it from the general equilibrium framework. The problem is that we have no equally complete representation of economic processes alternative to Walrasian general equilibrium theory. The Swedish sequential approach (Ohlin, 1937; Lundberg, 1937), carried further by Hicks (1956, 1965) with the combination of single-period with multiperiod analysis (the so called "continuation theory"), leaves a number of problems open. It is no accident that Keynes, after

⁷ See note 1.

referring to that method in 1931 and 1932 (Keynes, 1973, page 184), dropped it in his successive drafts of the *General Theory*. And Stiglitz and his co-authors, while frequently mentioning time lags in transactions and sometimes introducing sequences in cyclical phases (in particular, Greenwald, Kohn and Stiglitz, 1990), gave up this method in favor of a partial equilibrium approach.⁸

Brancaccio's idea of turning one or two of BA&G's key labor market parameters into endogenous variables, however, may also suggest a less ambitious way of bringing out the analytical fragility of the concepts of natural rate of unemployment and natural level of output. This way could avoid the restriction of economic policy effectiveness to short-term monetary policy measures, even if it does not suffice for the creation of the "building blocks" of alternative models.

Monetary and fiscal policy choices tend to affect the medium-term values of z and μ . Expansionary fiscal policy, for instance, which increases the total demand for goods in the short term (shifting the IS curve to the right) can make the entry of new firms advantageous in the medium term and thus change the degree of monopoly in the supply market. This is clearly equivalent to changing the value of μ . This same policy also lowers the unemployment rate, enhancing the unions' bargaining power. If this results in a modification of labor market rules and institutions in the medium term, this will obviously also change the value of z . Analogous effects are produced by expansionary monetary policy. By lowering the rate of interest (shifting the LM curve downwards), this type of policy too increases the total demand for goods in the short term and accordingly can modify z and/or μ , as just described. However, as we have seen, in BA&G's framework any change in z or μ will shift the wage and price curves, thus altering the natural rate of unemployment and consequently the natural level of output.

The end result is that monetary and fiscal policies affect the natural unemployment rate and output level. That is, they have an impact on the real economy beyond the short term. When they come to depend on policy choices, the very concept of "natural" rates and levels loses economic importance, in that they no longer function as the center of gravity of the medium-term equilibrium of the economy. Market mechanisms do not ensure optimal equilibria in the medium run.

⁸ See also: MESSORI M. (1999). Truth to tell, when Stiglitz and his co-authors sought to construct complete macroeconomic models, they even often ended up resorting to a traditional general equilibrium framework (see, for example STIGLITZ J.E. and WEISS A., 1992). But this weakened the results they achieved in the analysis of individual markets.

9. - Conclusion

The observations set out at the end of the previous section are interesting because they bring out some of the analytical weaknesses of the new macroeconomics and highlight its utter inability to speak to the problems posed by the crisis of 2007-09 and the European sovereign debt crisis. However, this *pars destruens* is not accompanied by an equally potent *pars construens*. If we drop the synthesis between the “real cycle” theory and the new Keynesian economics hinging on endogenous rigidities but we still incorporate the microfoundations based on the Walrasian general equilibrium theory, we are faced with an unsatisfactory *dilemma*: either go back to the original Keynesian framework, whose main results are empirically plausible but based on *ad hoc* theoretical hypotheses, or else return to a traditional framework with instantaneous price adjustments, which yields optimal equilibria even in the short term.

To escape this *dilemma* without relapsing into a macroeconomics shorn of microfoundations and consequently into the separation of micro- from macroeconomics that weighed so heavily on the development of economics until the mid-1970s, the only possible way out appears to be to abandon the foundations based on the Walrasian general equilibrium. This can be done by defining a more complex relationship between micro- and macroeconomics or by adopting a non-Walrasian microeconomics. The first of these approaches is largely unexplored and might require us to go beyond the borders of economic theory and bring in other social sciences in order to treat economic and social institutions as dependent instead of exogenous variables. The second, however, can rely on that microeconomics approach which features agency relations and contract theory. But as we have seen, this approach – at least as formulated by Stiglitz and his co-authors – has proven unable to construct a unitary macroeconomic framework and has mostly taken refuge in partial equilibrium analysis. The question, raised a good number of years ago but still not answered satisfactorily (see for example: Amendola and Gaffard 1988; Graziani, 2003; Messori, 1999), is whether a Hicksian sequential scheme can fit the multiple partial-equilibrium results into a unitary framework as general as that assured by Walrasian equilibrium models.

Whatever our answer, one indisputable point remains: it would be unrealistic and inappropriate to make these crude working hypotheses the basis of a textbook. First-year or second-year students in macroeconomics need to be introduced to tested analytical tools that furnish a simple method of learning and a basic discipline. Only once this groundwork has been laid does it become essential

to stimulate the student's critical spirit by pointing out the various weaknesses in this basic framework. Obviously, as noted earlier (Section 2), there is a number of macroeconomic textbooks which follow neither the standard traditional approach nor its most up-to-date versions and which are able to offer a basic guide to the discipline. However, in my view, BA&G stands out in providing a framework that is tested but at the same time open enough to criticism. This is why I think that BA&G will remain, for many years to come, the best available teaching tool for introductory macroeconomics. Those teachers who, like myself, are unsatisfied with its analytical underpinnings will have to be content with noting the fragility of some of its key concepts. In this regard, a useful approach is to undermine the notions of natural unemployment rate and natural level of output.

BIBLIOGRAPHY

- AMENDOLA M. - GAFFARD J.L., *The Innovative Choice*, Basil Blackwell, Oxford, 1988.
- BALL L. - ROMER D., «Real Rigidities and the Non-Neutrality of Money», *Review of Economic Studies*, vol. 57, 1990, pages 183-203.
- BARUCCI E. - MESSORI M., «Come evitare l'economia degli zombie. Introduzione all'edizione italiana», in QUIGGIN J., *Zombie economics. Le idee fantasma da cui liberarsi*, Università Bocconi Editore, Milano, 2012.
- BLANCHARD O. - AMIGHINI A. - GIAVAZZI F., *Macroeconomics. A European Perspective*, Pearson, Harlow, (2010), It. transl., il Mulino, Bologna, 2011.
- BLANCHARD O. - GALÌ J., «Real Wage Rigidities and the New Keynesian Model», *Journal of Money, Credit, and Banking*, vol. 39 (suppl.), 2007, pages 35-66.
- BLANCHARD O. - KIYOTAKI N., «Monopolistic Competition and the Effects of Aggregate Demand», *American Economic Review*, vol. 77, 1987, pages 647-666.
- BRANCACCIO E., *Anti-Blanchard. Un approccio comparato allo studio della macroeconomia*, F. Angeli, Milano, 2012.
- CLARIDA R. - GALÌ J. - GERTLER M., «Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory», *Quarterly Journal of Economics*, vol. 115, 2000, pages 147-180.
- COLANDER D.C., *Macroeconomics*, McGraw Hill, New York, 8th edition, 2010.
- DE GRAUWE P., «The Scientific Foundation of Dynamic Stochastic General Equilibrium (DSGE) Models», *Public Choice*, vol. 144, 2010, pages 413-443.
- GRAZIANI A., *The Monetary Theory of Production*, Cambridge University Press, 2003.
- GREENWALD B.C. - KOHN M. - STIGLITZ J.E., «Financial Market Imperfections and Productivity Growth», *Journal of Economic Behavior and Organization*, vol. 13, 1990, pages 321-345.
- GREENWALD B.C. - STIGLITZ J.E., «Keynesian, New Keynesian and New Classical Economics», *Oxford Economic Papers*, vol. 39, 1987, pages 119-132.
- , «Towards a Reformulation of Monetary Theory: Competitive Banking», *Economic and Social Review*, vol. 23, 1991, pages 1-34.
- HICKS J.R., «Method of Dynamic Analysis», in 25 Economic Essays in English, German and Scandinavian Languages, (1956), reprinted in HICKS J.R., *Money, Interest and Wages*, Harvard University Press, Cambridge Ma., 1982.
- , *Capital and Growth*, Oxford University Press, 1965.
- KEYNES J.M., *The General Theory of Employment, Interest, and Money*, (1936), The Collected Writings of John Maynard Keynes, vol. VII, MacMillan, London, 1973.
- , *The General Theory and After. Part I: Preparation*, The Collected Writings of John Maynard Keynes, vol. XIII, MacMillan, London, 1973.

- LUCAS R.E., «Expectations and the Neutrality of Money», *Journal of Economic Theory*, vol. 4, 1972, pages 103-124.
- LUNDBERG E., *Studies in the Theory of Economic Expansion*, (1937), Kelley, New York, 1964.
- MANKIW N.G., «Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly», *Quarterly Journal of Economics*, vol. 100, 1985, pages 529-539.
- MESSORI M., «New Keynesian Economics and Sequence Analysis», in MESSORI M. (ed.), *Financial Constraints and Market Failures. The Microfoundations of new Keynesian Macroeconomics*, E. Elgar, Cheltenham, 1999.
- MODIGLIANI F., «Liquidity Preference and the Theory of Interest and Money», *Econometrica*, vol. 12, 1944, pages 45-88.
- OHLIN B., «Some Notes on the Stockholm Theory of Savings and Investment I», *Economic Journal*, vol. 47, 1937, pages 53-69.
- PHELPS E.S., «Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time», *Economica*, vol. 34, 1967, pages 254-281.
- QUIGGIN J., *Zombie Economics. How Dead Ideas Still Walk Among Us*, Princeton University Press, 2010.
- SARGENT T.J., «Rational Expectations, the Real Rate of Interest and the Natural Rate of Unemployment», *Brookings Papers on Economic Activity*, vol. X, no. 2, 1973, pages 429-472.
- SHACKLE G.L.S., *The Years of High Theory: Invention and Tradition in Economic Thought, 1926-1939*, Cambridge University Press, 1983.
- STIGLITZ J.E., «The Causes and Consequences of the Dependence of Quality on Price», *Journal of Economic Literature*, vol. 25, 1987, pages 1-48.
- STIGLITZ J.E. - WEISS A., «Banks as Social Accountants and Screening Devices and the General Theory of Credit Rationing», in *Essays in Monetary Economics in Honor of Sir John Hicks*, Oxford University Press, 1992.
- TAYLOR J.B - WOODFORD M. (eds.), *Handbook of Macroeconomics*, North Holland, Amsterdam, 1999.

HOW “NEW” WOULD BE THE NEW TEXTBOOK APPROACH TO MACRO?

Alessia Amighini - Francesco Giavazzi

As it had repeatedly happened in the past, since the aftermath of the Great Depression, the recent crisis has raised a host of doubts about main-stream macroeconomic theories and inspired changes in the foundations of macroeconomics as well as in the way it is taught in the classroom.

Right in the middle of the crisis, a lively debate started among scholars about what, if anything, macroeconomic theory had missed and how to fix it (see among others Bernanke 2010, Blanchard 2010, Blinder 2010, Colander 2010 and 2011; Mankiw 2010). In mid-2009 *The Economist* magazine opened a discussion about what had gone wrong in modern macroeconomic theory – and therefore also in the teaching of macro – based on interviews with the authors of major textbooks (*The Economist*, July 18th 2009 and April 3rd 2010).

Already since 2010 the major macroeconomic textbooks had introduced brand new sections covering topics which until then had been neglected, in particular topics related to the functioning of financial markets. By 2011 – with the crisis behind us, although not its legacies – the teaching of macro around the world had been retooled through the introduction of new readings and up-to-the-minute presentations. Whether or not the current decade will eventually enter the history of economic thought as a new period of “high theory” is far too early to say. The debate around the state of macro has so far agreed on the fact that macroeconomists could and should have warned about the risks the world economy was running during the so-called “Great Moderation”, not necessarily that they were badly equipped to understand it. As a matter of fact, a few economists had shouted at the peril well before the beginning of the crisis, most notably Robert Shiller (2005) and Raghuram Rajan (2001).

Going back to our aim in these notes, to what extent this debate has or should have already made its way in introductory textbooks depends on whether each single issue can be taught to first-year students. The first European edition of Blanchard’s textbook (Blanchard, Amighini and Giavazzi, 2010) was among the first to include an entire chapter devoted to the crisis, and to adapt the core treatment of standard income/expenditure, IS-LM, and AD-AS models so as to be able to explain the origins of the crisis, how a financial crisis turned into a full-

edged economic crisis, and its legacies. By so doing, the new book revisited a set of topics which had become standard ingredients of introductory macro courses over the past 20 years. Among the major extensions that were added was – besides a section entirely devoted to Europe and the euro – an analysis of the crisis built on an extended version of the IS-LM model with a financial sector. This allowed to overcome the well known limitations of single-interest rate models in a very soft way - pedagogically speaking. Also, new concepts such as leverage, insolvency and illiquidity easily made their way into the text. In the same vein, Bofinger (2011) suggested that «it is relatively easy to reinterpret the basic model in such a way so that inconsistencies can be avoided». This has been made easy, in the context of Blanchard's textbook, by the pedagogical choices it has stuck to since its first edition:

- to organise the text around different time horizons since the very beginning, starting explaining how the economy works in the short-run and moving on to longer horizons later. This choice has made it easy to analyze the crisis distinguishing between the shocks, the immediate policy response and the longer term legacies;
- to put more emphasis on the mechanisms and policies that could bring the economy back towards equilibrium, rather than on the instability of the economy. Although aggregate demand shocks had hardly a place in the first editions of the text, the discussion of the recent crisis has given the opportunity to include such events.

Overall, Blanchard's approach has proved very successful also in that it can be expanded step-by-step to introduce more and more engaging and complex topics without losing consistency and confusing the picture – a pedagogical must. Extensions to the core chapters allow to deal with complex issues without making them unnecessarily difficult for undergraduate students. This is why students like the text, probably more than the most convinced among the teachers, and why it is one of the most appreciated introductory textbooks (Messori, 2012).

Ironically maybe, even the recent reactions by Brancaccio (2012) – provocatively titled as if they were at the opposite extreme of Blanchard's approach – were build on exactly the same standard framework, which suggests that the latter can be a common starting point for a discussion of competing ideas, theories, evidence.

Easy as it has been to adapt this textbook to explain the crisis, the question remains: where do we go from here? The next step, unfortunately, is far more dif-

difficult since it would require further integration of finance and macroeconomics. To dig deeper into the reasons behind the crisis one would need to introduce, as we said, concepts such as Leverage, Value at risk and similar. This is easily done in Mickey-mouse financial models but these are then difficult to integrate even in standard macro models. We have in mind the time-honored Holmstrom-Tirole model of financial intermediation, the Diamond-Dybvig model of bank runs, etc. One advantage of these models is that they illustrate in a transparent way some basic mechanism driving financial markets which can pave the way to macroeconomic shocks. Integrating these models in macro models – let alone in a standard textbook model - is a challenging task and today constitutes, in our opinion, the most challenging frontier of macro.

BIBLIOGRAPHY

- BERNANKE B., «Implications of the Financial Crisis for Economics», Princeton University Speech, 2010, <http://www.federalreserve.gov/newsevents/speech/bernanke20100924a> September 24.
- BLANCHARD O., «The State of Macro», *NBER Working Paper Series*, no. 14259, 2008.
- BLANCHARD O. - AMIGHINI A. - GIAVAZZI F., *Macroeconomics. A European Perspective*, Pearson Education, Essex, 2010.
- BLINDER A.S., «Teaching Macro Principles after the Financial Crisis», *Journal of Economic Education*, October-December, 2010, pages 385-390.
- BOFINGER P., «Teaching Macroeconomics After the Crisis», Wirzburg, *Economic Papers*, no. 86, 2011.
- BRANCACCIO E., *Anti-Blanchard. Un approccio comparato allo studio della macroeconomia*, Milano, Franco Angeli, 2012.
- COLANDER D., «Is the Fundamental Science of Macroeconomics Sound?», 2010, paper prepared for the ASSA Meetings, Jan, 2011.
- GARTNER M. - GRIESBACH B. - JUNG F., «Teaching Macroeconomics after the Crisis: A Survey among Undergraduate Instructors in Europe and the US», Universitat S. Gallen, *Discussion Paper*, no. 20, 2011.
- MANKIWI N.G., *Macroeconomics*, 7th ed., Worth Publishers, New York, (2010).
- RAJAN R., *Has Financial Development Made the World Riskier*, Federal Reserve Bank of Kansas City, 2001.
- SHILLER R.J., *Irrational Exuberance*, Princeton University Press, 2005.
- THE ECONOMIST, «The Other-Worldly Philosophers», July 16, 2009, <http://www.economist.com/node/14030288>
- .-, *Revise and Resubmit: The Crisis is Changing how Macroeconomics Is Taught*, March 31, 2010.

A COMPARATIVE APPROACH TO THE STUDY OF MACROECONOMICS

Emiliano Brancaccio

Will the economic crisis that broke out in 2008 and the ensuing crisis of the Eurozone lead to another revolution in the ideas of economists as regards the working of the market economy and the tasks of economic policy? A great many scholars have recently endeavoured to give an answer to this question. Some of them suggest that the so-called “mainstream” approach to macroeconomics has already been addressing the typical failures of the market that foster instability and recession for some time now. Economists should therefore be able to remedy any errors of prediction and assessment of the crisis by drawing upon studies already existing in the predominant literature (Tabellini, 2009). On this view, there would be no need to disrupt what Olivier Blanchard calls the “core” of the dominant macroeconomic theory and hence no need to rewrite the textbooks on which that core is based (Blanchard, Amighini, Giavazzi, 2010; for a specification of the basic elements of the core, see Blanchard, 2000, Chapter 30 and Taylor, 2000).

Though widely held among economists, this view appears in actual fact to overlook some problems that have recently emerged within mainstream debate on theory and economic policy. One example is provided by the doubts as to the relevance of analyses of economic policy that are based on non-observable magnitudes such as the “output gap” (Blanchard, Dell’Ariccia, Mauro, 2010). This question is in fact not only a practical matter. Although a pure “objectivist” approach to economic theory would be hardly conceivable, systematic reference to non-observable variables creates problems for the epistemological basis of the dominant theory (Leontief, 1991). Another example is provided by the contributions of some celebrated mainstream economists and major international institutions suggesting that the growth of income inequality can have depressive effects on aggregate demand and production in the long run (Fitoussi and Stiglitz, 2009; IMF-ILO, 2010). The simple fact that this interpretive stance has been adopted by influential mainstream economists calls into question the heuristic self-sufficiency of the prevailing macroeconomic theory, which rules out long-term causal relations from income distribution to aggregate demand and production. At the same time, there appears to be an evident similarity between this interpretation and some logical *schemata* of a heterodox character that have been

arousing renewed interest in the literature for some time now (Hein and Vogel, 2008; Stockhammer *et al.*, 2009). These and other elements therefore seem to indicate a weakening of the general consensus amongst economists that led Blanchard just a few years ago to note marked convergence at the level of mainstream method and theory, and hence to claim that «the state of Macro is good» (Blanchard, 2008).

Despite the spreading of cracks inside the core of the dominant approach, Marcello Messori (2012) rightly observes that, as things now stand, there is little tangible sign of any exchange of views between advocates of the different paradigms in the field of macroeconomics. This state of affairs does not appear to result, however, from any lack of alternatives to the dominant approach. In actual fact, there are rigorous programmes of research aimed at the development of competing paradigms both in macroeconomics and in the more general sphere of the foundations of economic theory, some of which have already reached a stage of conceptual arrangement (Pasinetti, 1977, 2007; Kurz and Salvadori, 1995; see also the preliminary work of Godley and Lavoie, 2006). The claim that alternative theories are the work of «sects of economists on the way to extinction», (Tabellini, 2006) therefore appears to be misleading. The problem is rather the almost total breakdown in communications between the various schools of thought that has existed for many years now. Suffice it to examine the number of times non-orthodox articles are cited in mainstream journals, once non-negligible but today close to zero. There are grounds for thinking that this form of separate development in watertight compartments has not enhanced the quality of economic research but, on the contrary, impoverished and damaged it in some respects. Blanchard himself points out the risk of «too much convergence» amongst economists, the resulting proliferation of technicalities of doubtful heuristic relevance, and the advisability of rehabilitating the use of «simple models» to examine once again the fundamental logical relations between economic variables (Blanchard, 2008). These are wholly acceptable observations. It appears difficult, however, to avert these risks and seize these opportunities in an historical phase when a certain degree of theoretical conformity seems to pay dividends also in terms of career while analyses devoted to the comparison of paradigms remain confined to a sort of no man's land.

While the claim that renewed interest in the comparative study of economic theories could serve to revitalise contemporary macroeconomics is thus hardly implausible, how can channels of communication be reopened between the various schools of thought after years with no exchange of ideas? How can the submerged

and forgotten comparative approach to economic theories be rehabilitated? One possible answer involves a return to a method of comparing alternative paradigms that was widely used during the 20th century. This method is based on the use of just one system of equations for all the theories examined; the transition from one theory to the other takes place simply through modification of the functional forms and reversal of the positions of exogenous and endogenous variables. The initial system of equations thus acts as a sort of *stereogram*: very different conclusions will be reached in terms of economic analysis and policy in relation to the viewpoint from which it is examined. While this method of comparison does not of course make it possible to examine the entire range of epistemological differences between the approaches compared, it presents the unquestionable advantage of immediacy by showing that apparently marginal changes in hypotheses can lead to completely different deductions. This is hardly surprising, as the choice of the exogenous variables, on which the method is based, is considered crucial by many to the correct specification of an economic theory (Dobb, 1973; Garegnani, 1990; Kurz and Salvadori, 2003; see also Brancaccio, 2010).

My short essay *Anti-Blanchard. Un approccio comparato allo studio della macroeconomia* is an attempt to return to the use of this particular method of comparing economic theories also in the sphere of teaching (Brancaccio, 2012). Despite its apparently challenging title, it makes no claim to replace the celebrated mainstream textbook of Blanchard, Amighini and Giavazzi (2010) but is designed rather to stand alongside it. The purpose is to show how simple modifications to the initial hypotheses of the dominant macroeconomic model can lead to substantial reversal of the logical relations characterising it. The aim in this sense is to provide a flexible teaching instrument aimed to reconcile the need to provide students with a preliminary grounding of the mainstream type with the need to foster rather than stifle their critical spirit. In this respect, the objectives of the book are more limited than those that characterized the textbooks of the Italian non-orthodox tradition (see for example Graziani, 2001, among many others). However, the results obtained so far appear encouraging. One of the reasons why students seem to appreciate the comparative approach of the *Anti-Blanchard* is the fact that it sheds light on the relations existing between theoretical hypotheses and the implications of the models in terms of economic policy, which the conventional teaching programmes usually overlook. The statistical appendix, based on a well-known OECD survey raising doubts as to the existence of a significant empirical relationship between systems of protection for workers and rates of unemployment, has also proved quite effective (Suppa, 2012). Albeit purely at the

level of an example, the statistical exercise offers interesting *stimuli* for theoretical reflection by calling into question the validity of the claim that less protection for labour would mean less unemployment, which is implicit in the dominant macroeconomic model.

Messori (2012) regards *Anti-Blanchard* as possessing various merits, one of which is clarification of the links between the short, medium and long run, which students can sometimes find somewhat hard to follow in the textbook of Blanchard, Amighini and Giavazzi. At the same time, Messori describes some of the criticisms of the dominant model put forward in the *Anti-Blanchard* as “empirically plausible” but based on «weak analytical justification». In particular, Messori indicates a weakness in the criticism of the standard inverse relation of monetary prices to aggregate demand. He points out that the debate on this subject has made great strides in recent years and suggests that the remarks made by Keynes long ago about the ambiguous effects of the flexibility of prices on demand do not provide adequate support for the view put forward. He then goes on to say that the weakness of the arguments against the mainstream inverse relationship between monetary prices and aggregate demand also weakens the other fundamental thesis of *Anti-Blanchard*, namely that the mark-up on costs can be influenced by monetary wage negotiations.

Messori's comments do not strike me as wholly acceptable. As a work designed for teaching purposes, *Anti-Blanchard* can obviously offer the reader no more than a very limited view of the evolution of debate on economic theory. However, its criticisms to current teaching refer to theses of great present-day relevance. Let us consider the ambiguity of the effects of price flexibility on aggregate demand. This subject has been broadly addressed in the mainstream literature in some advanced works that are characterised by precise micro-foundations and generally entail no particular assumptions about expectations. The differences with respect to the macro models based on canonical neo-Walrasian foundations regard the presence of market imperfections, information asymmetries and the heterogeneity of agents. On the basis of these hypotheses it is shown that greater flexibility of prices can accentuate the instability of the system rather than help to restore the equilibrium in the wake of a demand shock. This typically Keynesian subject, taken up again by De Long and Summers (1986); Hahn and Solow (1995); Eggertsson and Krugman (2010) and others, is now addressed also in the sphere of DSGE models (Bhattarai, Eggertsson, Schoenle, 2012). If there is an analytical weakness, it therefore appears to lie in the necessarily inverse relation of the level of monetary prices to aggregate demand found in the Blanchard, Amighini and

Giavazzi textbook and typical of mainstream teaching in general. It is also interesting to note that once this relationship is admitted to be uncertain, it is no longer possible to determine the mark-up indicated in the textbook on the basis of profit maximisation in imperfect competition (see for example Blanchard and Fischer, 1989, Chapter 8). If this buttress collapses, however, the mainstream idea of a mark-up unaffected by wage negotiations also proves analytically fragile. Contrary to what Messori suggests, it seems therefore in the conventional determination of the mark-up that an “ideological” component could be detected.

Though unusual in the context of teaching, the objections put forward in *Anti-Blanchard* to the predominant teaching cannot therefore be regarded as wholly extraneous to the mainstream literature. But, if things are in these terms, why should the book be labelled an example of alternative economic theory? To this regard, it is important to clarify that there are in fact no prior grounds for repudiating any interpretations seeking to reconcile the essay with the prevailing approach to macroeconomics. At the same time, it should be added that such readings would limit the scale of the criticism that can be derived from *Anti-Blanchard*. The reason is that this kind of interpretations would overlook the fact that the mainstream literature always rests ultimately on the “fundamentals” of the neoclassical economic theory. Suffice it to point out that if all the imperfections, asymmetries and heterogeneities were to vanish by magic into thin air, even the most advanced mainstream models would determine equilibrium on the basis of the traditional neoclassical triad of exogenous variables: scarce resources, preferences of agents and technology available. The problem is that this way of determining equilibrium leaves numerous criticisms unanswered on points such as the exogenous and non-observable character of individual preferences, the neoclassical theory of capital (Pasinetti, 2000; Kurz and Salvadori, 2008; see also Mas-Colell, 1989), and the methods of temporary and intertemporal equilibrium typical of modern neoclassical theory (Petri, 2004). The economists who consider these questions crucial may be able to take advantage of some buildings blocks recently developed within the dominant approach, such as the field of asymmetric information. They could also accept, for example, Messori’s very interesting suggestion of a possible influence of the trend of aggregate demand on the markup. They will, however, only consider analytical tools that can be transferred into a paradigm alternative to the neoclassical and based, among other things, on a different choice of exogenous variables and hence on a different general conception of equilibrium.

But which alternative paradigm? To this regard, explicit reference is made in *Anti-Blanchard* to the concept of “reproduction” typical of classical economists,

Marx and contemporary input-output and structural dynamics *schemata*, which today find numerous points of contact with the most advanced Post-Keynesian analyses. It is to this field of economic research that the book ideally belongs. Some specific evidence of this is to be found in the third chapter, which contains a simplified version of a recent alternative model of monetary policy theory (Brancaccio and Fontana, 2012). Through a “reversal” of the logical relations contained in the well-known mainstream “rule” formulated by Taylor (1993), the model suggests that the intervention of the central bank on interest rates is not necessarily designed to control inflation but to ensure the solvency of the economic system and hence, in more general terms, its reproducibility. We can consider it as one of many examples of the fruitfulness of a comparative approach to macroeconomic theory and its renewed possibilities of application in the teaching field. The *Anti-Blanchard* can be interpreted as a very preliminary attempt in this direction.

BIBLIOGRAPHY

- BHATTARAI S. - EGGERTSSON G. - SCHOENLE R., «Is Increased Price Flexibility Stabilizing? Redux», Federal Reserve Bank of New York, *Staff Reports*, New York, no. 540, January, 2012.
- BLANCHARD O., *Macroeconomics*, 2nd ed., London, Prentice Hall, 2000.
- .-, «The State of Macro», *NBER Working Paper Series*, no. 14259, 2008.
- BLANCHARD O. - AMIGHINI A. - GIAVAZZI F., *Macroeconomics: A European Perspective*, 1st ed., Pearson Education Limited, 2010.
- BLANCHARD O. - DELL'ARICCIA G. - MAURO P., «Rethinking Macroeconomic Policy», *IMF Staff Position Note*, February 12, 2010.
- BLANCHARD O. - FISCHER S., *Lectures on Macroeconomics*, Cambridge, Massachusetts Institute of Technology, 1989.
- BRANCACCIO E., «On the Impossibility of Reducing the Surplus Approach to a Neoclassical "Special Case". A Criticism of Hahn in a Solowian Context», *Review of Political Economy*, no. 22, vol. 3, 2010.
- .-, *Anti-Blanchard. Un approccio comparato allo studio della macroeconomia*, Milano, Franco Angeli, 2012.
- BRANCACCIO E. - FONTANA G., «Solvency Rule versus Taylor Rule. An Alternative Interpretation of the Relation between Monetary Policy and the Economic Crisis», *Cambridge Journal of Economics*, forthcoming, 2012.
- DE LONG J.B. - SUMMERS L.H., «Is Increased Price Flexibility Stabilizing?», *American Economic Review*, no. 76, vol. 5, 1986.
- DOBB M., *Theories of Value and Distribution since Adam Smith*, Cambridge University Press, 1973.
- EGGERTSSON G.B. - KRUGMAN P., *Debt, Deleveraging and Liquidity Trap*, Federal Reserve Bank of New York, New York, 2010.
- FITOUSSI J.P. - STIGLITZ J., «The Ways Out of the Crisis and a Building of a More Cohesive World», *The Shadow GN, Chair's Summary*, LUISS Guido Carli, Rome, 2009.
- GAREGNANI P., «Classical versus Marginalist Analysis», in BHARADWAJ K. - SCHEFOLD B. (eds.), *Essays on Piero Sraffa*, London, Routledge, 1990.
- GODLEY W. - LAVOIE M., *Monetary Economics: An Integrated Approach to Credit, Money, Income, Production and Wealth*, Palgrave-Macmillan, 2006.
- GRAZIANI A., *Teoria economica. Macroeconomia*, Napoli, Edizioni scientifiche italiane, V ed., 2001.
- HAHN F.- SOLOW R., *A Critical Essay on Modern Macroeconomic Theory*, Oxford, Blackwell Publishers, 1995.
- HEIN E. - VOGEL L., «Distribution and Growth Reconsidered: Empirical Results for Six OECD Countries», *Cambridge Journal of Economics*, no. 32, 2008.

- ILO-IMF, «The Challenges of Growth, Employment and Social Cohesion», *Joint ILO-IMF Conference*, Oslo, September 13, 2010.
- KURZ H.D. - SALVADORI N., *Theory of Production. A Long-period Analysis*, Cambridge University Press, 1995.
- .-, «Classical vs. Neoclassical Theories of Value and Distribution and the Long-Period Method», in HAHN F. - PETRI F. (eds.), *General Equilibrium. Problems and Prospects*, London, Routledge, 2003.
- .-, «Neo-Ricardian Economics», in DURLAUF S.N. - BLUME L.E. (eds.), *The New Palgrave Dictionary of Economics*, 2nd ed., Palgrave Macmillan, 2008.
- LEONTIEF W., «The Economy as a Circular Flow», *Structural Change and Economic Dynamics*, no. 2, 1991.
- MAS-COLELL, A., «Capital Theory Paradoxes: Anything Goes», in FEIWEL R. (ed.), *Joan Robinson and Modern Economic Theory*, London, Macmillan, 1989.
- PASINETTI L., *Lectures on the Theory of Production*, London, Macmillan, 1977.
- .-, «Critique of the Neoclassical Theory of Growth and Distribution», *BNL Quarterly Review*, no. 53, 2000.
- .-, *Keynes and the Cambridge Keynesians: A "Revolution in Economics" to be Accomplished*, Cambridge University Press, 2007.
- PETRI F., *General Equilibrium, Capital and Macroeconomics*, Cheltenham, Edward Elgar, 2004.
- STOCKHAMMER E. - ONARAN O. - EDERER S., «Functional Income Distribution and Aggregate Demand in the Euro Area», *Cambridge Journal of Economics*, no. 33, 2009.
- SUPPA D., «Appendice statistica», in BRANCACCIO E., *Anti-Blanchard. Un approccio comparato allo studio della macroeconomia*, Milano, Franco Angeli, 2012.
- TABELLINI G., «Osservazioni sulla nota di dissenso di Luigi Pasinetti», *CIVR Panel*, n. 13, *Consensus Group di Economia*, Appendice 5, 2006.
- .-, «Il mondo torna a correre. L'Italia non si fermi», in AA.VV., *Lezioni per il futuro*, Edizioni Il Sole 24 Ore, 2009.
- TAYLOR J.B., «Discretion versus Policy Rules in Practice», North Holland, *Carnegie-Rochester Conference Series on Public Policy*, no. 39, 1993.
- .-, «Teaching Macroeconomics at the Principles Level», *American Economic Review*, no. 90, vol. 2, 2000.

British and Italian Universities and the Anglo-Saxon Model

Alan R.S. Ponter*
University of Leicester, UK

In the '80's and '90's, British Universities were subjected to a sequence of changes that transformed their organisation and management. There are clear parallels with the challenges presented at the present time to Italian Universities. With discussions of the implementation of the so called Anglo-Saxon model to Italian Universities the paper argues that the particular characteristics of change within the British system may only be understood through the British Institutional structure, wherein outcomes arise from extended discussions between the Universities and Government. The history of change since 1980's is summarized and recent changes are also discussed.

[JEL Classification: H52; I23; I24; I25; I28].

Keywords: universities; organisation; institutions; policy; British; Italian.

* <asp@le.ac.uk>, Emeritus Professor.

1. - Personal Introduction

Perhaps I should first mention something of my background. During the difficult Thatcher period, which I discuss below, I served as President of the Association of University Teachers (the academic's trade union) at the University of Leicester. Later I became Pro Vice Chancellor of the University.

2. - Introductory Remarks on British Universities

Close comparisons of the way Britain and Italy run their Universities is interesting and confusing. British Universities have always been independent institutions free to recruit students, appoint academic and administrative staff, carry out research and find leadership. All academics are employees of their university and not employees of the state. Universities are Institutions under the Crown, allowed to carry out activities for the State, set out in their Statutes. International studies indicate that they remain, overall, the most autonomous set of Universities in the world. There are claims that our universities are the most cost effective in the world. Considering the difficulties of such comparisons I suspect they can only be made with Universities in countries with similar systems, *i.e.* Canada, United States, Australia and New Zealand and average salaries is an issue. With the exception of a single small private university, they are all state universities. There have been discussion of whether particular Universities can become private Universities but this never seems to develop beyond an ideological aspiration.

3. - The Thatcher Period

The beginning of the modern story of UK universities can be conveniently designated as 1980, the year when the newly elected Thatcher government imposed severe cuts in funding. This was a shock to a system ill prepared for such an eventuality. There may be some parallels with your circumstances today. The cuts were partly justified by the observation that the student cohort was on a downward curve due to reductions in birth-rate. But government was as much concerned with efforts to radically reduce public spending. Initial plans for a reduction in student numbers change in 83/84 to talk of expansion as unemployment rose sharply, particularly amongst young people. A severe recession and the de-regulation of the City of London in '86 set in train a seemingly irrevocable

change from manufacture to service industries. Such a change required a higher number of graduates. These years, 1980 to 1990, involved painful changes. University autonomy had come at a price in overall planning. There were overdue reforms overlaid with Thatcher cuts. University education was expensive by international standards. A primary issue, recognised before 1980, involved too many small, barely viable departments and efforts by many universities to give degrees courses in subjects with a low student demand. A need for changes in Engineering had been recognised in the Finniston Report of 1978. The need for expensive facilities in science suggested fewer departments in “Centres of Excellence”. There was an issue about the management of the Polytechnics, formed from groups of local trade and professional schools and funded by local rather than national government. There had been a lack of overall planning, exacerbated by unrealistic expectations of student expansion in particular subjects.

4. - The Process of Change

The process of readjustment was painful. National schemes of funding allowed amalgamation of departments, transfer of staff between universities and reviews of subjects, leading to plans for change over a number of years. No university was closed, many emerged stronger. Some contracted significantly. Attempts at amalgamation of immediately adjacent universities (e.g. Leicester and Loughborough) were much discussed but did not occur. A merging of University College and the University of Wales Institute of Science and Technology in Cardiff was forced by government as the former had driven itself into bankruptcy by careless financial management. Eventually, many years later, the two Universities in Manchester agreed to combined, resulting in a much strengthened institution. Broadly, such changes were aimed at benefiting the service the university could provide to students and enhancing the academic environment for staff research, all at reduced cost. Various market led schemes were discussed but not implemented. But such changes that did occur were not always advantageous for individual academics. There were many early retirements.

5. - Early Retirement

Because of tenure, invitations for retirement were extended to categories of academics, initially all those over 60 (the normal retirement age was 65) although

there was a criterion of “managerial interest”. For example, if the academic held a position that the University needed to immediately replace, the argument for early retirement was seriously weakened. University staff had (and still have) their own independent retirement scheme and, in the 1980’s it could absorb early retirements with some government supplementation. Other selective invitations to retire early took place, with fairly generous conditions. The number of retirements was such that recruitment of young (less expensive) staff continued, producing an age distribution of staff that was healthier than previously. Promotion of staff continued on merit without regard to cost. Many academics continued teaching after retirement on contracts with payment for hours taught.

6. - Later Changes - The Creation of New Universities

Finally in 1992 the process of turning the Polytechnics into independent universities with degree giving powers began. Virtually all tertiary education including teacher training, nursing and other paramedical discipline training was either taken over by a local university or combined to form new universities. From about 40 universities in 1980 the number has expanded to over 100 today. When I went to University only about 5% of my cohort were able to go to university whereas now it is closer to 40% and is now likely to stay at this level. During this period policy was driven by a desire to raise the average quality, reduce costs but to ensure that the very best remained so. A primary concern, within the older universities and amongst many ministers, was the prospect that a market system would lose the international standing that predated 1980. Indeed there are arguments that the great concentration of Nobel Prize winners found in such concentrations of talent as the Cavendish Laboratories in Cambridge may become a thing of the past. Remember that Thatcher is the first Prime Minister and graduate of Oxford who has never been awarded an honorary degree by Oxford. Many of the changes that took place during these years were necessary but not at the ruthless pace they were driven through. But problems of overall planning persisted. Universities keen to expand produce degree courses popular among school leavers who are not always well advised (or advisable). This had occurred in the 1960’s when courses in Sociology expanded rapidly only to collapse by 1980 where large departments found themselves with few students. A similar expansion in Media Studies has occurred recently, producing graduates for which there never had been much prospects of a job related to their studies. Difficult subjects with

a high demand for graduates, such as Languages, Science and Engineering were under subscribed. At the same time the prospect of higher salaries in the service industry attracted highly numerate students to less demanding courses in Management Studies and Economics. This was within a system where opportunities to change course, once begun, is limited. Most Universities have grouped Departments into Schools where degree courses are given by staff from a number of Departments. This has helped generate new “themed” courses and protected universities against fluctuations in student demand.

The turning of the Polytechnics into Universities remains controversial as many have tended, not always successfully, to emulate the aspirations of the old Universities. Their traditional role, practical preparation for work within the local, rather than national, economy has been eroded, with debatable advantage to either.

7. - Recruitment of Staff and Students

An aspect that clearly differs between our systems is the recruitment of both staff and students. All our Universities are residential and it is now an expectation that the school leaver will go to a University some distance from their home. This arose from demography; a high proportion of the population live in the small cities, towns and villages. There are certainly patterns of recruitment and the Scottish Universities still find most students from their own citizens, but the popularity of Exeter, Bath and Bristol Universities with the sons and daughters of the London professional classes has been attributed to Brunel’s railway and their distance from London, still in England and not North. University towns with a reputation for a good student life, such as Newcastle, are popular. Indeed, and this contrasts most strongly with Italy, cities in industrial recession where the cost of accommodation is low have an advantage. Declining industrial cities such as Leicester, Sheffield, Manchester and Newcastle have become University Cities where the presence of strong universities provide the best hope for economic recovery and are seen as such.

As staff recruitment is a responsibility of each university, movement of staff around the country and from abroad is normal and junior staff seek senior positions in a university with the highest reputation and a mature interest in their expertise. This explains the concentration of excellence in Oxford, Cambridge, London, and, increasingly, Manchester and it is no surprise that this is where the best students also wish to congregate. But many fundamental research discoveries

have occurred elsewhere: the first cloned mammal (Dolly the sheep) in Edinburgh, the discovery and first application to forensic practice of genetic fingerprinting in Leicester and, again in Leicester, the first experiment that identified the exact position of an intense X-ray source (*i.e.* a Black Hole).

8. - Tenure

There was a major battle over tenure during the Thatcher years when universities needed to remind government that a reduction in funding of more than about 2-3% a year was only possible by dismissing staff, which tenure prevented. After much discussion, Statutes were eventually changed to include “financial exigency” as a reason for dismissal, *i.e.* dismissal if there were no funds to pay staff. But satisfying the criteria for dismissal is difficult to argue. Generous voluntary early retirement schemes were, eventually, the primary mechanism for cost reduction following the Thatcher cuts. The case law of the subsequent years indicate that tenure is as secure as before. Certainly there is an increased proportion of academic staff on fixed term appointments resulting, for example, from the uncertainty in overseas student numbers and other time limited funding. Research staff is generally on fixed term appointments and this has always been the case.

9. - Patterns of Management

Our system is a product of history but there has always been, since the time of Charles I and then Cromwell in the 17thC, a suspicion of overly centralised power. Our inability to run Collectivism efficiently after WW2 and the memories of the economic and social problems of the 1970's has strongly influenced views. For universities this has resulted, effectively, in a system of contracts where the government contracts individual universities to educate prescribed numbers of students in certain broad subject areas at a fixed cost and provides background support for research by a mixture of specific and general grants. Support for staff research time comes from funds resulting from the well known Research Assessment Exercise (RAE) where the research quality and productivity of all departments is assessed every 4-5 years with funding concentrated in Departments with the highest standing. Strong independent research universities form the core of our system are highly valued and always have been. In the recent spending cuts, funds for university research have been protected. There is an acceptance that

this is essential for a country with few natural resources. The RAE has resulted in an increased competition for the best staff, whether from within the UK or abroad, particularly the EC. For example, in Mathematics recruitment from Germany and Russia has been particularly strong. Support for research workers is gained by applying, competitively, for funds from Research Councils and various charities such as the Leverhulme Trust and from industry. In my own area, Engineering, it is quite common to have a project with funding from more than one source.

10. - The Role of the Professor

The demands the evolved system places on senior staff has resulted in an acceptance that effective leadership is required. The number of Professorships has increased. Individual universities make such appointment independently, assisted by external assessors who are leading professor from other universities. An academic with an international reputation and a recognised contributor to his subject would normally expect to be given the title Professor. Certain universities, Oxford is an example, have a clear distinction between Established Chairs that have distinct leadership roles and Titular Professorship, which is given for the academic's personal achievements but without specific responsibilities and does not necessarily command an increased salary. The offer of a Professorship to a young and very talented academic can be enough to persuade movement from one university to another. In my own Department of Engineering at Leicester University there were two professor in 22 academic staff in 1980 whereas today where there are 10 Professors amongst a total of 30 academic staff. Typically a third of staff would now be Professors in a research intensive Department in the top 20 Universities. At Oxford the proportion in the Department of Engineering Sciences is one half (36 Professors within 72 academic staff).

11. - The Importance of Assessment Information

The retention of autonomy has resulted in schemes of assessment that allow all potential university partners and potential students to understand the qualities and capacities of individual departments and universities. The results of the Research Assessment Exercise (RAE), Teaching Quality Assurance Agency reports, Student Satisfaction Surveys, data on entrance qualifications and degree results,

research funding, qualifications of staff, expenditure on libraries etc. are all freely available. If you Google “assessment of UK universities” you find a number of university quality surveys, mainly done by newspapers, variously produced from this published data. Such surveys influence student choice. But for Italian Universities I find only the international lists; competition for students is less of an issue when the student expects to attend their local university.

12. - The Anglo-Saxon Model

The extent to which the so called Anglo Saxon model of a market led economy has been applied to universities is debatable. There is open competition for students but only on the criteria of academic ability. There is vigorous competition for good staff and internal pressure on those who fail to perform well. Competition for research funds that allow fundamental work is intense. The best universities protect staff whose research interests may not be the flavour of the month, but there is a general expectation that each research group will find its own research funds. But the present government has begun to introduce new elements that have a stronger market element through new schemes for student fee and student grants.

13. - Recent Developments: The Student Loans Scheme

The new student grant scheme results in a high proportion of the cost of a student's tuition and living costs provided as a loan, repayable over their subsequent career. The tuition fee element is a fixed sum for each university, supplemented by direct grants to the university to make up for the variation in subject costs. As there is no supplement for Humanities, Social Sciences and Mathematics degrees, it could be said that the fee level is set at that of the courses with the lowest cost. For students from poorer backgrounds there are grants. On graduation, for those on wages less than the national average wage, no repayments are required until their salary rises above this threshold. Any remaining debt is cancelled after 25 years (increasing next year to 30 years). The scheme is complex and amounts to an attempt to introduce an entirely novel concept, a personalised taxation system collected through the tax scheme. UK citizens pay their taxes. The argument in favour attempts to show that the higher earnings of a graduate more than compensates for this higher tax. The scheme is a radical alternative to existing schemes,

a Graduate Tax as practiced in Australia and conventional student loan schemes run by banks. The scheme resulted from a long debate about who should pay for a mass higher education system, the state through general taxation or the beneficiary, the student. Objectors from both the Left and Right, for their own reasons, had pointed out that the general taxation method resulted in those unable to access higher education were making the major contribution to the education of the middle classes. Indeed there was much evidence that where you lived and the income of your parents had become, increasingly, the criterion for access to university. The argument continues that having a system that effectively took a major component of the cost away from general taxation, more funds could be spent on correcting the increasing diversity of educational opportunity in schools. Well, that's the theory and I rather suspect it would not have been introduced in this form except for the radicalism of the Liberal Democrats in government and the need to restructure national debt. Already there are signs of a reduction in applications from mature students. There are accusations of imposing a utilitarian view of higher education and that students will be discouraged from taking degrees in the Humanities as they accumulate the same debt as, say, a graduate in Engineering, with possibly lower salary expectations. This means, of course, that they are likely also to repay less. We will have to wait and see.

14. - Recent Developments: The Student Fee Scheme

The other component was the offer to the Universities of the opportunity to set the fee component, with a maximum of £9,000 (10,700€). A university would have the prospect, if they offer a lower fee, of additional student numbers. Students would be attracted by a lower debt at graduation. Universities have always operated under an arrangement that the fee the state pays for a student at Oxford should be the same at a less illustrious university and it was no surprise that when the time came to declare their fee, most chose £9,000 or nearly so. The scheme is generally regarded as a serious policy mistake on the part of the government and, at the present time, the resolution is not clear. I would anticipate that, eventually, some of the post 1992 universities may reduce their fee, but there is a general revulsion about the idea of a market for cheap courses.

15. - The Universities within British Society

Britain finds itself with a highly constrained and heavily regulated higher education market where free market principles are allowed where they are perceived to encourage efficient delivery and student choice without sacrificing either quality or equality of access. At the same time the needs of society as a whole must be protected, both culturally and economically. Whether this is achieved is open to question but it has certainly encouraged a more adaptable and dynamic system. How this is perceived by an individual academic is very variable. The management style has become much more managerial and less collegiate.

Undoubtedly, the free market reforms of Thatcher and Blair have produced a wealthier but materially less equal British society. I like to believe that a consensus view continues (except on the far Right of the Conservative Party) that equality of access is a citizen's right in two areas, education and health, and that it is the responsibility of government to maintain this through policy. This contract between government and people was forged in 1945 and has survived despite Thatcherism, but is a constant source of conflict. There is no doubt that the greatest impediment to access to higher education is the variability of school provision in the state sector and this has proved very difficult to change, despite significant investment during the Blair years. The break-up of the extended family seems to be a major factor and I perceive that this is not a problem in Italy. The advantage of private education is an anomaly that is difficult to dislodge in a free society. Private (boarding) school fees are typically £33,000 (39,300€) a year, about 1.5 of the average wage. The aim has been to make state education in schools as good as private education but this has not been achieved sufficiently widely.

16. - The Anglo-Saxon Model and the Italian Universities

So if your government wishes to apply the so called Anglo-Saxon market led model to your University, then they should do their homework, it just isn't that simple. Certainly the obvious difference in structure of British and Italian society makes many of our practices irrelevant. Concerning your current North/South issue, I read the address of Prof. Ing. Nicola Costantino, Rector of the Politecnico of Bari, with great interest and admiration, particularly the following:

«Today, many speak of young people, their ambitions and their hopes: we work with our young people every and all day, together to build a future that is worthy of their expectations. And for this reason we contest that the model of

“market University” which would result in a failure to provide, to the poorer regions, the right to high quality training».

I honestly believe you would have difficulty in finding a British academic or, for that matter, a government minister who would disagree with Rector Costantino. A policy of cuts that clearly discriminates against the Universities of the South in Italy violates this absolute: that equal educational opportunity is a right of all citizens. It seems to me, as policy, if pursued, it is likely to divide the nation. This is intellectually lazy policy making. Such policies were a feature of the later Thatcher years and contributed to her demise.

17. - Open Debate was the Key

The 1980's and 1990's were uncomfortable decades for British universities but they have emerged as more efficiently run with, on average, improved standards. A primary characteristic of this period was the intense level of debate and an acceptance that change would occur and, in many areas, should occur. The national forums of the time, the Higher Education Funding Council (which funds) the Committee of Vice Chancellors and Principals (which debates and agrees across universities), the Association of University Teachers (the trade union) and their local Associations, Professional Accrediting Bodies, subject Associations, and the pages of the Times Higher Education Supplement all have had a role in forging a national consensus within the universities. Local disputes were exhaustively argued at *ad-hoc* meetings and at University Senates. Many market views were dissected to oblivion in the process.

18. - Conclusion

I hope this brief summary will be sufficient to show that the notion of the Anglo-Saxon free market model came directly into conflict with an older and more fundamental understanding of the role of the university. Britain has always been very proud of its traditions and achievements. This of course, we have in common with Italy. In the years prior to 1980, Britain had begun to lose its confidence and Margaret Thatcher has to be seen in that context. She spoke to the anxieties of most ordinary people and was willing to act decisively. She attacked vested interests with vigour. It was her brazen devotion to change that was so breathtaking and shocking. But the processes of change that her government ini-

tiated would, eventually, allow our universities to occupy a more significant role in British life.

19. - Acknowledgement

The author extends his thanks to Dr. Michael Ciavarella of the University of Bari. The contents of this paper arose from our private discussion and subsequent discussions on the internet forum iMechanica.org. A form of this paper was presented for the author by Dr. Ciavarella at the conference “The Future of Italian Public Universities” held at the Polytechnic University of Turin in March 2012.

Preventing and Resolving Banking and Sovereign Crises: Toward a New Institutional Framework

Jean Tirole*

Toulouse School of Economics

Weak European institutions bear some responsibility for the current Eurozone crisis. The paper discusses various alleys for reforming European institutions, such as those in charge of monitoring countries and banks. It then discusses the future of solidarity within the Eurozone, as well as proposed reforms impacting risk sharing. It concludes with steps that might help restore confidence in the European Monetary Union.
[JEL Classification: E62; F34; H63].

Keywords: European debt crisis; golden rules; banking regulation; joint liability; bailouts; private sector involvement.

* <jean.tirole@tse-fr.eu>. This paper is the text of a lecture delivered by Jean Tirole on the occasion of his being conferred an Honorary Degree by the University of Rome 2 (“Tor Vergata”) on May 18, 2012.

1.- Introduction

The Eurozone faces a dual debt-and-competitiveness crisis. This crisis has been facilitated by the weakness of European institutions.

On the Dual Nature of the Crisis: Southern European countries have jeopardized their future both through the accumulation of large debts (official or hidden such as pension liabilities) and through their failure to implement the structural reforms that would make it credible that this debt will be reimbursed in the future.

Our countries have also let wages increase faster than productivity, and seen their competitiveness substantially reduced over the last decade. While the Eurozone is roughly in external balance, Southern European countries lost from 10% to 40% competitiveness relative to Germany over that period. As much trade is intra-European, substantial current account imbalances in Southern European countries emerged, leading to a concern about the sustainability of their debt.

Weak Institutions: The EcoFin Council, which is in charge of monitoring member countries' deficits, does not exert sufficient pressure. In contrast with, say, the IMF, the EcoFin council is too political and has a conflict of interest. As an outcome, no sanction has ever been applied despite 68 outright violations of the SGP.

The political bias toward *laissez-faire* is understandable:

- first, there is little for a country to gain by insisting on imposing discipline on another country that does not comply with common rules; the former is unlikely to be pivotal in the 17-member Eurozone decision-making process while it can get into the latter country's bad books by taking an adversarial stance.
- second, political benefits may dominate reputational concerns; a case in point is Europe's deliberate ignorance of accounting gimmicks (documented by Eurostat and others) in order to allow some countries to join the Eurozone.
- third, "enforcers" may feel that they will be granted a similar favor when their turn comes. Interestingly, France and Germany themselves violated the rules in 2003.

Regarding sanctions, let me add that I view financial sanctions, as adopted in the Maastricht Treaty and picked up in the recent fiscal compact, as generally inappropriate: they increase financial pressure at a time at which the country is already in financial straits. Rather, and as is indeed standard practice for IMF programs (as well of the "Troika" -EC, ECB and IMF- programs in Europe), troubled countries should see their sovereignty reduced rather than face monetary

sanctions: their budget should be more carefully monitored and structural reforms promoting competitiveness required, while protecting the poorest inhabitants.

European institutions also have had a bad record in the current crisis:

- attempts at shifting the blame (blaming rating agencies and speculators; to be certain the latter speed up crises, but blame shifting does not enhance investor confidence),
- lenient stress tests for banks (for example, Dexia was deemed solvent just before it defaulted),
- loss of credibility stemming from some confusing announcements proclaiming an inflexible determination to honor fully sovereign commitments and ruling out private sector involvement prior to Greek-debt restructuring on July 21, 2011. Furthermore, announcements have often been made before authorities had drawn a precise plan making these announcements credible.
- buying time. All Eurozone meetings- at least in their resulting decisions- have been mainly preoccupied by the short-run goal of avoiding an imminent collapse (there is here a striking analogy with Daniel Ellsberg's account of US policymaking during the Vietnam War).

2. - A Broader Remit for the Stability Pact?

It has long been understood that Europe, with its limited labor mobility and its *quasi*-absence of an automatic redistribution mechanism (a federal budget), is not an optimal monetary zone. It seems unlikely that fiscal federalism will come about, as it would presumably involve large and predictable transfers across countries.

Short of fiscal federalism, European countries will still have to accept a substantial loss of sovereignty if they are to continue living together. Fortunately, there is still some scope for making institutions more compatible with the existence of a monetary union.

First Alley for Reforms: Fiscal Rules and Independent Fiscal Councils

There has been much discussion lately about using balanced budget rules or "golden rules" specifying a (cyclically adjusted) target for the budget deficit. Among other things, the March 2, 2012 "Fiscal Compact", signed by all of the member states of the European Union (EU) except the Czech Republic and the United Kingdom, calls for

- a binding golden rule (specifying a maximum primary deficit – *i.e.* gross of debt service – below 0.50% over the cycle),
- automatic sanctions (reverse qualified majority voting), and
- enforcement by the European Court of Justice.

Another interesting aspect of this fiscal compact innovation is the relaxation of the unanimity rule that so far had hampered European efforts in fiscal matters; namely, a minimum of 12 out of Eurozone's 17 member states have to ratify the text for it to come into force.

To be certain, golden rules raise complex design and enforcement issues and are no *panacea*. Capturing debt sustainability through a single cap requires converting contingent liabilities and revenues into a single number. In theory, investments that will improve tax revenue or the current account in the future should not be treated as current consumptions and therefore not be subject to cash accounting (that is, one should not charge investments to the budget). In practice, however, separating the investment component and the consumption component of public policies is awfully difficult.

Defined-benefit pensions loom particularly large in the mis-measurement of public debt (both in Europe and in the US). They usually are off-balance-sheet and thus not treated as public debt on the ground that they are only contingent liabilities. To be sure, and as we observe now in some Southern European countries, pensions can be reduced; the scope for adjustment remains rather limited, though.

Debt sustainability is a complex phenomenon and the detection of accounting tricks requires expertise; regardless of the institutions in charge of enforcing them, golden rules require the public sphere to boost its economic analysis capability. The need for technical expertise and the prevention of manipulations calls for the creation of independent fiscal councils endowed with professional bodies.

Another key factor for the success of golden rules is a political and popular support for budget discipline. Countries such as Sweden, Germany or Chile benefit from a broad consensus in the matter and reforms instituting golden rules, independent fiscal councils and other checks on fiscal discipline, have by and large been *bi-partisan*. A serious concern in countries like France and some other countries in Southern Europe is that golden rules are the objects of maneuvering and clashes within a political establishment rather unkeen to see its prerogatives reduced. This absence of bi-partisanship, together with the recent concerns about the US (driven mainly by a political inaptitude to come to terms with budget realities), serves as a reminder that institutions cannot deliver optimal outcomes unless they gather a minimum level of political consensus.

Should the Eurozone mandate the adoption of balanced-budget rules? The answer is “yes”. Obviously, the credibility of such rules would be much enhanced if the Eurozone countries appropriated these rules (ownership) as they did in the US. Yet, the spontaneous adoption of rules in the US in the mid-19th century followed a strict no-bailout policy, which in Europe is unlikely.

Second Alley for Reform: Better Banking Regulation

There are of course many banking reforms to be undertaken (and some of them are in the process of being undertaken); I have no time to cover these, and will focus on an institutional one: the need for regulating the financial institutions at European level.

Despite the creation of several European-level institutions (such as the European Banking Authority and the European Systemic Risk Board), financial supervision and default resolution still operate according to the country-of-origin rule: national regulators supervise the financial institutions chartered in their own country and their branches abroad¹. The European Banking Authority (created in 2010) for example cannot substitute for a full-fledged European banking regulator. The Authority focuses on the “single rule book”, has a limited staff, and relies on national regulators for information.

There are at least two rationales for centralizing prudential regulation at the European level and for creating an independent supervisory authority. First, it is unlikely that all 27 financial supervision authorities in Europe (17 in the Eurozone only) be sufficiently well-staffed to match the sophistication of private banks.

Second, and more to the point for the focus of this talk, there is growing awareness that private debt is now public debt; the banking fragility following the construction booms in Ireland and Spain is a case in point; Ireland and Spain did not have extravagant public deficits and public debts to start with. But banking bailouts threatened to take a heavy toll on public finances. More generally, many exchange rate and sovereign debt crises start with a credit expansion and a real estate bubble that authorities treat leniently.

Concomitantly with the allocation of the supervisory task to a real European Banking Authority should be created a European Guarantee Fund that would guarantee retail deposits across Europe; as for the Eurobonds, which we will later

¹ *N.b.*: After this speech was delivered, Eurozone leaders decided (on June 29, 2012) to create a joint banking supervisor for the 17-nation common currency area, and to locate it at the European Central Bank. Most of the crucial details remain to be worked out.

discuss, a key difficulty here is that we do not stand behind the veil of ignorance and a European Guarantee Fund would imply large transfers toward Southern European countries. Presumably the negotiation will separate the treatment of legacy issues from the sharing of future bank risk. Similarly the institutions for resolution will be hotly debated.

3. - Solidarity within the Eurozone

The second focus of this lecture is the current debate on the solidarity among Eurozone countries, which has taken center stage in the policy discussions about the current crisis. To this purpose, let me start with some background remarks:

Insurance/Moral Hazard Tradeoff

Countries understandably want to be insured against adverse shocks; on the other hand, insurance reduces accountability and countries can avail themselves of many ways to “expropriate” foreign investors: through default, devaluation, or extractive activities (taxation, straight asset expropriation). Policies leading to such expropriations fall into two groups:

- policies leading to a lack of competitiveness: inefficient labor laws, investments in non-tradable (real estate typically) rather than competitiveness-enhancing investments, protection of specific professions...;
- policies leading to public debt concerns: high public spending and low taxes – or insufficient tax collection efforts –, lax banking supervision, failure to reform the pension system...

This familiar trade-off between insurance and moral hazard is complicated by the presence of liquidity crises. Indeed, economists distinguish between liquidity and solvency problems:

- a solvency problem arises when the primary deficit and competitiveness problems make the path of sovereign debt unsustainable.
- a “pure liquidity problem” arises when a country is actually on a sustainable path, but self-fulfilling realizations force it to pay very high interest rates, making its debt grow fast and in the end making it indeed difficult to reimburse the debt. The liquidity view of sovereign crises thus emphasizes self-fulfilling realizations of insolvency; the country’s fundamentals are such that the debt could

be sustainable, and indeed there is another, unrealized market equilibrium in which the country maintains an easy access to the international debt market.

Whether a country is suffering from a liquidity crisis or from a solvency problem is often hard to tell; both lead to high interest rates and ultimately a lack of access to international capital markets. Unfortunately, desirable policies depend radically on the diagnostic. Illiquidity problems call for widespread guarantees against country default, while insolvency problems require interference with country policies and acceptance of a potential default.

Allocation of Country Risk between the Market and the International Community

Who, of markets and the official sector, should bear the burden of default? In the current context, who foots the bill also determines who monitors:

- if *markets* are to monitor countries by removing their access to financing when bad policies are selected, then the private sector should be made accountable: banks should bear the burden for losses they incur when lending to the country; and prudential regulation should abandon the Sovereign risk-free tag; that is, it should treat risky sovereign debt as a risky asset when computing capital adequacy requirements.
- if the *official sector* is in charge of controlling debt sustainability, then it should foot the bill and also intervene whenever a country's indebtedness path may not be sustainable.
- none of these alternatives prevailed prior to the Eurozone crisis. Regardless of one's view as to who (international community or market) is a better monitor, one will agree that the recent no-man's land has proved rather unsatisfactory: governments have imposed no discipline on each other and markets have long thought that their lending to weak sovereigns would go unpunished. Put differently, no-one felt accountable.

Solidarity

Solidarity among countries raises at least two issues: the choice of mechanism through which this solidarity operates and the solidarity area.

As for the mechanism, should solidarity take the form of:

- *ex-post* bailouts (the dominant approach hitherto)?
- a mechanism in which Eurozone countries are jointly liable for each other's default, for example through Eurobonds or a banking union?

- a Eurozone stability fund with well-defined rules for access?
- or alternative cross-insurance mechanisms (taking the form of liquidity provision to countries), that do not necessarily involve insurance among countries within a monetary zone? For example the IMF's Contingent Credit Line and its replacement, the Flexible Credit Line. The Chiang Mai Initiative, a multilateral currency swap arrangement pooling \$120 B and launched in 2010, encompasses the ten members of the Association of Southeast Asian Nations (ASEAN), the People's Republic of China, Japan, and South Korea. This scheme involves insurance among countries with different currencies, unlike the Eurozone scheme.

As for the solidarity area, that is the allocation of risk within the official sector, it is generally assumed in the European context that other Eurozone countries are the natural providers of insurance. This assumption, which is reflected in negotiations and current bailout policies, is at first sight startling. After all, insurance economics points at the desirability of spreading risk broadly, and not allocating it to a small group of countries, which moreover may well face correlated shocks.

It is fair to say that the economics of country solidarity are still in their infancy. Nonetheless, I would like to venture into a few considerations, based on recent research².

Bailouts are driven by the fear that spillovers from the distressed country's default negatively affect the rescuers. Negative externalities from default can be

- economic: reduced trade, subsidiaries and banking exposures, run on other countries,
- or other: empathy, jeopardy of European construction, distressed country's nuisance power.

Collateral damages of a country's default are *de facto* collateral for the country. In this sense, countries that are deeply intertwined within a monetary zone may be more natural providers of insurance than less connected countries. Consider the solidarity area "puzzle": while insurance economics teaches us that risks should be spread widely rather than be shared among a small number of countries (facing somewhat correlated shocks to boot), solidarity in the recent European crisis has

² TIROLE J., *Country Solidarity, Private Sector Involvement and the Contagion of Sovereign Crises*, mimeo, 2012.

involved essentially Eurozone countries, not even other European countries. Eurozone countries however are more exposed to each other's default, for several reasons: first, there has been a substantial expansion of financial integration within the Eurozone; second, the Euro project was in large part a political vision, and abandonment of this project is perceived to be very costly.

Joint-and-several liability mobilizes further collateral. In effect, a country's default becomes its guarantor's default if the latter fails to abide by its obligation to stand by the defaulting country. So the guarantor's incentive to pay is larger than in the absence of joint liability. But joint liability also makes domino effects/systemic risk more likely: if the guarantor country does not abide by its pledge to rescue the distressed country, the guarantor country itself is in default.

Standard liquidity provision or risk sharing models presume that accord is reached behind the veil of ignorance. Indeed, it can be shown that joint-and-several liability emerges in an environment in which

- a) countries are in symmetrical situation, with similar probabilities of facing hardship- they are then less concerned about cross-subsidies and more willing to enter into a mutual insurance agreement;
- b) these countries face shocks that are relatively independent; and
- c) spillovers from default are substantial.

However, the veil of ignorance fiction is unrealistic for Europe, and the suspicion of a cross-subsidy from Northern to Southern Europe may thwart efforts to reach a consensus on these proposals. Indeed the same research shows that, once the veil of ignorance is lifted, healthy countries have no incentive to accept obligations beyond the implicit ones that arise from spillover externalities. Put differently, it is not in the self-interest of healthy countries to accept joint-and-several liability, even though they realize that it will be in their interest to *ex post* offer *some* solidarity in order to prevent/limit spillovers of sovereign default. Could one arrange an *ex-ante* transfer from distressed countries to healthy ones to compensate the latter for accepting the future liability? The answer is "no", as financing this transfer would just add to the distressed countries' indebtedness and thus the compensation would be in funny money. The current reality is that Germany is on the hook, and any reform proposal has to address the question: what is in it for Germany?

Furthermore, even leaving aside Northern European countries' willingness to enter such a scheme, the current proposals for a reinforcement of European solidarity through joint liability do not address the possibility that even Germany

may not prepare itself adequately to its role as guarantor. Somehow, either Germany needs to be monitored and become “super solvent”, or some second defense that involves non Eurozone parties must be planned: substantial credit line from the IMF or (with all associated hazards) ECB-led devaluation.

4. - Wrapping Up

With few options left on the table, what is the way forward for the Eurozone? Let me make four concluding remarks in this regard:

- 1) *Europe needs long-term discipline rather more than that short-term austerity.* Austerity can have adverse consequences given the current low growth and high unemployment. That said, politicians’ promises that they will return to discipline in a few years completely lack credibility. Besides, as the magazine *The Economist* pointed out (on May 5, 2012), «*Calling for growth is like advocating world peace: everybody agrees that it is a good thing, but nobody agrees how to do it*».

So the challenge is to create a real commitment. To this purpose, our countries can do three things: structural reforms, institutional reforms, and reorientation of public spending toward policies that will promote growth and make it credible that debts will be repaid: investments in tradable-good industries and in growth-promoting education, R&D and infrastructure; rather than the creation of public jobs, the accumulation of new debts or support to non-tradable industries such as real estate.

- 2) *It is urgent to undertake structural reforms*, again to restore confidence in our ability to compete and repay.
 - We must overhaul Southern European labor market institutions so as to put an end to labor market duality and exclusion, and to foster job creation (for example, Olivier Blanchard and I have offered a blueprint to this purpose).
 - We must guarantee our pension systems’ sustainability by increasing the retirement age, adding flexibility and promoting fairness (substantially increasing the retirement age for skilled employees, who start their career late).
 - We must make a number of professions more competitive.

I recognize that implementing reforms in a recession is not straightforward. Allocating European funds to this purpose could help smooth the process.

- 3) *I have discussed long-term institutions that could restore the sustainability of the Eurozone.* Because we do not stand behind the veil of ignorance, the prospect

of sizeable, well-identifiable cross-subsidies hampers solidarity among countries. For this reason, and also because of moral hazard concerns, fiscal integration or full-fledged Eurobonds are not really on the agenda.

Short of these, countries must draw their lessons of the failure of the Stability and Growth Pact. Countries must give up some of their sovereignty. They must accept

- to devolve powers to independent fiscal councils,
- a financial regulation truly at the European-level,
- and when push comes to shove, further interference in the management of their economy.

More discipline in turn increases the scope for solidarity.

- 4) Finally, I have not discussed *the potential role of the Central Bank in indirectly monetizing sovereign debts*. A breakthrough of the last decades has been the successful fight against inflation; a return to inflation, even temporary, would be a setback. This comeback of inflation will however become unavoidable if the sovereign crisis spreads to large European countries; put differently, the ECB will have to stand by as lender of last resort on an even larger scale than today. Hopefully the required institutional reforms will occur sufficiently soon so as to reestablish trust in Europe and thwart the contagious spiral, and if they are insufficient to prevent such an adverse development, to ensure that inflating will be a “once in a lifetime event”.

La manifattura in Italia nell'ultimo decennio, sulle spalle dei... nani

Daniela Venanzi*

Università Roma Tre

How has the Italian manufacturing changed in the last decade? The empirical evidence shows that large size did not outperform the medium-small size. “Giants” appear less profitable, riskier and financially weaker when compared to “dwarfs”, that appear better performers, financially stronger and more able to face the crises (the gap increases in troubled periods). Furthermore, the medium-sized firms, which operate in district areas, outperformed the counterparts operating in the context of large firms. If we distinguish myth from reality, it seems clear that size effect does not win in terms of performance and risk. Policy makers could derive useful insights.

[JEL Classification: G32; L25; L60; M21].

Keywords: size effect on performance and risk; trends in Italian manufacturing; MSEs in industrial districts.

* <venanzi@uniroma3.it>, Dipartimento di Scienze Aziendali.

1. - Il pensiero dominante sulla manifattura italiana

Si parla da anni della debolezza dell'economia italiana, attribuibile secondo molti alla specializzazione settoriale, focalizzata sulla manifattura e su settori a basso contenuto tecnologico e per questo maggiormente esposti alla concorrenza di aree a minore costo del lavoro (anche se ormai i paesi emergenti sono ben presenti anche nelle produzioni a maggiore contenuto tecnologico), e alla prevalenza di imprese di medio-piccola taglia e finanziariamente fragili, perciò particolarmente sensibili agli *shock* economici. Recentissimo il pronunciamento di esponenti del governo in tal senso¹ e nella stessa direzione anche il recente Rapporto della Commissione Europea (European Commission, 2011) sulla competitività degli stati membri che, pure citando l'apporto delle imprese medie e medio-grandi alla competitività dell'economia italiana e la loro capacità di rimanere competitive anche nei settori ad alta intensità della forza lavoro, attraverso l'*upgrading* qualitativo dei prodotti offerti, definisce una priorità la crescita dimensionale delle PMI italiane (incluse quindi le "medie" imprese) e individua nell'indebitamento finanziario, eccessivo e sbilanciato sul breve termine, il vincolo alla loro crescita dimensionale.

Il pensiero dominante (cfr. ultime Relazioni annuali della Banca d'Italia; Visco, 2009; Ciocca, 2009; Rossi, 2006 e 2009; Baldwin, Barba Navaretti, Boeri, 2007) assume dunque il mito della grande dimensione, che consentirebbe incrementi di produttività e di efficienza, favorirebbe l'innovazione e renderebbe possibile l'adozione delle TIC e il pieno sfruttamento del loro "potenziale razionalizzatore". Considera le imprese maggiori meno rischiose e più solide finanziariamente, le minori troppo indebitate, soprattutto a breve termine, quindi più vulnerabili e rischiose. Individua nella finanza i vincoli alla crescita, sposando il mito del *financial gap*, secondo cui le imprese minori sarebbero tali per vincoli *supply-side* all'accesso al capitale (Berger, Udell, 1998).

In questo studio si effettua un'analisi della manifattura in Italia nel decennio 2001-2010, cercando di verificare se le esposte ragioni della sua debolezza trovino supporto nei bilanci delle imprese, ovvero se si tratti di falsi miti figli di un eccesso di astrazione dell'analisi economica, che non aggiorna modelli concettuali e interpretazioni alla varietà e variabilità delle realtà indagate e che soprattutto non

¹ «Una grande debolezza è la piccola dimensione delle nostre imprese, aggravata dalla scarsa patrimonializzazione» lo ha detto l'allora vice ministro dell'Economia, Vittorio Grilli, nel corso di un'audizione alla Camera, sottolineando che «questo aspetto di fragilità e debolezza è stato un freno alla crescita delle imprese» (Ansa, 23 aprile 2012).

distingue le imprese per le strategie produttive e di mercato perseguite; e che, sul lato finanza, trascura completamente il punto di vista *demand-side* (Watson, 2010), secondo cui la taglia media o piccola può essere frutto di scelta su ragioni di opportunità economico-strategica o comunque legate a obiettivi e progetti della proprietà/*management*.

L'analisi è condotta preliminarmente su un insieme ampio e rappresentativo della manifattura: abbiamo cioè utilizzato i dati cumulativi dell'ultima indagine dell'Ufficio Studi di Mediobanca su 2030 società italiane (Mediobanca, 20011) relativamente al periodo 2001-2010². Si tratta di un insieme che contiene sempre le stesse imprese e quindi consente un'analisi non distorta dei *trend* nel decennio considerato: 1.679 imprese manifatturiere (escludendo quindi i comparti energia e costruzioni) che rappresentano il 43% del fatturato, il 31% degli occupati e il 55% delle esportazioni (dato 2007) della manifattura (con riferimento all'universo Istat delle imprese con più di 20 addetti) e contiene tutte le imprese italiane con più di 500 addetti e un sesto delle medie imprese.

La prospettiva assunta nell'analisi è duplice. Dentro l'insieme citato, confrontiamo gli 83 maggiori gruppi a controllo italiano (imprese con fatturato superiore a 3 md di euro o che appartengono a gruppi con fatturato superiore a 3 md di euro o con azionista di controllo pubblico) con 504 imprese medio-grandi (almeno 500 dipendenti oppure fatturato superiore a 330 milioni, ma inferiore a 3 md di euro), con 430 imprese italiane a controllo estero e con 662 medie imprese – d'ora in poi MI – cioè, secondo la definizione Mediobanca-Unioncamere, società che soddisfano contemporaneamente i seguenti requisiti: dipendenti compresi nell'intervallo 50-499, fatturato tra 15 e 330 milioni di euro e non appartenenti a gruppi di grande dimensione³. La **prospettiva** quindi è **dimensionale**, ma non in senso canonico perché, diversamente da quanto avviene nelle ri-

² La base dati, per alcune parti non pubblica, relativa sia ai dati cumulativi che all'universo delle MI (cfr. oltre) è stata fornita, per gentile concessione, dall'Area Studi di Mediobanca.

³ I bilanci considerati in questa analisi sono bilanci che non consolidano le attività non svolte in Italia: nel 2010, i maggiori gruppi hanno in media oltre 2.300 addetti e circa 1 md di fatturato, le imprese medio-grandi 500 addetti e 180 milioni di fatturato, le imprese a controllo estero 525 addetti e 225 milioni di fatturato e le medie imprese 165 dipendenti e 65 milioni di fatturato. I criteri definitori degli aggregati, sopra citati, sono soddisfatti a livello di bilancio consolidato. Da notare che si tratta di bilanci aggregati, per le diverse sub-aggregazioni considerate nell'analisi, e che quindi si tratta di dati medi dei diversi indici, ponderati al peso che le imprese dell'aggregato hanno sul denominatore dell'indice, e quindi maggiormente influenzati dai valori assunti dalle imprese maggiori (per la variabile al denominatore dell'indice) dell'aggregato di riferimento.

levazioni statistiche, la dimensione non è riferita alle unità produttive fisiche, ma alle unità economico-aziendali, le uniche cui è corretto riferire strategie e scelte di gestione, nonché *performance* e rischio conseguenti⁴. Confrontiamo quindi non solo dimensioni diverse, ma anche modelli di *business* diversi, di cui la dimensione è un aspetto qualificante. Inoltre assumiamo la **prospettiva italiana**, utilizziamo cioè bilanci singoli che considerano le sole attività localizzate in Italia, con l'intento di evidenziare l'apporto alla manifattura italiana e al suo sviluppo nel decennio considerato.

Va precisato che qui l'impresa minore è la media impresa e non la piccola o micro-impresa. Parliamo cioè di imprese con almeno 50 dipendenti, che sono 10 volte più grandi della media micro-impresa del mezzogiorno d'Italia (5,8 dipendenti in media nella manifattura) e 6 volte della media micro-impresa nazionale (8,5 dipendenti in media), micro-imprese che occupano in Italia il 47,4% dei dipendenti contro una media UE del 29,8% (dati European Commission, 2011). Inoltre le MI incluse nell'aggregato indicato sono quelle di taglia maggiore, una volta e mezza la "mediana" media impresa dell'universo delle MI. Questo aspetto può evidentemente distorcere l'analisi dell'effetto dimensione su *performance* e rischio. Un primo passaggio è stato allora quello di estendere l'analisi all'universo delle MI manifatturiere⁵, al fine di aumentare la copertura della manifattura italiana e di meglio analizzare l'impresa di taglia inferiore. L'insieme analizzato (in media oltre 5.000 imprese), infatti, rappresenta il 60% circa di fatturato e valore aggiunto, il 49% circa degli addetti e il 64% degli investimenti della manifattura italiana, con riferimento agli ultimi dati Istat disponibili di imprese con più di 20 addetti, relativi al 2008 (Mediobanca, 2012). Le MI considerate sono imprese più piccole di quanto l'intervallo di loro definizione (50-499 addetti) faccia supporre. Sono polarizzate sulle classi inferiori (50-99 e 100-249 addetti), che rappresentano il 90% delle imprese (il 45% circa la prima classe) e oltre l'80% di fatturato e valore aggiunto e la media impresa della classe superiore si attesta vicino alla soglia inferiore dell'intervallo (circa 293 addetti). La mediana dell'universo considerato è di 108 addetti, che è un quinto circa della dimensione media delle medio-grandi e delle imprese a controllo estero e un ventesimo di quella dei maggiori gruppi (lo 0,5% se ragioniamo in termini di fatturato consolidato minimo), quindi un differenziale di dimensione sufficiente a indagarne le implicazioni.

⁴ Si noti come molte imprese di piccola e media taglia, in senso canonico, facciano parte di grandi gruppi.

⁵ Si tratta di oltre 3.200 società nel 2010 (4.097 il numero medio nel decennio considerato).

Restano comunque fuori dal quadro le imprese di piccole dimensioni (PI), ma l'analisi di questo insieme incontra elevati margini di errore. Si tratta di imprese i cui bilanci presentano gradi elevati di soggettività e non attendibilità, sia per fenomeni rilevanti di evasione fiscale e sommerso, ampiamente rilevati e analizzati (Coltorti, 2011), sia per il minore dettaglio dei bilanci in forma abbreviata, che queste imprese possono adottare, che rendono non distinguibile in alcune aree la gestione operativa da quella finanziaria, oltretutto più opachi gli effetti economico-finanziari delle scelte gestionali. A questi si aggiunge la maggiore difficoltà a distinguere i confini tra l'impresa e la compagine proprietaria e i rispettivi patrimoni, soprattutto quando questa opacità è fonte di vantaggi dal punto di vista fiscale. Si pensi, solo per fare alcuni esempi, ai fondi forniti all'impresa dalla proprietà a titolo di debito ovvero ai finanziamenti bancari garantiti dai depositi dei proprietari presso la banca finanziatrice o, in senso lato, dal patrimonio della famiglia, o ancora a flussi trasferiti alla famiglia sotto forma di maggiori costi. In aggiunta, la turbolenza più accentuata dell'universo delle MI si ha verso la soglia inferiore, cioè i movimenti di piccole imprese che diventano medie o di medie che tornano piccole hanno rappresentato nel decennio il 75-80% della consistenza a fine 2000. Questa "contaminazione" tra i due insiemi fa ritenere che ci sia una certa similarità di fondamentali aziendali delle piccole imprese con le medie, che sono appunto in prevalenza "ex piccole". Contaminazione che passa anche per i flussi di beni e servizi tra MI e PI, sia per l'indotto che le MI attivano, sia per la rilevante presenza delle MI negli agglomerati locali: nel 2009 circa il 40% delle MI hanno sede in distretti e in altri sistemi produttivi locali e due terzi delle MI hanno sede in province distrettuali (Mediobanca-Unioncamere, 2012).

Il confronto è stato fatto in termini relativi, ovvero confrontando per ciascun aggregato le dinamiche dei fondamentali aziendali nel decennio considerato, caratterizzato oltretutto dalla crisi recente degli anni 2008-2009, dalla "piccola crisi" del 2003 e dagli anni *post* euro e della sua rivalutazione rispetto al dollaro, che hanno comportato una perdita di competizione delle imprese italiane sui mercati internazionali, almeno per quanto riguarda la competizione da *devaluation* (Bugamelli, Schivardi, Zizza, 2010). Ovviamente questo tipo di confronto evita il *non sense* del confronto assoluto tra aggregati di imprese che fanno mestieri diversi, ma non è esente dal diverso impatto delle fasi del ciclo economico sui settori di attività, cioè dalla sua componente *industry-specific*.

2. - La dinamica della manifattura italiana nell'ultimo decennio

Il Grafico 1 riassume il *trend* dei principali aggregati dell'insieme di 1.679 imprese.

Il *trend* rispecchia l'andamento del ciclo nel decennio considerato, in particolare i due punti di svolta inferiore del 2003 e del 2009, il punto di svolta superiore posizionato tra il 2007-2008 e la "ripresina" del 2010. Il fatturato, dopo anni di crescita trascurabile, assume una dinamica crescente a partire dal 2004, raggiunge nel biennio 2007-2008 oltre il 130% del livello del 2001 e ritorna nel 2009 a livelli reali più o meno pari a quelli del 2004⁶. La dinamica del valore aggiunto ha andamento simile, ma è meno accentuata: nel decennio, infatti, l'incidenza del valore aggiunto sul fatturato scende dal 22% dei primi tre anni al 19,5% del 2008 e recupera un 1% nell'ultimo biennio⁷. Le ragioni sono principalmente due:

- a) la disintegrazione verticale della produzione, attraverso esternalizzazione di fasi del processo produttivo ovvero loro delocalizzazione (fuori dell'Italia) e il riordino delle filiere. Gli occupati sono stabilmente, anche se debolmente, in calo (nel 2010 al 90% circa del numero di addetti del 2001) e gli investimenti (in immobilizzazioni materiali) in Italia si riducono sensibilmente (solo nel 2008 il livello reale è di poco superiore al valore del 2001). L'incidenza degli investimenti sul valore aggiunto tocca il minimo nel 2010 (15%), cioè 2/3 circa della quota del 2001 (in valore assoluto sono pari a 3/4 del livello del 2001).
- b) la riduzione dei margini sulle vendite per mantenere la quota di mercato e fronteggiare il ristagno della domanda, soprattutto interna: il rapporto MOL/fatturato ha un *trend* decrescente nel decennio considerato, generalizzato per tutte le imprese. Idem l'incidenza del MOL sul valore aggiunto: sul totale, nel 2010 il 10% in meno che nel 2001, ma con crescita negli anni 2006-2007.

⁶ Si è considerato il deflatore del VA della manifattura (NACE C) dell'Eurozona (dati Eurostat), per limitare le distorsioni evidenziate da alcuni (COLTORTI F., 2011) nel calcolo del deflatore Istat, per effetto della diversa qualità dei beni, cui può essere almeno in parte riferito l'aumento dei prezzi.

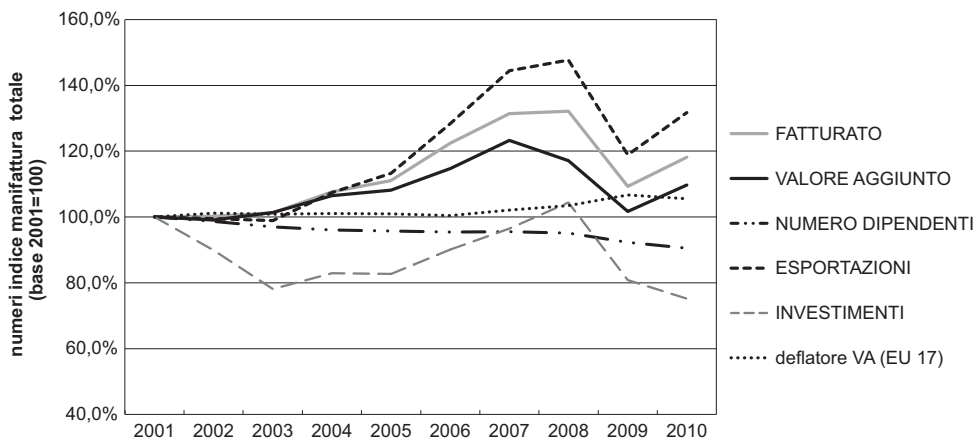
⁷ I bilanci considerati sono quelli non consolidati delle società con sede in Italia: non consolidano quindi le attività estere. Analogamente l'*export* include solo le esportazioni in senso stretto di beni prodotti in Italia e non comprende le vendite estero su estero (beni prodotti all'estero e venduti in paesi diversi da quelli in cui ha sede la casa madre) e queste posso assumere anche rilevanza per effetto della delocalizzazione produttiva. Analogamente gli investimenti includono solo quelli effettuati in Italia. Evidentemente questo fatto influenza anche i margini (la profitabilità può essere diversa da quella domestica). Idem la raccolta di capitali realizzata da finanziarie ubicate all'estero (per esempio in mercati più favorevoli).

Le esportazioni mostrano una dinamica più vivace ed evidenziano un incremento a partire dal 2005 della quota di *export* sul fatturato che non flette nei periodi di crisi: per l'insieme totale è pari in media al 40% nel decennio.

Quindi, sul totale dell'insieme considerato, appare trovare conferma un processo di decoinvolverimento della manifattura in Italia, attraverso esternalizzazione della produzione o sua delocalizzazione all'estero.

GRAF. 1

TREND PRINCIPALI AGGREGATI
(1.679 IMPRESE)



3. - Performance e rischio della manifattura: nani e giganti sono uguali?

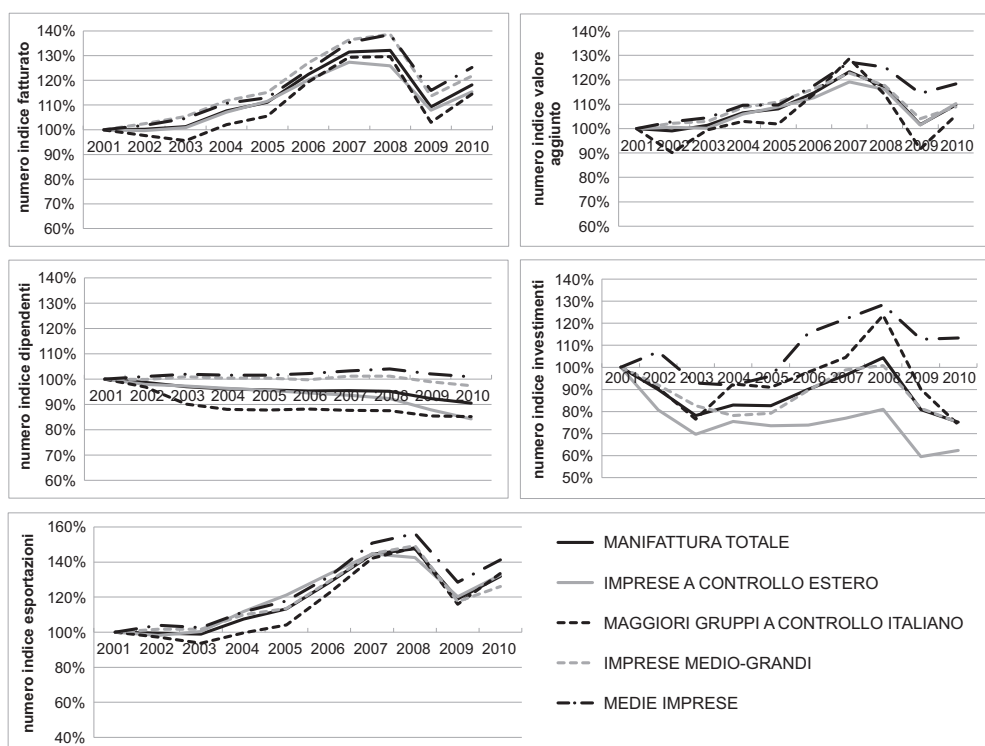
Ma le categorie di imprese considerate si comportano nel decennio allo stesso modo? Il Grafico 2 mostra andamenti molto differenziati.

Spicca la **migliore performance delle MI** e la **peggiore performance dei maggiori gruppi**, rispetto alle altre categorie. La dinamica per le MI è sempre più virtuosa: fanno meglio negli anni di svolta superiore e contengono la caduta negli anni di crisi. La forbice è più ampia per valore aggiunto, numero di dipendenti e investimenti. Nel 2010 rispetto al 2001 (base) il valore aggiunto delle MI cresce del 18%, cioè il doppio delle medio-grandi (e poco meno del doppio delle imprese a controllo estero) e più del triplo dei maggiori gruppi. Il numero di addetti, mentre decresce per tutti rispetto all'occupazione del 2001 (lieve la caduta per le medio-grandi nell'ultimo biennio, più marcata per i maggiori gruppi e le imprese a controllo estero: -15%), tiene nelle MI (mai sotto i livelli del 2001). Gli investimenti nel 2010 regi-

strano un aumento del 13% sul 2001, mentre per tutte le altre categorie diminuzioni marcate (-25% le imprese grandi e grandissime, -40% le imprese a controllo estero). Le MI nel 2001 avevano una quota di investimenti sul valore aggiunto pari al 70% di quella dei maggiori gruppi, nel 2010 è circa pari (96%): metà quella delle imprese a controllo estero e 3/4 quella delle imprese medio-grandi.

GRAF. 2

TREND PRINCIPALI AGGREGATI DELLE CATEGORIE DIMENSIONE-PROPRIETÀ



Analizziamo ora redditività e rischio. Le Tavole 1 e 2 riportano i dati per gli anni più rilevanti del decennio considerato: i due punti di svolta inferiore del ciclo (2003 e 2009) e il punto di svolta superiore (il 2007), la media del primo biennio e del triennio 2004-2006 (aggregati per la relativa similarità), e gli anni 2008 e 2010, nonché media e volatilità nell'intero periodo. Il confronto tra aggregati è evidentemente influenzato dal diverso *mix* settoriale. La dinamica tendenziale è simile tra i diversi aggregati di imprese, salvo un andamento, per alcuni indicatori,

anticipatore del ciclo economico da parte dei maggiori gruppi, evidentemente per il maggiore peso che hanno per loro le produzioni di beni intermedi.

Che cosa emerge?

In termini di **redditività**, sia netta che operativa (cfr. Tavola 1), i maggiori gruppi a controllo italiano fanno peggio di tutti, con valori medi più bassi e di segno negativo in molti anni del decennio. Le imprese differiscono sensibilmente in termini di contributo delle diverse aree della gestione alla redditività netta. La redditività netta *ante* imposte (ROE *ante* imposte) delle MI è la migliore (oltre 7 volte quella media dei maggiori gruppi), e soprattutto lo è negli anni più bui (2003 e 2009) in cui doppiano la *performance* delle imprese a controllo estero, che sono quelle con i valori più lusinghieri negli altri anni, mentre risultano penalizzate da livelli di *tax rate* più elevati, dovuti ad una maggiore incidenza dell'IRAP: perdono quindi nel confronto se consideriamo la redditività netta *post* imposte piuttosto che *ante* imposte (Mediobanca, 2011; Barbaresco, 2012)⁸. Elevata e molto variabile è per i grandi gruppi l'incidenza della gestione *extra corrente* (misurata in termini di rapporto tra risultato prima delle imposte e risultato corrente, escludendo quindi le imposte) per effetto di operazioni di rivalutazioni/ svalutazioni e di plusvalenze da realizzi. Il maggiore peso della gestione *extra corrente*, oltreché dei proventi finanziari (che pesano in media il 4,5% contro valori più bassi per le altre imprese), avvantaggia i grandi gruppi in termini di IRES, per l'esenzione di dividendi e plusvalenze. Di contro l'influenza della gestione *extra corrente* è praticamente nulla per le MI (e i proventi finanziari pesano meno dell'1%), con guadagni considerevoli dal punto di vista della trasparenza dei loro conti per investitori e analisti (trattandosi di voci di bilancio estremamente soggettive e più difficili da prevedere).

La redditività operativa, calcolata rapportando MON (margine operativo netto) e proventi finanziari al capitale investito, è migliore nelle imprese a controllo estero, seguite dalle medio-grandi e dalle MI. Il migliore *trade-off* prezzi/costi, misurato dal MOL (margine operativo al lordo degli ammortamenti, che sono frutto di politiche in parte discrezionali, e influenzati dalle politiche di rivalutazione dei cespiti ammortizzati) rapportato al valore aggiunto, eliminando così l'effetto della differente integrazione verticale delle imprese considerate (sia per il diverso *mix* settoriale, che per scelte strategiche assunte), lo ottengono le

⁸ Secondo le ultime stime Mediobanca nel 2010 il *tax rate* complessivo dei grandi gruppi è stato del 22,3%, del 28,7% per le medio-grandi, del 30,6% per le imprese a controllo estero e del 34,6% per le medie imprese (divari relativi più o meno simili anche negli anni precedenti, ad eccezione del 2009).

medie e medio-grandi imprese (42%), leggermente inferiore è quello delle imprese a controllo estero (40%) e decisamente su valori più bassi si posizionano i maggiori gruppi a controllo italiano (27,5%). La differenza si assottiglia nell'ultimo triennio per le prime tre tipologie, mentre si accentua rispetto ai maggiori gruppi. Resta la maggiore stabilità nel decennio delle MI e delle imprese a controllo estero. La redditività delle vendite (ROS) è simile per le imprese a controllo estero e per le medio-grandi, su livelli leggermente inferiori per le MI, ma più stabili nel decennio e soprattutto con maggiore tenuta nella flessione del 2009.

Le imprese differiscono anche per la struttura degli attivi e per la sua dinamica. In termini di scadenza, le attività correnti e quelle immobilizzate hanno pari peso all'incirca per medio-grandi e imprese a controllo estero, mentre per i maggiori gruppi in media l'attivo immobilizzato pesa quasi due terzi (64%), mentre nelle MI solo il 40%. Si rileva nella parte finale del decennio rispetto agli anni iniziali un peso crescente degli attivi immobilizzati e un peso decrescente degli attivi circolanti, frutto anche di processi di rivalutazione dei cespiti, eccetto che per le imprese a controllo estero, per le quali il *trend* è opposto e potrebbe essere connesso a uno spostamento verso le fasi distributive. Differente è anche la composizione per natura degli attivi e la loro dinamica: gli attivi industriali si riducono nel decennio a favore degli attivi finanziari (partecipazioni e crediti finanziari, verso consociate e terzi) del 10-12% per le medio-grandi imprese e per le imprese a controllo estero (dal 70% del 2001 al 62% circa del 2010), mentre rimangono pressoché invariati nel caso dei grandi gruppi e delle MI, che però differiscono sensibilmente per peso relativo, rispettivamente il 45% e l'80%.

TAV. 1

REDDITIVITÀ NETTA E OPERATIVA
(1.679 IMPRESE)

	ROE	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		0,3%	-2,7%	7,8%	10,0%	5,7%	0,8%	6,4%	4,4%	1,05
Imprese a controllo estero		6,9%	0,2%	8,7%	9,2%	4,9%	-0,3%	8,0%	6,2%	0,57
Maggiori gruppi a controllo italiano		-11,7%	-13,6%	4,3%	10,5%	6,5%	-2,4%	6,8%	-0,3%	38,23
Imprese medio-grandi		8,0%	4,7%	10,4%	10,2%	6,0%	3,0%	5,3%	7,7%	0,37
Medie imprese		8,5%	6,3%	8,1%	9,8%	5,0%	4,7%	5,2%	7,2%	0,25
	ROE ante imposte	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		5,0%	2,4%	13,8%	16,4%	9,4%	3,6%	10,0%	9,3%	0,61
Imprese a controllo estero		15,1%	6,2%	16,9%	18,7%	10,9%	4,8%	14,3%	13,6%	0,38
Maggiori gruppi a controllo italiano		-11,3%	-10,2%	7,2%	13,9%	7,5%	-1,4%	9,1%	1,8%	4,56
Imprese medio-grandi		13,9%	10,2%	17,3%	16,4%	9,9%	5,5%	8,2%	13,0%	0,36
Medie imprese		16,1%	13,2%	15,8%	17,7%	9,8%	8,6%	9,0%	13,8%	0,26
	INCID. GESTIONE									
	EXTRACORRENTE (no tax)	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		0,44	0,25	1,09	0,94	0,89	0,49	0,97	0,77	0,46
Imprese a controllo estero		0,90	0,48	1,00	0,89	0,69	0,39	0,79	0,81	0,28
Maggiori gruppi a controllo italiano		7,50	-8,45	1,33	0,88	1,32	-0,48	1,38	1,36	4,48
Imprese medio-grandi		0,82	0,68	1,09	1,00	0,90	0,74	0,93	0,92	0,18
Medie imprese		0,99	0,95	1,10	1,05	0,90	1,00	1,00	1,02	0,07
	MOL/FATTURATO	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		8,5%	8,4%	8,5%	8,8%	7,5%	6,7%	7,3%	8,1%	0,09
Imprese a controllo estero		9,8%	9,3%	9,4%	9,4%	8,9%	8,0%	9,1%	9,2%	0,06
Maggiori gruppi a controllo italiano		3,9%	5,2%	5,5%	6,9%	4,9%	2,9%	4,2%	4,9%	0,25
Imprese medio-grandi		11,0%	10,2%	10,1%	9,9%	8,4%	8,0%	8,1%	9,7%	0,12
Medie imprese		8,8%	8,3%	8,1%	8,4%	7,5%	7,7%	7,2%	8,1%	0,06
	MOL/VA	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		38,6%	38,1%	39,6%	42,3%	38,5%	32,6%	35,7%	38,3%	0,07
Imprese a controllo estero		41,1%	39,2%	40,6%	42,3%	40,3%	35,9%	40,1%	40,2%	0,04
Maggiori gruppi a controllo italiano		21,9%	27,4%	31,0%	38,0%	30,7%	17,6%	24,9%	27,5%	0,23
Imprese medio-grandi		44,9%	42,6%	43,5%	44,7%	40,5%	35,6%	37,0%	42,1%	0,08
Medie imprese		43,7%	41,9%	42,0%	44,6%	41,7%	38,9%	38,4%	41,9%	0,05
	ROS	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		4,0%	3,8%	4,5%	5,5%	4,3%	2,8%	3,8%	4,2%	0,17
Imprese a controllo estero		5,7%	5,0%	5,7%	6,4%	5,8%	4,7%	6,2%	5,7%	0,09
Maggiori gruppi a controllo italiano		-1,1%	0,3%	1,1%	3,4%	1,5%	-1,5%	0,3%	0,5%	2,97
Imprese medio-grandi		6,2%	5,2%	6,1%	6,4%	5,0%	3,7%	4,2%	5,5%	0,17
Medie imprese		5,1%	4,4%	4,7%	5,5%	4,7%	4,3%	4,0%	4,7%	0,09
	ROA	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		8,2%	7,8%	9,2%	12,3%	9,6%	6,4%	7,8%	8,9%	0,19
Imprese a controllo estero		11,4%	9,5%	11,2%	13,4%	11,7%	9,1%	11,7%	11,2%	0,12
Maggiori gruppi a controllo italiano		1,9%	3,3%	5,8%	12,6%	8,2%	4,4%	6,4%	6,0%	0,54
Imprese medio-grandi		11,1%	10,0%	10,7%	11,5%	9,4%	6,2%	6,7%	9,7%	0,20
Medie imprese		10,0%	9,0%	9,0%	10,7%	8,8%	6,3%	6,2%	8,7%	0,18

Per quanto riguarda il **rischio** (cfr. Tavola 2), i gruppi maggiori sono quelli che presentano la maggiore volatilità nel decennio (misurata rapportando la deviazione *standard* al valore medio) di tutti gli indicatori mentre le MI mostrano l'assetto più stabile. Il costo del capitale di rischio delle medie imprese è minore di quello delle medio-grandi, di quelle a controllo estero e dei maggiori gruppi. Ciò si può facilmente ricavare guardando alla variabilità nel decennio considerato del ROE, che risulta per le medie imprese molto inferiore a quella dei grandi gruppi (meno dell'1%) e delle imprese medio-grandi (due terzi) e a controllo estero (meno della metà). Gli azionisti delle medie imprese hanno cioè sopportato un minore rischio rispetto a quelli delle altre categorie di imprese, misurato in termini di variabilità del rendimento (redditività contabile) percepito sul loro investimento.

4. - Nani solidi e giganti fragili⁹

Per approfondire questo aspetto, abbiamo misurato il rischio recuperando l'ottica *ex ante* dell'investitore, che guarda ai fondamentali da cui il rischio dipende. Abbiamo calcolato (cfr. Tavola 2) nei 10 anni per i quattro aggregati di imprese:

•

la **leva finanziaria**, cioè l'indebitamento finanziario e lo *spread* tra redditività operativa (cioè quanto rende il capitale investito nell'azienda, misurato al lordo delle imposte, per evitare l'effetto distorsivo sopra citato) e costo del debito. Un maggiore rapporto di indebitamento comporta una maggiore variabilità del rendimento del capitale di rischio, data la variabilità del rendimento degli attivi dell'impresa. Se lo *spread* è positivo, la variabilità è nella parte alta della distribuzione (gli azionisti cioè guadagnano non solo sui capitali direttamente forniti, ma si appropriano anche del delta tra quanto rende e quanto costa il capitale investito da terzi); se negativo, la variabilità è verso il basso. Evidentemente in ottica *ex ante* del rischio percepito è importante quanto questo *spread* si sia mantenuto positivo e stabile nel tempo;

- la **leva operativa**, cioè il moltiplicatore della redditività operativa, data la volatilità dei ricavi. Dipende dalla rigidità della struttura produttiva, in termini di incidenza dei costi fissi: data cioè una certa variabilità sistematica dei ricavi dell'impresa, maggiore è l'incidenza dei costi fissi, maggiore sarà la conseguente variabilità del margine operativo, cioè la sua sensibilità al variare delle vendite;

⁹ A conferma, su dati recenti, delle considerazioni di BECATTINI G. - BELLANDI M. (2002).

- la **ciclicità dei ricavi**, cioè la variabilità sistematica dei ricavi rispetto al PIL. Abbiamo calcolato la correlazione tra valore aggiunto delle imprese e valore aggiunto della manifattura dell'eurozona a 17 paesi, entrambi depurati dell'effetto dell'inflazione (deflatore sempre eurozona per evitare le distorsioni già evidenziate in precedenza nel calcolo del deflatore per l'Italia). Abbiamo fatto riferimento alla variabilità sistematica, ipotizzando l'ottica di un investitore che possa diversificare il suo portafoglio, come insegna la teoria finanziaria.

Emerge chiaramente che la maggiore dimensione non fa premio in termini di rischio e di costo del capitale. Il rischio operativo sistematico, approssimato dal prodotto tra leva operativa e sensibilità del valore aggiunto al ciclo economico della manifattura, è più contenuto nel caso delle medie imprese (meno rigida la struttura produttiva e meno sensibile al ciclo economico il valore aggiunto), maggiore nel caso delle medio-grandi e delle imprese a controllo estero (oltre il 40% in più) e dei grandi gruppi (il 30% in più).

Per quanto attiene alla leva finanziaria, nei gruppi lo *spread* è molto variabile e negativo in tutti gli anni tranne che nel 2007 e nel 2010 (in cui è prossimo allo zero), oltretutto negativo il dato medio; più elevato e sempre positivo nel caso delle imprese a controllo estero, su livelli simili per medio-grandi e medie imprese, migliore per queste ultime nell'ultimo triennio. Il capitale netto sull'attivo netto pesa in media un terzo per tutte le imprese, ma il rapporto registra un aumento costante per le MI nel decennio, più sensibile negli ultimi tre anni, in cui aumenta di 7 punti percentuali sul 2001: nel 2010 nelle MI è pari al 38%, come le imprese medio-grandi (35% circa i maggiori gruppi e 33% le imprese a controllo estero). L'indebitamento finanziario, misurato rapportando il debito finanziario al capitale netto tangibile¹⁰, è su livelli simili nel 2010 per tutte le imprese (il debito finanziario è pari all'80% del CN tangibile, cioè all'incirca il 45% del totale dei mezzi forniti da azionisti e creditori), tranne che per le imprese a controllo estero, per le quali invece è pari circa all'unità.

¹⁰ È preferibile calcolare il rapporto di indebitamento facendo riferimento al capitale netto tangibile, cioè al capitale netto detratte le immobilizzazioni immateriali, quando si confrontano imprese che crescono con modalità diverse. Infatti, spesso le immobilizzazioni immateriali compaiono per valori rilevanti nell'attivo quando l'impresa cresce per vie esterne: nelle acquisizioni vengono evidenziati gli avviamenti, cui non corrisponde necessariamente un maggiore valore dell'attivo.

TAV. 2

RISCHIO OPERATIVO E FINANZIARIO
(1.679 IMPRESE)

	SPREAD	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		3,0%	3,0%	4,6%	6,1%	1,7%	1,6%	3,6%	3,7%	0,41
Imprese a controllo estero		8,2%	6,6%	9,1%	10,4%	7,1%	5,5%	9,5%	8,3%	0,19
Maggiori gruppi a controllo italiano		-4,6%	-2,6%	-0,9%	3,0%	-4,5%	-1,9%	0,2%	-1,4%	1,70
Imprese medio-grandi		6,2%	5,1%	5,8%	5,5%	2,4%	1,8%	2,9%	4,6%	0,38
Medie imprese		5,0%	4,4%	4,9%	5,4%	2,8%	2,7%	3,2%	4,3%	0,25
LEVA FINANZIARIA										
	(DF/CN tangibile)	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		0,99	1,04	0,86	0,88	0,89	0,87	0,85	0,90	0,08
Imprese a controllo estero		1,13	1,17	1,03	1,25	1,10	1,03	0,98	1,08	0,08
Maggiori gruppi a controllo italiano		0,81	0,89	0,67	0,63	0,78	0,80	0,82	0,75	0,14
Imprese medio-grandi		1,05	1,07	0,87	0,85	0,86	0,85	0,80	0,90	0,11
Medie imprese		1,13	1,12	1,04	1,03	0,89	0,82	0,82	0,99	0,12
CAPITALE										
	NETTO/ATTIVO NETTO	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		32,9%	32,4%	35,0%	34,3%	35,1%	35,7%	35,8%	34,4%	0,04
Imprese a controllo estero		34,4%	34,0%	35,1%	31,8%	33,0%	33,0%	32,9%	33,9%	0,04
Maggiori gruppi a controllo italiano		33,1%	31,8%	35,8%	35,8%	35,3%	35,3%	35,5%	34,7%	0,06
Imprese medio-grandi		32,0%	31,6%	34,9%	35,7%	36,2%	37,2%	37,9%	34,7%	0,07
Medie imprese		31,1%	32,0%	32,5%	32,7%	36,3%	38,5%	37,8%	33,7%	0,08
COPERTURA ONERI										
	FINANZIARI (MON/OF)	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		1,87	1,90	2,47	2,42	1,55	1,33	2,10	2,04	0,21
Imprese a controllo estero		3,15	2,93	3,81	3,51	2,70	2,64	4,41	3,39	0,17
Maggiori gruppi a controllo italiano		-0,44	0,14	0,52	1,13	0,37	-0,54	0,11	0,19	3,01
Imprese medio-grandi		2,70	2,34	2,91	2,71	1,82	1,65	2,24	2,49	0,19
Medie imprese		2,79	2,68	3,23	3,14	2,46	3,05	3,73	3,03	0,12
COSTO DEBITO										
	FINANZIARIO (OF/DF)	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		6,7%	6,3%	6,5%	8,4%	9,6%	5,9%	5,4%	6,8%	0,20
Imprese a controllo estero		7,1%	6,4%	6,2%	7,7%	8,5%	6,4%	5,5%	6,7%	0,14
Maggiori gruppi a controllo italiano		6,7%	6,3%	7,5%	11,2%	13,7%	6,9%	7,0%	8,2%	0,32
Imprese medio-grandi		6,4%	6,4%	6,5%	7,6%	8,3%	5,3%	4,6%	6,5%	0,18
Medie imprese		6,2%	5,6%	5,1%	6,4%	6,8%	4,1%	3,5%	5,3%	0,21
LEVA OPERATIVA										
		2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		18,7%	19,0%	17,6%	16,1%	16,0%	18,2%	17,3%	17,7%	0,06
Imprese a controllo estero		19,4%	19,7%	18,4%	17,0%	17,2%	18,6%	17,7%	18,4%	0,05
Maggiori gruppi a controllo italiano		18,5%	18,8%	16,9%	15,4%	14,9%	17,6%	16,7%	17,1%	0,08
Imprese medio-grandi		19,5%	19,8%	18,4%	16,7%	16,7%	19,5%	18,6%	18,5%	0,06
Medie imprese		15,8%	16,2%	15,4%	14,0%	13,9%	16,1%	15,4%	15,3%	0,05
ATTIVO IMMOBILIZZATO/										
	ATTIVO NETTO	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità
Manifattura totale		52,8%	52,5%	52,8%	53,0%	54,8%	55,2%	55,4%	53,5%	0,02
Imprese a controllo estero		49,5%	51,5%	51,3%	49,3%	49,4%	47,9%	46,8%	49,8%	0,03
Maggiori gruppi a controllo italiano		62,2%	60,6%	61,1%	62,6%	64,9%	66,3%	66,6%	62,9%	0,04
Imprese medio-grandi		49,7%	49,9%	50,9%	52,3%	54,2%	54,7%	55,8%	51,9%	0,05
Medie imprese		39,6%	38,6%	38,1%	37,8%	42,7%	44,7%	45,0%	40,2%	0,07

SEGUE TAV. 2

ATTIVO INDUSTRIALE/										
ATTIVO NETTO										
	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità	
Manifattura totale	61,1%	61,6%	61,0%	60,2%	60,4%	59,8%	57,9%	60,5%	0,02	
Imprese a controllo estero	69,6%	68,0%	66,6%	65,8%	63,9%	62,7%	61,4%	66,1%	0,04	
Maggiori gruppi a controllo italiano	42,9%	45,5%	45,7%	44,7%	44,8%	44,7%	42,6%	44,5%	0,03	
Imprese medio-grandi	67,1%	66,0%	65,2%	63,7%	64,5%	63,7%	62,0%	65,0%	0,03	
Medie imprese	80,1%	78,4%	78,1%	78,6%	81,3%	81,1%	78,8%	79,3%	0,02	
DEBITI FIN. BT/DEBITI										
FINANZIARI										
	2001-2002	2003	2004-2006	2007	2008	2009	2010	media	volatilità	
Manifattura totale	61,9%	56,1%	57,1%	54,7%	52,6%	53,5%	52,9%	56,5%	0,05	
Imprese a controllo estero	65,2%	62,0%	55,0%	54,0%	48,6%	54,6%	55,4%	57,0%	0,09	
Maggiori gruppi a controllo italiano	66,3%	51,0%	60,5%	55,1%	50,9%	51,7%	47,9%	57,1%	0,11	
Imprese medio-grandi	53,7%	52,9%	54,4%	52,2%	52,2%	51,6%	52,7%	53,2%	0,02	
Medie imprese	65,5%	62,6%	61,6%	61,9%	64,7%	60,9%	60,2%	62,6%	0,03	

Ma va evidenziato come questo livello sia stato raggiunto dalle MI con una costante riduzione del rapporto di indebitamento finanziario, che nel 2010 è a -30% del livello del 2002. Anche le medio-grandi l'hanno ridotto (-24%), le imprese a controllo estero l'hanno ridotto meno (-13%) e i grandi gruppi l'hanno mantenuto.

Nel decennio **cambia anche la struttura per scadenza del debito finanziario**. Si riduce in media per tutto l'insieme considerato il peso del debito finanziario a più breve scadenza del 20% nel 2010 rispetto al 2001, con riduzioni più marcate per i maggiori gruppi (di un terzo), solo del 3% per le medio-grandi e del 10% circa per le MI: in media i debiti finanziari a breve termine pesano di più per queste ultime (63% del totale dei debiti finanziari, ovvero 10 punti percentuali in più delle medio-grandi e 6 punti percentuali in più delle altre). Tuttavia, anche la struttura per scadenza degli impieghi delle MI è più orientata sul breve termine, come visto in precedenza. Per le MI, infatti, il 60% dell'attivo è capitale circolante netto, circa il 50% per medio-grandi e a controllo estero e solo il 36% per i maggiori gruppi, che lo hanno ridotto imponendo ai fornitori dilazioni più estese a fronte di minori dilazioni concesse alla clientela.

Nel decennio, infatti, i maggiori gruppi sono gli unici per i quali la dilazione dai fornitori è superiore (e anche in maniera rilevante) alla dilazione concessa alla clientela (rispettivamente, in media, 120 giorni a debito contro 90 giorni a credito), mentre per le altre imprese la prima è inferiore alla seconda, in misura più marcata per le MI, che quindi scontano un maggiore fabbisogno di circolante (più o meno rovesciato il rapporto tra le due dilazioni). La media sul totale per entrambe è di 105 giorni.

Per quanto precede, ne discende che il *matching* per scadenza tra impieghi e fonti è in media per le MI più equilibrato: le fonti consolidate coprono per intero gli attivi immobilizzati e un quinto di quelli correnti, mentre nei maggiori gruppi arrivano a coprire solo il 90% degli attivi immobilizzati. Nel caso delle imprese a controllo estero e delle medio-grandi la qualità del *matching* è intermedia tra MI e maggiori gruppi: in entrambe in media capitale proprio e debito finanziario a media-lunga scadenza eccedono la copertura dell'attivo immobilizzato, coprendo circa un decimo dell'attivo circolante (rispettivamente l'8% e il 13%).

È importante sottolineare che il miglioramento dell'esposizione debitoria delle medie imprese, in un periodo di congiuntura avversa, derivi, come emerge dai dati indagati in precedenza in termini di occupazione e investimenti, non da una contrazione dell'attività (le MI sono quelle che maggiormente tengono nell'ultimo triennio, in termini di occupazione e di investimenti fissi: l'incidenza di questi sul fatturato nell'ultimo triennio si riduce per tutte le imprese rispetto alla media del decennio, mentre aumenta per le MI del 15% e supera quella dei maggiori gruppi), ma dal reinvestimento in misura rilevante (in media la metà nel decennio e i quattro quinti nel 2010, percentuale massima) degli utili conseguiti, mentre le medio-grandi reinvestono meno (in media un quarto e nel 2010 la metà), e i maggiori gruppi e le imprese a controllo estero pagano dividendi anche quando chiudono l'esercizio in perdita, utilizzando riserve di utili precedentemente accumulate.

Emerge dai dati che dietro la differente dimensione, c'è una differente strategia. La proprietà della media impresa crede nell'impresa e, anche negli anni di maggiore difficoltà, continua a fare investimenti per mantenere la sua competitività, salvaguarda la sua forza lavoro, che detiene un *know-how* cruciale per il suo successo, rinuncia ai margini per mantenere quote di mercato e rinuncia ai dividendi, per autofinanziare gli investimenti o ridurre l'indebitamento e rafforzare la solidità. La **maggiore solidità finanziaria** che le medie imprese nel periodo considerato hanno perseguito ed ottenuto si evince anche dal grado di copertura degli oneri finanziari, che misura la capacità di fare fronte agli interessi passivi sul debito con il margine operativo della gestione, che è maggiore di oltre un terzo rispetto al valore del 2001, mentre è diminuito per le medio-grandi, prossimo allo zero per i maggiori gruppi, su livelli più elevati per le imprese a controllo estero, ma comunque con una dinamica nel decennio meno virtuosa. Indice di copertura che assume valori più elevati nelle medie imprese per tutto il decennio considerato, sia rispetto ai maggiori gruppi che rispetto alle medio-grandi imprese. Le MI fanno peggio solo delle imprese a controllo estero (in media 3 delle prime contro 3,4 delle seconde), ma va sottolineato il migliore valore nell'anno peg-

giore del decennio, cioè il 2009¹¹. Sono quindi imprese più solvibili¹², come dimostra anche indirettamente il minore costo del debito rispetto alle altre tipologie di imprese che i terzi finanziatori hanno loro caricato: un punto/un punto e mezzo percentuale in meno delle medio-grandi e di quelle a controllo estero, circa tre punti percentuali in meno dei maggiori gruppi in media nel decennio.

È interessante notare che il costo del debito è correlato negativamente con il grado di copertura degli oneri finanziari (come sarebbe logico attendersi, essendo quest'ultimo un indicatore molto utilizzato nel misurare il rischio di insolvenza del debito, per esempio nel calcolo dei *rating*) solo nel caso delle MI e delle imprese a controllo estero, mentre questa correlazione non è statisticamente significativa per le altre due categorie di imprese. Anzi, per i maggiori gruppi esiste una debole correlazione positiva. Pur se il costo del debito è in parte distorto da oneri di natura operativa ricompresi negli oneri finanziari, come il saldo, quando negativo, delle differenze di cambio, che potrebbero assumere in alcune imprese, per esempio i maggiori gruppi, valori più consistenti, resta possibile che il sistema bancario finanzia le maggiori imprese senza guardare alla loro solvibilità (quindi sulla base di altri criteri) e i maggiori gruppi scarichino il maggiore rischio finanziario sulle banche che poi, in periodi più favorevoli, recuperano mantenendo alto il costo del debito e quindi appropriandosi dei loro margini (operando una sorta di *smoothing* dei loro risultati netti). Questa ipotesi non tiene conto del fatto che la valutazione del merito di credito potrebbe essere fatta guardando ai bilanci consolidati dei gruppi, comprensivi anche delle attività fuori dell'Italia.

Da notare infine che nel caso delle MI sembra trovare conferma la POH (*pecking order hypothesis*) come strategia finanziaria. Si rileva infatti nel decennio considerato un legame molto stretto tra fabbisogno finanziario per investimenti non autofinanziato, variazione delle disponibilità liquide e variazione del debito finanziario: correlazione elevata positiva (+90%) tra variazione del debito finanziario non autofinanziato al netto della variazione delle disponibilità liquide e correlazione elevata negativa (-80%) tra fabbisogno non autofinanziato e variazione della liquidità. Ciò fa supporre che le imprese preferiscono usare l'autofinanziamento per finanziare il fabbisogno per investimenti e usano il debito quando il fabbisogno per investimenti fissi eccede il flusso monetario generato dalla gestione; di contro rim-

¹¹ Se misuriamo la copertura includendo oltre al MON anche i proventi finanziari, il quadro relativo non cambia, tranne che le imprese medio-grandi raggiungono le MI e le imprese dei gruppi esteri si staccano in positivo.

¹² Il modello di *scoring* R&S-Uniocamere conferma la solidità delle MI, aumentata nel 2009 (BARBARESCO G., 2012; MEDIOPANCA-UNIONCAMERE, 2012).

borsano il debito quando l'autofinanziamento eccede il fabbisogno (ovvero l'indebitamento si riduce per gli utili accantonati e non distribuiti). Usano o creano in parte liquidità (cassa e titoli a reddito fisso presenti nell'attivo) come cuscinetto (risorse di mobilità finanziaria) prima di ricorrere al debito (in caso di fabbisogno eccedente l'autofinanziamento) o prima di rimborsarlo (in caso di eccedenza netta). Questo legame non è significativo o solo debolmente significativo per le altre imprese. Questa strategia finanziaria fonda la preferenza verso i fondi generati internamente per finanziare il fabbisogno con la volontà di autonomia, autosufficienza e controllo da parte della proprietà/*management*, coerente con il modello di impresa "progetto di vita" (Becattini, 2001), che ricorre al debito in ottica conservatrice, entro limiti che garantiscano ampi margini di sicurezza ai finanziatori (elevato valore del rapporto MON/OF), ed evita il ricorso all'emissione di capitale di rischio, e la conseguente diluizione del controllo.

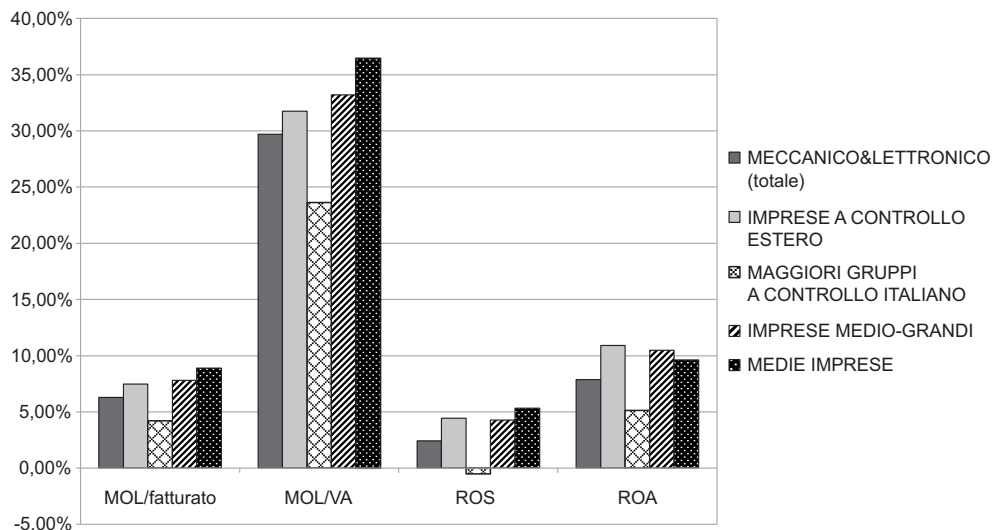
5. - Dimensione e produttività: falsi miti?

Il diverso *mix* settoriale delle imprese considerate influenza redditività e rischio, soprattutto quando riferiti alla gestione operativa e ai fondamentali da cui questa dipende (la produttività dei fattori), più strettamente legati alle caratteristiche specifiche del business. Per evitare questa distorsione nel confronto, abbiamo ristretto l'analisi *within industry*, perdendo in generalità, ma guadagnando in significatività. Abbiamo fatto riferimento al settore meccanico ed elettronico, che è il più importante per l'insieme considerato, per quota di fatturato (2/3 per i maggiori gruppi e il 30% circa per le altre imprese) e di esportazioni (oltre 3/4 per i maggiori gruppi e oltre il 40% per le altre) e per numero di imprese (tranne che per le MI per le quali pesano maggiormente i settori alimentare e beni per la persona e la casa), oltre ad essere un settore di punta della manifattura italiana. Sull'insieme totale della manifattura qui considerato, le imprese del settore (che sono il 30% del totale) producono il 43% del valore aggiunto e il 54% delle esportazioni e occupano il 47% degli addetti.

Dal Grafico 3 emerge che i maggiori gruppi a controllo italiano fanno peggio di tutte le altre categorie di imprese, mentre le imprese di taglia inferiore fanno meglio in termini di *trade-off* prezzi-costi. La redditività operativa del capitale investito delle MI è in media inferiore a quella delle imprese a controllo estero e delle medio-grandi, ma rispetto a queste ultime è superiore negli anni più recenti.

GRAF. 3

REDDITIVITÀ OPERATIVA - SETTORE MECCANICO & ELETTRONICO
(MEDIA 2001-2010)



Per confrontare la produttività dei fattori, essendo differente il grado di integrazione verticale¹³ delle imprese confrontate (minore per i maggiori gruppi a controllo italiano, che maggiormente esternalizzano e/o delocalizzano all'estero la produzione e su livelli simili per le altre imprese: diminuisce per i primi nel decennio del 12% a fronte di un aumento per le altre imprese), è più sensato ragionare in termini di valore aggiunto piuttosto che di fatturato. Inoltre, per evitare distorsioni nel confronto, dal capitale investito sono state escluse le partecipazioni e le altre attività di natura finanziaria, essendo esclusi i rispettivi proventi dal valore aggiunto.

Dalla Tavola 3 emerge chiaramente che la **dimensione non fa premio in termini di produttività**, anzi. Le imprese di taglia inferiore hanno una maggiore produttività del fattore lavoro. Infatti, se il VA per addetto è superiore in media per le imprese a controllo estero, è anche maggiore in questo caso il costo unitario del lavoro: evidentemente perché delocalizzano alcune fasi della produzione in paesi a minore costo del lavoro, svolgendo in Italia le fasi a più alto valore aggiunto, cioè quelle a monte e a valle della filiera, quali per esempio la progetta-

¹³ Abbiamo usato il rapporto (VA-MON)/FATTURATO per misurarli, ovvero costo del lavoro e ammortamenti su fatturato.

zione e la commercializzazione, che però richiedono un fattore lavoro più *skilled*. Le MI sono quelle con la minore incidenza del costo del lavoro sul valore aggiunto: questo indicatore è quello che meglio misura la produttività non per unità di forza lavoro, che potrebbe non essere omogenea come abilità e competenze possedute e quindi tipo di lavorazione svolta che le richiede (significative sono infatti le differenze del costo del lavoro per dipendente), ma per euro di costo del fattore lavoro.

La produttività del capitale investito è su livelli simili per MI e maggiori gruppi, mentre fanno meglio le imprese a controllo estero. Ma questo dato è a vantaggio delle MI perché nel decennio (in particolare nel periodo 2002-2004) i maggiori gruppi hanno effettuato svalutazioni rilevanti dei cespiti (per valori anche superiori al 10% del fatturato), così come le medio-grandi e le imprese a controllo estero (anche se in misura meno rilevante), mentre questo non è avvenuto per le MI (le svalutazioni aumentano il rapporto, perché riducono il denominatore, senza modificare il numeratore). Per tutte le imprese nel decennio aumenta del 50% circa il capitale investito per dipendente (del 40% per i maggiori gruppi e del 60% per le medio-grandi imprese). Nel caso dei maggiori gruppi e delle imprese a controllo estero, influisce la riduzione dei dipendenti; non per le medie imprese, che li hanno aumentati dell'8%, né per le medio-grandi che li hanno mantenuti invariati. Anche il credito di fornitura aumenta, pur se in parte compensato dal debito di fornitura, in misura rilevante nell'ultimo biennio per le medio-grandi imprese (difficoltà a vendere), mentre i maggiori gruppi a controllo italiano e le imprese a controllo estero l'hanno diminuito (stabile per le MI). Aumenta l'intensità del capitale fisso dei processi produttivi anche per effetto di nuovi investimenti, la cui incidenza sul fatturato è il 4,4% per i maggiori gruppi in media nel decennio, la metà circa per le imprese a controllo estero, il 3,9% per le MI e inferiore (3,2%) per le medio-grandi. Solo le MI, tuttavia, l'aumentano nell'ultimo triennio, mentre tutti gli altri la riducono.

TAV. 3

PRODUTTIVITÀ SETTORE MECCANICO & ELETTRONICO
(DATO MEDIO 2001-2010)

SETTORE MECCANICO & ELETTRONICO	(VA-MON)/ FATTURATO	VALORE AGGIUNTO x DIPENDENTE ('000 €)	FATTURATO x DIPENDENTE ('000 €)	COSTO LAVORO x DIPENDENTE ('000 €)	COSTO LAVORO/ FATTURATO	COSTO LAVORO/ VALORE AGGIUNTO	FATTURATO/ CAPITALE INVESTITO MEDIO	VALORE AGGIUNTO/ CAPITALE INVESTITO MEDIO	MONVA
TOTALE	18,7%	65,94	311,15	46,13	14,9%	70,3%	1,61	0,34	11,4%
IMPRESA A CONTROLLO ESTERO	19,1%	72,21	304,90	49,12	16,1%	68,3%	1,62	0,38	18,8%
MAGGIORI GRUPPI A CONTROLLO ITALIANO	18,2%	60,96	344,94	45,75	13,4%	76,4%	2,02	0,36	-3,2%
IMPRESA MEDIO-GRANDI	19,2%	65,31	279,89	43,84	15,7%	66,8%	1,17	0,27	18,2%
MEDIE IMPRESA	19,2%	65,03	266,53	41,42	15,6%	63,6%	1,46	0,36	21,8%

Quindi, a conclusione, il pensiero dominante che lega la produttività dei fattori e l'efficienza alla dimensione non trova supporto nei bilanci delle imprese qui analizzate. Appare evidente che è inadeguato ragionare in termini di economie interne all'impresa e di produttività fisica dei fattori. L'adozione di modelli differenti di organizzazione della produzione, che utilizza reti (anche internazionali) e filiere distrettuali, come pure la non omogeneità qualitativa dei beni prodotti fanno saltare questi legami. Non sono produzioni di massa quelle dei settori che maggiormente contribuiscono al valore aggiunto della manifattura in Italia (beni per la cura della persona e della casa, alimentare, meccanica e metallurgia), ma produzioni di nicchia; non beni di qualità omogenea, ma prodotti differenziati, che evolvono in relazione ai mutamenti della domanda o per contrastare la concorrenza, per i quali la redditività non dipende tanto dall'efficienza produttiva che abbassa i costi, ma piuttosto dalla qualità, dal servizio al cliente, dalla differenziazione e dalla personalizzazione, che consente di alzare i prezzi (Coltorti, 2008 e 2006).

La maggiore stabilità dei margini sulle vendite delle MI nel decennio considerato e la loro migliore tenuta anche nelle fasi calanti del ciclo economico, fa capire quanto sia forte l'immagine di qualità dei prodotti di queste imprese e la loro forza competitiva sui mercati, anche internazionali. Caratteri che poco o nulla hanno a che vedere con la grande dimensione, con le economie interne di scala e varietà e con i rendimenti crescenti. Nelle produzioni di nicchia i rendimenti sono decrescenti all'aumentare della quantità, l'innovazione è prevalentemente di prodotto e non di processo e l'aumento della produttività si persegue in termini di maggiore valore del bene prodotto, di snellezza della struttura produttiva che esternalizza alcune fasi della filiera (in prevalenza ad imprese locali collegate a rete) per garantire l'efficienza e ripartire tra più imprese capitale investito e rischio, la distribuzione si avvale di reti monomarca coerenti con la differenziazione dei prodotti commercializzati. Le economie possono essere cioè esterne all'impresa e interne al sistema di imprese, inteso come distretto o catena (anche internazionale) della fornitura (Coltorti, 2011). Ed è sempre la qualità del bene prodotto che spinge l'*export*: nel decennio la quota di *export* sul fatturato è aumentata per le MI nella stessa misura (di un quinto circa) dei maggiori gruppi e delle imprese a controllo estero, entrambi potenzialmente avvantaggiati per massa critica (le medio-grandi l'hanno mantenuta sostanzialmente stabile, ma su valori più elevati di tutte le altre imprese).

L'analisi condotta e i dati utilizzati riguardano le attività in Italia delle imprese considerate. Si potrebbe dire che per le grandi imprese, che operano sui mercati internazionali e delocalizzano fuori dell'Italia fasi o interi processi produttivi, il

quadro che ne discende è parziale e potrebbe apparire più soddisfacente se si guardasse ai bilanci consolidati. Tuttavia, per lo sviluppo e la competitività della manifattura italiana contano i risultati delle attività svolte in Italia e i dividendi guadagnati sulle partecipazioni nelle controllate e consociate estere, entrambi inclusi nella redditività operativa e netta qui calcolata.

Quindi, i grandi assenti sembrano i maggiori gruppi a controllo italiano, che hanno conseguito risultati insoddisfacenti e non hanno alimentato investimenti e occupazione in Italia. Verrebbe quindi più naturale chiedersi perché le grandi imprese abbiano così debolmente svolto il loro ruolo nella manifattura in Italia piuttosto che perché le imprese di taglia minore non abbiano voluto imitarle, diventando grandi. È la grande impresa che ha fallito nella manifattura italiana nel decennio considerato e stupisce che il pensiero dominante non lo evidenzi, né i rapporti della Commissione europea indichino questa come priorità da affrontare per l'Italia. Emerge anche che le MI restano tali non per vincoli finanziari esterni alla crescita, ma per scelta strategica, economicamente valida.

6. - Le medie imprese italiane: tutte prime della classe?

Ma l'universo delle MI è omogeneo al suo interno? E' un insieme vasto, in media circa 4100 imprese nel decennio 2000-2009, ultimo anno disponibile, che presenta differenze dimensionali anche marcate: le imprese della fascia minore occupano in media 76 addetti nel decennio, quasi il doppio (145 addetti) le imprese della classe intermedia e quasi il quadruplo (293 addetti) quelle della classe superiore. La dimensione media relativa (rispetto alle altre due classi) della classe inferiore per fatturato è maggiore di quanto mostri il divario dimensionale in numero di addetti: 25 milioni in media nel decennio contro 36 della classe intermedia e 70 della classe superiore. Le imprese della fascia minore sono, infatti, meno integrate verticalmente: circa il 25% in meno delle altre due classi.

Sono imprese abbastanza simili per alcuni aspetti strutturali: in media il capitale circolante pesa il 59% e il 41% il capitale immobilizzato, e limitato è il peso delle partecipazioni. La distribuzione settoriale è abbastanza simile, con maggiore peso nelle due classi superiori del settore meccanico (40% circa rispetto al 33% della classe inferiore) e minore nel settore chimico-farmaceutico (il 7-10% verso il 13% della classe inferiore). Questa similarità autorizza un confronto *cross-industry* delle tre classi dimensionali. Trattandosi di universo delle MI, le imprese variano negli anni. Non è quindi corretto analizzare il *trend*: i confronti temporali

vanno interpretati come fotografie dell'universo in anni differenti. La dinamica delle imprese dell'universo è piuttosto marcata negli anni, soprattutto per effetto di entrate e uscite dal e verso l'area delle piccole imprese, mentre più contenuta è la dinamica da e verso le imprese grandi. L'ultimo biennio ha visto una riduzione sensibile del numero delle imprese (nel 2009 il 70% del 2007), sia per revisione delle soglie che per effetto della crisi, che ha fatto tornare piccole le imprese della classe inferiore per il non rispetto della soglia di fatturato; sono gli unici due anni in cui il saldo uscite-entrate nell'area delle piccole è positivo e rilevante (420 imprese nel 2008 e 628 nel 2009). Inoltre nel decennio la taglia media della classe dimensionale inferiore in termini di numero di addetti si riduce del 10% circa, mentre aumenta per le altre due classi, in misura più marcata per la maggiore (+20%) (Mediobanca-Unioncamere 2012).

La dinamica della redditività delle tre classi dimensionali è piuttosto simile: redditività più elevata negli anni iniziali 2000-2002 e nel picco del 2007, in cui però è su livelli inferiori. Gli anni più negativi sono il 2009 e il 2003, il primo peggiore del secondo¹⁴.

Per quanto riguarda il ROE, le minori fanno peggio delle altre nel 2000, poi all'incirca ottengono risultati simili alle maggiori negli anni migliori (2002 e 2007) e soprattutto tengono maggiormente negli anni di crisi (2003 e 2009), in cui contengono il calo. Questa evidenza si conferma anche per gli altri indicatori di redditività operativa e netta, salvo il ROS in cui è la classe maggiore ad avere livelli leggermente più elevati delle altre classi dimensionali. Per tutti gli indicatori, comunque, nel 2009 la classe minore è quella che contiene meglio la crisi. La maggiore differenza tra maggiori e minori si ha per lo *spread* tra redditività operativa e costo dei mezzi di terzi, in cui le MI maggiori fanno meglio (5% verso 4%), tranne che nel 2009 (2% verso il 3,5%), ma le differenze sono contenute. Inoltre, le più piccole sono più indebitate delle intermedie e delle grandi (16% in più in media), e questo, dato lo *spread* positivo e più stabile, le avvantaggia in termini di ROE per l'effetto leva finanziaria. Per quanto attiene al *trade-off* prezzi-costi (MOL su valore aggiunto), le minori fanno meglio delle maggiori: in media nel decennio il 41,5% contro il 37-38% delle classi superiori. Inoltre, tengono maggiormente nell'ultimo biennio: invariato a fronte di un declino dell'8% delle altre due classi.

Tuttavia, come già detto, il *trend* dei rapporti MOL/fatturato e MOL/valore aggiunto è decrescente per tutte le MI dal 2000 al 2009 (come lo è stato anche

¹⁴ Per brevità si omettono le Tavole dei dati del decennio.

per gli altri aggregati considerati nella precedente analisi), ma con dinamica sensibilmente diversa: al 2009 tutte le MI raggiungono lo stesso valore dell'8% del MOL sul fatturato, ma le classi maggiori registrano valori più elevati nel 2000. Scendono quindi i margini nel decennio considerato. La percentuale di MOL sul valore aggiunto scende nel decennio solo del 4% per la classe inferiore, mentre il calo è maggiore per le classi intermedia e maggiore, -15% e -25% rispettivamente. Sono cioè le due classi superiori quelle che hanno ridotto maggiormente i margini, per mantenere le quote di mercato e contrastare il ristagno/calò della domanda interna e internazionale: sono, infatti, le MI più orientate ai mercati esteri, con una quota di *export* che è in media pari al 38% per la classe maggiore (in lieve calo nel decennio), 36% la classe intermedia (in aumento nel decennio) e 31% la classe inferiore (sostanzialmente stabile nel decennio).

Per quanto attiene alla produttività, in media nel decennio la produttività (cfr. Tavola 4) per dipendente (fatturato e valore aggiunto) è migliore per la classe inferiore: 328.000 euro di fatturato e 63.000 di valore aggiunto, a fronte di 237.000 e 59.000 (rispettivamente +38% e +7%). Ma la produttività del capitale investito è superiore per le due classi superiori. Le MI più piccole hanno un capitale investito per addetto più elevato rispetto alle altre due classi: in media nel decennio 204.000 euro, circa il 20% in più delle altre due classi. Sono simili nelle tre classi dimensionali le dilazioni da fornitori e verso clienti, come pure la rotazione del magazzino e la liquidità, mentre la rotazione del capitale fisso in termini di valore aggiunto è minore in media di 8 punti percentuali nel decennio nelle MI di minore dimensione rispetto alle classi di dimensione maggiore. Quindi le MI della classe inferiore si concentrano su un minore numero di fasi della filiera produttiva, a maggiore intensità di capitale fisso, con margini più elevati e più orientate al mercato interno.

TAV. 4

PRODUTTIVITÀ E ROTAZIONE DEL CAPITALE DELLE MI (MEDIA 2000-2009)

	(VA-MON)/ FATTURATO x DIPENDENTE ('000 €)	COSTO LAVORO x DIPENDENTE ('000 €)	COSTO LAVORO/ VALORE AGGIUNTO	FATTURATO/ CAPITALE INVESTITO	VALORE AGGIUNTO/ CAPITALE INVESTITO	DEBITI COMMERCIALI/ ACQUISTI BENI E SERVIZI	CREDITI COMMERCIALI/ FATTURATO	SCORTE/ COSTO VENDUTO	VALORE AGGIUNTO/ CAPITALE FIX	MON/VA	
UNIVERSO MI	17,4%	266,54	36,69	61,3%	1,67	37,7%	29,7%	34,2%	21,6%	0,94	22,5%
CLASSE 50-99 ADDETTI	14,4%	327,86	36,90	58,5%	1,78	34,4%	29,2%	34,1%	20,5%	0,89	25,2%
CLASSE 100-249 ADDETTI	18,6%	247,83	36,66	62,6%	1,63	38,6%	30,0%	34,6%	22,0%	0,96	21,3%
CLASSE 250-499 ADDETTI	19,5%	236,96	36,54	61,7%	1,61	40,4%	29,9%	33,3%	22,4%	0,97	21,9%

Per quanto attiene al rischio, tutti gli indicatori di redditività sono più stabili nelle imprese della classe dimensionale minore e più volatili nelle imprese della maggiore, anche se comunque entro un quadro generale di stabilità, proprio della media impresa, già emerso dall'analisi sull'insieme delle 1679 imprese. La classe dimensionale minore ha un minore grado di leva operativa (specchio della minore integrazione verticale): i costi fissi sono il 15% dei costi totali, cioè un terzo in meno delle altre classi. Inoltre il VA delle MI minori ha anche una minore correlazione con il VA della manifattura dell'eurozona (entrambi opportunamente deflazionati) rispetto alle classi maggiori: ne risulta un rischio sistematico operativo inferiore del 30% circa per le MI di minore dimensione rispetto alle altre due classi¹⁵.

Le MI minori sono più indebitate, sia in termini di indebitamento finanziario che totale, ma le differenze sono molto ridotte (dell'ordine di 2 punti percentuali). La solvibilità (MON/OF) è migliore per le classi maggiori (3,1 verso 2,65), ma nel 2009 è superiore quella delle imprese più piccole. Tutte sono comunque solvibili, con il margine operativo netto che è sempre almeno doppio degli oneri finanziari. È interessante notare che per tutte le tre classi, il costo del debito aumenta all'aumentare dell'indebitamento finanziario, mentre solo per le imprese minori c'è una correlazione negativa statisticamente significativa tra costo del debito e solvibilità dell'impresa. Come se i terzi finanziatori nel caso delle MI di taglia inferiore correlassero il costo del debito alla capacità dell'impresa di far fronte al servizio del debito, ma non fanno altrettanto per le classi dimensionali maggiori, pur se nella stessa taglia di media dimensione.

Quindi, anche analizzando l'universo delle MI, non emerge una superiorità delle MI di taglia più grande: la minore dimensione sembra fare premio in termini di maggiore flessibilità ed adattamento, che consente *performance* più stabili e meno sensibili agli andamenti congiunturali avversi.

7. - I luoghi influenzano la *performance* delle medie imprese?

Indaghiamo ora le differenze dentro l'universo delle MI connesse ai luoghi in cui esse operano. Seguendo la classificazione di Becattini e Coltorti (2004) tra aree di grande impresa (PGI) e aree distrettuali (PDI)¹⁶, assumiamo che le MI

¹⁵ Questa analisi temporale è però distorta dal fatto che le classi dimensionali non sono insiemi chiusi.

¹⁶ Definite classificando le province nell'una o nell'altra tipologia, se la percentuale di addetti attribuibile alla grande impresa (con più di 500 addetti) o ai distretti industriali supera la percentuale media nazionale (aree residue le altre).

nelle PGI operino tutte nell'indotto della grande impresa come subfornitrici (si ritiene cioè che la grande impresa, come unità complessa, sia in grado di influenzare i contesti in cui opera), mentre nelle PDI facciano parte o siano comunque connesse alla filiera distrettuale. La classificazione sottende quindi differenze che attengono al modello di *business*, alla tipologia di prodotti offerti, all'organizzazione produttiva e distributiva, all'interscambio in senso lato con il territorio che, almeno nel caso delle MI distrettuali, non riguarda solo beni e risorse (reali e finanziarie), ma è di natura socio-culturale e attiene a conoscenze tecniche/etiche/estetiche radicate nei luoghi e trasferite tra produttori e consumatori e tra produttori e loro fornitori di macchinari e servizi (Becattini, Coltorti, 2004).

L'analisi è stata condotta con riferimento ad uno specifico settore: sull'insieme o su aggregazioni settoriali (per esempio il *made in Italy*) il confronto risulterebbe distorto dal diverso peso dei diversi settori manifatturieri dentro gli aggregati delle MI nelle PDI e nelle PGI. Abbiamo scelto il settore meccanico perché settore tra quelli con migliori *performance* dentro l'universo delle MI (migliore redditività, sia netta che operativa e maggiore orientamento all'*export*) e minore rischio (minore indebitamento finanziario, maggiore solvibilità e simile rischio totale, a fronte di un maggiore rischio operativo per una maggiore incidenza dei costi fissi sui costi totali) e perché rilevante come peso: quasi il 40% per numero di imprese e valore aggiunto, un terzo per fatturato. Inoltre, sia per PDI che per PGI, rappresenta il settore più importante (40% delle PGI e il 36% delle PDI)¹⁷.

Per dimensione e struttura le due tipologie appaiono simili: addetti in media 130-134 nel decennio e fatturato tra 29 e 30 milioni di euro; simile il grado di integrazione verticale (intorno al 21%), che mostra una dinamica simile: più elevato negli anni 2001-2004, torna nel 2005 sugli stessi livelli del 2000, si riduce nel periodo 2006-2008 e torna più o meno ai livelli del 2000 nel 2009. Stesso peso dei costi fissi sui costi totali (22% il grado di leva operativa) e stessa composizione degli attivi (in media capitale fisso 35% e circolante 65%).

Tuttavia, le **MI delle aree distrettuali conseguono migliori performance**: superiore sia la redditività netta che la redditività operativa (Tavola 5¹⁸). Questo vantaggio assume maggiore rilevanza negli anni della crisi recente (il biennio 2008-2009): il ROE è circa doppio (differenza di 4-5 punti percentuali), a fronte

¹⁷ Le principali PDI della meccanica ed elettro-meccanica sono Brescia, Vicenza, Bergamo, Varese, Pordenone, Lecco, Treviso, Modena e Reggio Emilia; le principali PGI sono Torino, Parma, Cuneo, Genova, Venezia, Ferrara.

¹⁸ Per brevità, di alcuni anni relativamente omogenei si riporta il dato medio, di altri il dato puntuale, oltre a media e volatilità (deviazione *standard* su media) dell'intero periodo.

di un differenziale di 2 punti percentuali e mezzo nella media del decennio. Il divario è molto marcato anche nel 2002 (ROE triplo delle PDI), e dipende anche dal fatto che le MI delle aree di grande impresa anticipano il ciclo economico e questo potrebbe essere legato al fatto che si posizionano nelle fasi più a monte della filiera produttiva o nella fornitura di beni intermedi/servizi, come è anche confermato dal fatto che la quota di *export* è molto minore (producono in misura maggiore per il mercato interno) rispetto a quella delle MI delle aree distrettuali: in media nel decennio il 36% contro il 49%.

La Tavola 6 evidenzia, a parità di costo unitario del lavoro, una maggiore produttività sia del fattore lavoro che del capitale industriale delle MI distrettuali, anche se la differenza non è molto marcata, ma si accentua nel 2009 (anno in cui il divario nell'incidenza del costo del lavoro sul valore aggiunto è di 5 punti percentuali). Per le MI distrettuali è minore la durata della dilazione alla clientela e dai fornitori e meno marcato l'aumento nel 2009. La dilatazione dei tempi di fornitura nella crisi recente sembra dovuta a difficoltà di vendita, dato il peggioramento della rotazione delle scorte, molto marcato per le MI delle PGI (aumento di circa 20 giorni). Da notare inoltre che i due insiemi nell'ultimo biennio si riducono sensibilmente in termini di numero di imprese appartenenti: per entrambi solo il 60% delle imprese del 2007. Quelle che restano sono imprese mediamente più grandi per numero di addetti (+8%) e con maggiore intensità del capitale investito per dipendente. A parte quanto già detto sul circolante, l'aumento più marcato si registra per il capitale fisso, che nell'ultimo biennio è circa il 50% e il 33% più elevato del 2007, rispettivamente per le MI delle aree distrettuali e delle aree di grande impresa.

TAV. 5

REDDITIVITÀ E RISCHIO MI AREE DISTRETTUALI E DI GRANDE IMPRESA
(SETTORE MECCANICO)

	ROE	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		7,8%	8,2%	4,7%	6,6%	8,8%	4,8%	3,2%	6,5%	0,28
MI Meccanico		9,3%	8,7%	5,5%	8,7%	12,6%	9,3%	5,1%	8,6%	0,26
MI Meccanico PDI		9,9%	9,7%	5,9%	8,8%	13,2%	10,7%	5,9%	9,1%	0,25
MI Meccanico PGI		6,1%	3,3%	4,0%	8,9%	12,4%	5,7%	2,2%	6,7%	0,50
	ROA	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		10,9%	10,2%	8,8%	9,1%	10,6%	8,4%	6,4%	9,3%	0,15
MI Meccanico		13,0%	11,5%	10,1%	11,3%	13,8%	11,6%	8,2%	11,5%	0,15
MI Meccanico PDI		13,2%	11,7%	10,2%	11,4%	14,3%	12,5%	8,7%	11,8%	0,15
MI Meccanico PGI		11,9%	9,4%	9,4%	10,9%	12,6%	10,2%	6,5%	10,5%	0,18
	ROS	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		5,5%	5,4%	4,6%	4,9%	5,6%	4,8%	4,3%	5,1%	0,09
MI Meccanico		6,7%	6,0%	5,3%	6,0%	7,1%	6,5%	5,8%	6,2%	0,10
MI Meccanico PDI		6,9%	6,2%	5,5%	6,2%	7,4%	7,0%	6,3%	6,5%	0,10
MI Meccanico PGI		6,0%	4,8%	4,7%	5,5%	6,3%	5,4%	4,1%	5,4%	0,14
	SPREAD	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		4,9%	5,1%	3,8%	4,8%	5,2%	2,3%	2,4%	4,3%	0,26
MI Meccanico		6,7%	6,1%	4,8%	6,9%	8,8%	6,1%	4,2%	6,4%	0,22
MI Meccanico PDI		6,8%	6,1%	4,8%	7,0%	9,2%	6,9%	4,7%	6,6%	0,22
MI Meccanico PGI		5,5%	4,0%	4,3%	6,6%	8,0%	4,6%	2,2%	5,4%	0,34
	LEVA FINANZIARIA(DF/CN_{tang})	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		1,27	1,20	1,19	1,14	1,24	1,03	0,97	1,16	0,09
MI Meccanico		1,15	1,07	1,05	1,00	1,09	0,92	0,85	1,03	0,09
MI Meccanico PDI		1,15	1,06	1,05	1,00	1,04	0,85	0,79	1,01	0,12
MI Meccanico PGI		1,24	1,18	1,13	1,01	1,17	1,04	1,05	1,11	0,09
	CAPITALE NETTO/ATTIVO NETTO	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		27,7%	28,8%	29,4%	29,4%	28,3%	32,8%	35,0%	29,8%	0,03
MI Meccanico		27,0%	28,1%	28,9%	28,5%	27,3%	31,6%	34,2%	29,0%	0,03
MI Meccanico PDI		27,5%	29,1%	30,0%	29,5%	28,1%	33,3%	36,6%	30,1%	0,04
MI Meccanico PGI		24,9%	24,7%	25,7%	25,6%	25,6%	29,2%	29,8%	26,2%	0,02
	PAYOUT	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		57,8%	79,8%	65,9%	64,2%	49,7%	60,2%	70,4%	63,4%	0,15
MI Meccanico		50,6%	89,9%	60,0%	59,0%	42,4%	38,2%	52,0%	56,1%	0,26
MI Meccanico PDI		44,1%	80,7%	58,6%	57,0%	42,6%	36,1%	46,3%	52,4%	0,26
MI Meccanico PGI		76,4%	223,5%	67,5%	66,0%	32,9%	53,1%	99,7%	82,7%	0,65
	COPERTURA ONERI FINANZIARI (MON/OF)	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		2,59	2,85	2,53	3,12	2,90	2,29	2,72	2,78	0,11
MI Meccanico		2,96	3,11	2,81	3,79	3,99	3,49	3,58	3,43	0,14
MI Meccanico PDI		2,97	3,11	2,77	3,82	4,18	3,90	4,01	3,54	0,16
MI Meccanico PGI		2,59	2,49	2,63	3,74	3,73	2,86	2,39	3,05	0,21
	COSTO DEBITO FINANZIARIO (OF/DF)	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		7,3%	6,5%	6,1%	5,5%	6,7%	7,0%	4,7%	6,2%	0,14
MI Meccanico		8,2%	7,1%	6,8%	6,1%	7,0%	6,9%	5,0%	6,7%	0,15
MI Meccanico PDI		8,2%	7,3%	6,9%	6,2%	7,1%	7,0%	5,0%	6,8%	0,15
MI Meccanico PGI		8,1%	6,8%	6,6%	5,9%	6,6%	7,0%	5,2%	6,6%	0,15

SEGUE TAV. 5

	MON/OF corretto	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Meccanico PDI		3,17	3,25	3,01	4,13	4,65	4,07	4,21	3,79	0,17
MI Meccanico PGI		2,77	2,61	2,81	3,95	3,98	2,93	2,43	3,22	0,22
	OF/DF corretto	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Meccanico PDI		7,4%	6,8%	6,1%	5,5%	6,2%	6,6%	4,7%	6,2%	0,15
MI Meccanico PGI		7,2%	6,3%	5,9%	5,5%	6,0%	6,7%	5,0%	6,1%	0,13
	% EXPORT	2000-2001	2002	2003	2004-2006	2007	2008	2009	media	volatilità
MI Universo		34,0	33,7	33,3	34,3	36,0	36,2	35,1	34,5	0,03
MI Meccanico		44,3	43,7	43,9	44,7	46,3	47,5	48,5	45,2	0,04
MI Meccanico PDI		47,2	46,9	47,8	48,6	49,9	51,2	53,0	48,9	0,04
MI Meccanico PGI		35,7	36,0	34,1	34,9	36,4	38,7	37,7	35,9	0,05

Nell'ultimo biennio peggiora la produttività del capitale investito, in maniera più sensibile per le MI delle aree di grande impresa (all'80% circa del livello del 2007) e nel 2009 anche la produttività del fattore lavoro (solo leggera la flessione nel 2008): per le MI delle aree di grande impresa l'incidenza del costo del lavoro sul valore aggiunto aumenta del 10%, del 6% per le distrettuali (da considerare la vischiosità del ricorso alla cassa integrazione durante la crisi).

Guardando alla struttura finanziaria, le **MI distrettuali sono meno indebitate e più solvibili**: in media nel decennio il capitale proprio finanzia il 30% del capitale investito contro il 26% delle altre. Anche l'indebitamento finanziario è più contenuto. Il motivo principale sta nel maggiore reinvestimento degli utili delle prime rispetto alle seconde: in media nel decennio reinvestono il 48% degli utili, quasi il triplo delle altre (17%) (il tasso di reinvestimento degli utili è il complemento a 100 del *payout*). Queste differenze si accentuano in misura rilevante nei periodi di crisi. In particolare nel 2009 il capitale netto arriva al 37% delle prime contro il 30% delle seconde; l'indebitamento finanziario è in media il 10% in meno nel decennio, un terzo in meno nel 2009. La copertura degli oneri finanziari è sempre maggiore e più stabile e anche in questo caso il divario si accentua nel biennio 2008-2009.

Si potrebbe spiegare questa evidenza con una maggiore avversione al rischio delle MI delle aree distrettuali per le più gravi ripercussioni che il dissesto o l'espulsione di un anello della catena, in presenza di intense relazioni produttivo-commerciali connesse alla specializzazione per fasi delle imprese, comporta lungo la filiera del distretto, ovvero per il maggiore danno che ne deriva in termini di reputazione-immagine all'interno della comunità distrettuale. Il rischio totale che ne deriva, per effetto del minore indebitamento finanziario, è minore per le MI delle aree distrettuali, a fronte di un rischio operativo relativamente simile.

Abbiamo depurato gli oneri finanziari della differenza cambi, che in genere, quando negativa, viene in essi ricompresa, e abbiamo ricalcolato per le MI delle PDI e delle PGI sia il costo del debito che il grado di copertura degli OF (in questo caso decurtando il MON della differenza cambi, considerata costo operativo) corretti (cfr. Tavola 5), per evitare che il confronto possa risultare distorto, trattandosi di imprese con diverso peso dell'*export*. Per le MI sia delle PDI che delle PGI risulta forte la correlazione positiva tra costo del debito e indebitamento finanziario (+76%), mentre solo per le MI distrettuali il costo del debito è inversamente correlato all'indice di copertura degli oneri finanziari. Per le MI distrettuali, inoltre, il costo del debito è maggiore¹⁹: il divario, che appare trascurabile, è a ben vedere più rilevante se si tiene conto del fatto che le MI distrettuali sono meno indebitate e più solvibili. Inoltre, se incrociamo queste evidenze con quanto emerge da uno studio recente della Banca d'Italia (Finaldi Russo-Olivieri-Pastorelli, 2011), appare ancora più evidente la "virtuosità finanziaria" delle MI distrettuali e più marcata la loro penalizzazione da parte del sistema bancario. Lo studio rileva, infatti, un maggiore indebitamento finanziario e un minore costo del debito per le imprese del nord-est rispetto al nord-ovest: la differenza risulterebbe spiegata non da determinanti *firm-specific*, ma dal maggiore sviluppo del sistema bancario e dalla maggiore frammentazione dei rapporti di credito, che caratterizzano il nord-est rispetto al nord-ovest. Questa distinzione geografica si sovrappone (anche se solo in parte) a quella tra PDI e PGI²⁰ e accentua il divario tra MI delle due aree. Quindi troverebbero conferma sia la maggiore avversione al rischio delle MI delle aree distrettuali, già evidenziata, che il maggiore rischio, percepito dal sistema bancario, del credito alle MI distrettuali, o per le maggiori asimmetrie informative che deriverebbero dalla stretta interconnessione dei flussi economico-finanziari tra le imprese delle aree distrettuali, oppure per la concentrazione settoriale e geografica degli affidamenti (se da parte di banche locali). In alternativa, si potrebbe avanzare l'ipotesi che nelle aree di grande impresa, il maggiore potere contrattuale dei grandi gruppi verso il sistema bancario avvantaggi anche le MI dell'indotto.

¹⁹ Nell'ultimo biennio il differenziale si inverte (OF/DF corretto).

²⁰ L'analisi qui condotta si differenzia tuttavia perché considera un numero di imprese molto minore, di taglia media e fa riferimento al dato medio dell'aggregato; si avvantaggia però di una maggiore omogeneità del *business* delle imprese analizzate (il settore è una determinante rilevante del *leverage*).

TAV. 6

PRODUTTIVITÀ MI DI AREE DISTRETTUALI/GRANDE IMPRESA
(MECCANICA - MEDIA 2000-2009)

	(VA-MON)/ FATTURATO	FATTURATO x DIPENDENTE ('000 €)	COSTO LAVORO x DIPENDENTE ('000 €)	COSTO LAVORO/ VALORE AGGIUNTO	FATTURATO/ CAPITALE INVESTITO	VALORE AGGIUNTO/ CAPITALE INVESTITO	DEBITI COMMERCIALI/ ACQUISTI	CREDITI COMMERCIALI/ FATTURATO	SCORTE/ COSTO VENDUTO	VALORE AGGIUNTO/ CAPITALE FX	MON/VA
MI UNIVERSO	17,4%	266,54	36,69	61,3%	1,67	37,7%	29,7%	34,2%	21,6%	0,94	22,5%
MI MECCANICO	20,5%	224,23	37,95	63,5%	1,56	41,8%	33,9%	36,6%	27,3%	1,27	23,3%
MI MECCANICO PDI	20,5%	222,79	37,70	63,0%	1,60	43,0%	33,3%	36,0%	26,2%	1,30	24,0%
MI MECCANICO PGI	20,9%	219,69	37,46	65,1%	1,56	40,9%	37,0%	39,1%	26,9%	1,17	20,4%

Quindi, le **medie imprese distrettuali fronteggiano meglio la crisi** di quanto facciano le MI delle aree di grande impresa: l'appartenenza all'area distrettuale opera come un marchio di qualità che consente di reggere la competizione nelle fasi di ciclo calante e di contenere gli effetti negativi della crisi, grazie anche all'effetto sociale del "luogo" e ai rapporti virtuosi che consente di instaurare, che funzionano da ammortizzatori nelle fasi avverse. Nelle aree di grande impresa, le difficoltà appaiono maggiori e impattano maggiormente sui conti delle MI, presumibilmente in virtù dell'effetto traino delle peggiori *performance* delle imprese più grandi, cui sono legate da rapporti di fornitura; inoltre, quando queste delocalizzano (e abbiamo visto che i gruppi maggiori lo fanno largamente), le difficoltà attengono alla ricerca di nuovi clienti, il più delle volte esteri (come per esempio è avvenuto per l'indotto Fiat).

8. - Conferme e novità

Dall'analisi condotta emergono le seguenti evidenze:

- la **minore dimensione non perde** il confronto con la grande: nel decennio 2000-2010, confrontando le imprese di media taglia con i grandi gruppi (con controllo in Italia o estero) e le medie imprese minori con le maggiori, la minore dimensione non è causa di minore redditività né di maggiore rischio. Anzi, la minore dimensione favorisce l'adattamento alle alterne vicende della congiuntura economica, consentendo di contenere gli effetti negativi negli anni di crisi;
- le **imprese minori non sono più fragili finanziariamente** di quelle maggiori: anche quando più indebitate (ma il loro indebitamento finanziario si è ridotto costantemente nel decennio considerato, fino a raggiungere i livelli delle altre imprese), sono solvibili quanto le grandi (ma spesso di più) e mostrano un *matching* più equilibrato per scadenza tra impieghi e fonti;
- la **manifattura italiana** nel decennio considerato **poggia sulle imprese di taglia media**, in termini di occupazione e investimenti. I maggiori gruppi hanno de-verticalizzato e de-localizzato i processi produttivi (solo il 45% degli attivi in Italia sono attivi industriali), conseguendo bassa redditività e sopportando alti rischi; le imprese medio-grandi e a controllo estero lo hanno fatto in parte (2/3 in media gli attivi industriali in Italia), ma più intensamente nel decennio considerato, le MI hanno mantenuto attivi industriali pari all'80%, hanno conservato i dipendenti e aumentato gli investimenti;
- nelle produzioni di nicchia dei settori della manifattura italiana, la **minore di-**

- menzione consente una maggiore produttività dei fattori produttivi:** non è vincente la *leadership* di costo, ma la differenziazione e la specializzazione qualitativa dei prodotti, che consente prezzi e margini più elevati e favorisce l'*export*;
- **le MI delle aree distrettuali fanno meglio delle MI delle aree della grande impresa**, pur svolgendo lo stesso mestiere e adottando una struttura simile, e questo è vero soprattutto nei periodi di congiuntura negativa. Sono più redditizie, meno rischiose e finanziariamente più solide. Il “luogo” distrettuale, quindi, produce benefici effetti su redditività e rischio: un “marchio di qualità” che contrasta il calo della domanda e un “contesto sociale” che ammortizza l’impatto della crisi.

Quanto emerge vale, evidentemente, nei limiti degli insiemi di imprese considerati e della natura dei dati utilizzati, che sono medie di aggregati.

BIBLIOGRAFIA

- BALDWIN R. - BARBA NAVARETTI G.- BOERI T. (a cura di), *Come sta cambiando l'Italia*, Il Mulino, 2007.
- BANCA D'ITALIA, *Relazione del Governatore all'assemblea dei partecipanti*, Banca d'Italia, vari anni.
- BARBARESCO G., *Le medie imprese italiane (2000-2010)*, presentazione dell'indagine annuale sulle medie imprese industriali italiane (2000-2009), Milano, 13 aprile 2012.
- BECATTINI G., «Alle origini della campagna urbanizzata», *Bollettino del Dipartimento di Urbanistica e Pianificazione Territoriale*, Università di Firenze, 2001
- BECATTINI G. - BELLANDI M., «Forti pigmei e deboli vatuzzi. Considerazioni sull'industria italiana», *Economia italiana*, n. 3, 2002.
- BECATTINI G. - COLTORTI F., «Aree di grande impresa ed aree distrettuali nello sviluppo post-bellico dell'Italia: un'esplorazione preliminare», *Rivista Italiana degli Economisti*, n. 1 (supplemento), 2004.
- BERGER A.N. - UDELL G.F., «The Economics of Small Business Finance: The Roles of Private Equity and Debt Markets in The Financial Growth Cycle», *Journal of Banking & Finance*, no. 22, 1998.
- BUGAMELLI M. - SCHIVARDI F. - ZIZZA R., «The Euro and Firm Restructuring», in ALESINA A. - GIAVAZZI F. (eds.), *Europe and the Euro*, NBER, 2010.
- CIOCCA P., «L'impresa e le sorti dell'economia», *Aperta Contrada*, 2009, www.aperta-contrada.it
- COLTORTI F., «Il capitalismo di mezzo negli anni della crescita zero», *Economia Italiana*, n. 3, 2006.
- .-, «Il Quarto Capitalismo tra passato e futuro», prolusione al master CUOA per imprenditori di pmi, *Quaderni CUOA*, n. 4, 2008.
- .-, *L'industria italiana tra declino e trasformazione: un quadro di riferimento*, Relazione tenuta alla 52° Riunione Scientifica Annuale SIE, Roma, 15 ottobre 2011.
- EUROPEAN COMMISSION, «Member States Competitiveness Performance and Policies», 2011, <http://ec.europa.eu>
- FINALDI RUSSO P. - OLIVIERI E. - PASTORELLI S., «La struttura finanziaria delle imprese nel Nord Est», in Banca d'Italia, *L'economia del Nord Est*, 2011.
- MEDIOBANCA, *Dati cumulativi di 2.030 società italiane*, Mediobanca, 2011.
- MEDIOBANCA - UNIONCAMERE, *Le medie imprese industriali italiane (2000-2009)*, Mediobanca -Unioncamere, 2012.
- ROSSI S., *La regina e il cavallo*, Laterza, 2006.
- .-, *Controtempo*, Laterza, 2009.
- VISCO I., *Investire in conoscenza. Per la crescita economica*, Il Mulino, 2009.
- WATSON J., *SME Performance. Separating Myth from Reality*, Edward Elgar, 2010.

SAGGI

Crescita economica ed economia sommersa. Un approccio di simulazione e verifica empirica per l'Italia

Roberto Dell'Anno*
Università degli Studi di Foggia

This paper examines the role of shadow economy, tax enforcement and taxation, on economic growth. These effects are studied by simulating a modified version of an endogenous growth model proposed by Loayza (1996). Econometric results and calibration are based on the Italian economy over the period 1982-2004. Empirical evidence confirms the theoretical model. We show that the relationship between shadow economy, direct tax burden and the growth rate of GDP is negative in the long run equilibrium.

[JEL Classification: O17; E26; O43].

Keywords: shadow economy; economic growth; institutions; tax enforcement.

* <r.dellanno@unifg.it>, Dipartimento di Economia. L'Autore desidera ringraziare i *referee* anonimi e l'editor Prof. G. Piga, per gli utili commenti ricevuti. L'Autore rimane il solo responsabile di ogni eventuale errore.

1. - Introduzione

La relazione tra crescita economica ed economia sommersa (ES) è un tema di grande attualità nel dibattito pubblico. Questo lavoro si pone l'obiettivo di analizzare gli effetti che il livello di tassazione, le politiche di contrasto al sommerso e la qualità delle istituzioni potrebbero avere sul tasso di crescita del PIL ufficiale di lungo periodo quando nel sistema economico opera un settore informale.

L'approccio teorico di riferimento è quello della teoria neoclassica della crescita endogena. In particolare si propone un modello ispirato a Loayza (1996) adattato per essere idoneo ad un'applicazione simulativa. Il modello *à la* Loayza (1996) ha due caratteristiche che ne hanno giustificato la preferenza rispetto ad altre formulazioni proposte in letteratura. Il primo è di avere una struttura di base semplice (1 settore, 1 bene, 1 agente consumatore-produttore, rendimenti costanti, ecc.) e flessibile che consente di focalizzare l'attenzione su altre caratteristiche del fenomeno analizzato (ad es. il sistema sanzionatorio, la relazione con l'ES, ecc.) usualmente iper-semplificate per agevolare la trattabilità analitica. Il secondo pregio del modello proposto da Loayza (1996) consiste nella capacità di individuare, per un sistema economico in cui coesistono settore formale ed informale, condizioni di equilibrio di stato stazionario idonee ad un utilizzo per fini simulativi. In particolare l'obiettivo della simulazione numerica sarà di valutare l'entità dell'effetto che variabili fiscali, di *enforcement* e istituzionali potrebbero avere sia sul tasso di crescita del PIL (formale) procapite che sulla quota di ES in relazione al PIL (formale).

L'approccio analitico offrirà una spiegazione solo parziale del complesso legame tra crescita economica economia sommersa o informale (ES)¹ ed istituzioni. In particolare nulla si dirà circa l'endogenità delle relazioni tra istituzioni e crescita

¹ Il termine di economia sommersa è spesso utilizzato come sinonimo di economia informale, economia irregolare, evasione fiscale, ecc. L'uso di questi termini però non è equivalente e, soprattutto nei lavori di natura empirica, può causare problemi di comparabilità delle stime. In generale, si ritiene sufficientemente rigorosa la definizione adottata nel 1993 dal Sistema di Contabilità Nazionale. Essa definisce l'aggregato "Economia non direttamente osservabile" (ENO) che include il valore aggiunto creato, ma non contabilizzato, dal sistema di contabilità nazionale (OECD, 2002). Questa definizione comprende: il sommerso statistico (causato ad esempio da carenze nel sistema di rilevazione), il sommerso economico (ad es. evasione fiscale, lavoro irregolare ed informale) e le attività illegali (se basate su uno scambio consensuale, ad esempio prostituzione, contrabbando). In questo lavoro, con il termine ES si fa riferimento esclusivamente al sommerso economico la cui dimensione, per le economie avanzate, è in larga misura costituita dall'evasione fiscale. Per una spiegazione della relazione esistente tra il concetto di sommerso economico e quello di evasione fiscale si veda BOVI M. (2006).

o istituzioni ed ES. La rassegna di questa letteratura è estranea agli obiettivi di questo lavoro, ciò nonostante ci si limita ad evidenziare come sia largamente condiviso in letteratura attribuire alle istituzioni un ruolo fondamentale per la crescita economica. In realtà, aderendo ad una definizione estensiva del concetto di istituzioni, questa ricerca potrebbe essere inquadrata direttamente in questo filone della letteratura. Ciò sarebbe possibile includendo tra le determinanti della crescita, accanto alle istituzioni formali (ad es. forza della legge, sistema normativo, ecc.), anche le interazioni che avvengono al di fuori del sistema delle regole formalizzate come è appunto l'operare nell'ambito dell'ES. Nel settore informale, infatti, il sistema di tutela giudiziaria dei contratti ed il rispetto delle norme fiscali e amministrative hanno una valenza solo parziale. Secondo questa impostazione l'ES, in quanto istituzione informale, concorrerebbe con le altre istituzioni (formali ed informali) a determinare il tasso di crescita economica.

La relazione tra crescita economica ed ES è fonte di acceso dibattito tra gli studiosi non tanto sul "se" i settori formale e informale interagiscano, ma sul tipo di interazioni. In questo senso, l'inclusione dell'ES tra i fattori che influiscono sulla crescita è supportata da una letteratura, consolidata (ad es. de Soto, 1989; Johnson, Kaufmann, Mcmillan e Woodruff, 2000; Friedman, Johnson, Kaufmann e Zoido-Lobaton, 2000; Carillo e Pugno, 2004; Banerjee e Duflo, 2005; Cimoli, Primi e Pugno, 2006). Più controverso è invece affermare se l'ES abbia una funzione pro- o anti- ciclica, se gli effetti siano indipendenti dal grado di sviluppo economico, se relazioni di (Granger) causalità esistano da o verso l'ES, ecc.²

In sintesi, il contributo di questo lavoro si può sintetizzare nel tentativo di formalizzare, simulare e poi testare empiricamente per l'Italia un modello di crescita endogena dove sia le istituzioni formali (ad es. sistema sanzionatorio penale e amministrativo, qualità delle istituzioni, pressione fiscale, efficienza settore pubblico, ecc.) che informali (ES) giocano un ruolo nel determinare la crescita di lungo periodo.

Il lavoro è organizzato come segue: il secondo paragrafo presenta il modello teorico che sarà calibrato per l'economia italiana (terzo paragrafo); il quarto paragrafo presenta un'analisi di statica comparata attraverso una simulazione numerica; il quinto paragrafo testa empiricamente se le ipotesi del modello sono confermate con riferimento all'economia italiana nel periodo 1982-2004. Il sesto paragrafo riassume i principali risultati dell'analisi ed offre alcune conclusioni di

² Con solo riferimento ai lavori recenti che trattano questi temi si segnalano: SCHNEIDER F. e ENSTE D.H. (2000); GILES D.E.A. e TEDDS L.M. (2002); SCHNEIDER F. (2005); DELL'ANNO R. (2008); BOVI M. e DELL'ANNO R. (2010).

ordine generale. Due appendici forniscono maggiori dettagli per quanto concerne la costruzione del database e l'analisi statistica di non-stazionarietà e cointegrazione propedeutici all'analisi empirica.

2. - Un modello di crescita endogena in presenza di economia sommersa

In questo paragrafo si presenterà il modello di crescita utilizzato come riferimento per l'analisi simulativa ed empirica degli effetti di alcune variabili di politica economica sul tasso di crescita del PIL. Tale obiettivo è perseguito adattando il modello di crescita endogena proposto da Loayza (1996). La rappresentazione dell'economia proposta consente di esaminare come l'ES reagisce a variazioni, *ceteris paribus*, della pressione fiscale, dell'*enforcement*, della qualità delle istituzioni, della produttività del settore pubblico e, attraverso queste, delle implicazioni che esse hanno sul tasso di crescita del PIL procapite di lungo periodo.

Il modello è costruito ipotizzando una tecnologia di produzione che è funzione dell'offerta di beni pubblici congestionabili. L'economia è composta da agenti che offrono sul mercato una misura di capitale che include sia capitale fisico che capitale umano. La tecnologia di produzione esibisce rendimenti di scala costanti e produce un unico bene utilizzato, indifferentemente, per consumo e per investimento. Seguendo l'impostazione di Barro e Sala-i-Martin (1992), si assume che il tasso di rendimento del capitale dipenda dal rapporto tra l'ammontare di beni pubblici (G) e la produzione totale (Y). In base a queste ipotesi, il prodotto dell'agente i -esimo Y_i è dato da:

$$(1) \quad Y_i = A \left(\frac{G}{Y} \right)^\alpha k_i \quad \text{con } 0 < \alpha < 1 \text{ e } A > 0$$

dove: A è il parametro che misura la produttività (esogena); k_i la dotazione di capitale dell'agente i -esimo; α è l'elasticità della produzione a variazioni del rapporto (G/Y).

Esistono due settori nell'economia:

- il settore formale (y_i^F), dove gli agenti pagano un'imposta proporzionale sul reddito con aliquota τ ed il cui gettito viene utilizzato dal governo per finanziare la fornitura di servizi pubblici produttivi (G), l'attività di *enforcement* e le altre spese improduttive.

$$(2) \quad y_i^F = (1 - \tau) A \left(\frac{G}{Y} \right)^\alpha k_i, \text{ con } 0 < \tau < 1$$

- il settore informale (y_i^I), dove gli agenti rinunciano ad una quota π^e del loro reddito a titolo di sanzione ed il cui ammontare viene impiegato dal governo per finanziare esclusivamente l'attività di controllo fiscale. Si ipotizza che l'operatore dell'ES sconti la sanzione per la probabilità attesa di accertamento per cui l'aliquota sanzionatoria incorpora la percezione soggettiva del contribuente di essere sottoposto a controllo.

Gli operatori del mercato informale avranno accesso soltanto ad una quota (δ) dei servizi pubblici offerti dal settore pubblico (G) il cui utilizzo è consentito senza pregiudicare lo *status* di irregolarità dell'attività economica (beni non escludibili)³. In simboli:

$$(3) \quad y_i^I = (1 - \pi^e) A \left(\frac{\delta G}{Y} \right)^\alpha k_i, \text{ con } 0 < \pi < 1 \text{ e } 0 < \delta < 1$$

La fornitura dei servizi pubblici è finanziata attraverso l'imposizione sulla produzione ufficiale:

$$(4) \quad G = \eta(q, \lambda, \rho) (\tau, Y^F), \text{ con } 0 < \eta \leq 1$$

dove $\eta(\cdot)$ è la quota di gettito destinata al finanziamento dei servizi pubblici produttivi rivali e $1 - \eta(\cdot)$ è la parte destinata al finanziamento delle spese improduttive tra cui rientrano anche i controlli fiscali. Si ipotizza che $\eta(\cdot)$ sia positivamente correlata con la qualità delle istituzioni (q), ciò è motivato dal fatto che migliori istituzioni permettono di amministrare l'apparato burocratico in modo più efficace ed efficiente [$(\partial \eta / \partial q) > 0$].

Particolare attenzione è stata rivolta alla formulazione delle variabili che determinano l'attività di contrasto all'ES (π^e). Rispetto a Loayza (1996), dove un incremento dell'*enforcement* era possibile solo riducendo le risorse pubbliche destinate ad acquistare beni e servizi pubblici "produttivi", nel modello qui proposto, si ipotizza che l'incremento dell'attività di controllo (π^e) produca un aumento del gettito

³ Ad esempio la spesa pubblica per finanziare i servizi per la tutela giudiziale dei contratti, il sistema di sicurezza sociale e previdenziali per i lavoratori (formali), i sussidi pubblici alle imprese, ecc.

maggiore del costo del servizio (improduttivo). Da ciò deriva che il maggior gettito riscosso può essere utilizzato anche per spese pubbliche propedeutiche alla crescita. In particolare, se i controlli fiscali (p) o i procedimenti penali (λ) diventano più frequenti e/o le sanzioni amministrative (σ) o penali (f) più dure, si riduce il settore informale $[(\partial y^I/\partial \pi^e) < 0]$. Si ipotizza, inoltre, che l'ammontare di risorse necessario a migliorare il sistema di *enforcement* sia decrescente al migliorare della qualità delle istituzioni e che il rendimento, in termini di quota di gettito destinato alla spesa pubblica produttiva, della procedibilità in sede penale dei reati tributari (λ) abbia rendimenti decrescenti si scala⁴. In termini formali la quota di gettito destinata al finanziamento della spesa pubblica produttiva (G) è data da:

$$(5) \quad \eta = p\lambda(q - \lambda)$$

dove, imponendo il vincolo sulla qualità delle istituzioni $q > 2\lambda$, si ottengono le seguenti derivate parziali:

$$\frac{\partial \eta}{\partial \lambda} > 0; \frac{\partial \eta}{\partial p} > 0; \frac{\partial \eta}{\partial q} > 0; \frac{\partial^2 \eta}{\partial \lambda \partial q} > 0; \frac{\partial^2 \eta}{\partial p \partial q} > 0; \frac{\partial^2 \eta}{\partial \lambda^2} < 0$$

Indicando con S la dimensione relativa dell'economia sommersa rispetto a quella totale ($Y = Y^F + Y^I$):

$$(6) \quad S \equiv \frac{Y^I}{Y}$$

Dalla (4) e (6) possiamo ricavare il rapporto tra la fornitura di servizi pubblici e la produzione totale:

$$(7) \quad \frac{G}{Y} = \eta(q, \lambda, p)(\tau(1 - S))$$

Per la contemporanea presenza di fenomeni di congestione nel servizio pubblico e a causa della non contribuzione dell'ES al finanziamento dei beni pubblici produttivi, avremo che, *ceteris paribus*, all'aumentare della dimensione dell'ES si riduce la produttività di tutti gli agenti.

⁴ Questa ipotesi è introdotta perché il sistema giudiziario è un esempio tipico di servizio pubblico soggetto a congestione.

Un ruolo fondamentale nel modello proposto in questo lavoro è svolto dall'aliquota sanzionatoria attesa (π^e). Essa determina la partecipazione o meno dell'agente all'ES e la sua formalizzazione costituisce il principale elemento di differenziazione rispetto a Loayza (1996). Nel modellare il sistema sanzionatorio⁵ si assume che il produttore, nel decidere se operare nell'ES, valuti soggettivamente⁶ tanto le conseguenze pecuniarie dirette – sanzione amministrativa (σ) – quanto quelle di natura penale – probabilità di denuncia all'autorità giudiziaria (λ) e costi ad essa associati (f). La forma funzionale dell'aliquota sanzionatoria attesa (π^e), è esplicitata come relazione lineare sia del livello di diffusione del sommerso che della probabilità di accertamento. Questi fattori incidono proporzionalmente sulle due componenti⁷ che determinano l'ammontare della sanzione tributaria (attesa). La prima, collegata al costo monetario diretto (sanzione amministrativa pecuniaria proporzionale all'imposta evasa: $\sigma\tau$) e, la seconda, collegata all'eventuale denuncia all'autorità giudiziaria conseguente all'atto evasivo accertato. In termini economici, essa può interpretarsi come la quota di produzione necessaria a finanziare l'attività difensiva del contribuente in sede giudiziaria (o anche come prezzo ombra del costo psicologico del dover sostenere un procedimento giudiziario). Ipotizzando, per semplicità, un costo fisso per l'attività difensiva o costo psicologico (f), questo sarà moltiplicato per la probabilità che all'accertamento tributario segua una denuncia dell'autorità giudiziaria (λ). In termini formali:

$$(8) \quad \pi^e = (\sigma\tau + \lambda f) pS$$

⁵ Questa formulazione cerca di sintetizzare le caratteristiche peculiari del sistema sanzionatorio tributario italiano.

⁶ L'aliquota sanzionatoria attesa è infatti una variabile non direttamente osservabile, per cui, la partecipazione o meno al settore informale è motivata dalla differente percezione che ciascun agente realizza circa l'entità delle variabili σ , τ , p , S , λ , f . In questo senso, l'esistenza di due settori diventa la conseguenza delle differenti percezioni delle variabili fiscali e di *enforcement* quali ad esempio il grado di illusione fiscale a cui sono sottoposti i contribuenti, la valutazione soggettiva del costo psicologico o reputazionale del partecipare all'ES; la percezione della probabilità di essere accertati, ecc.

⁷ Si è ipotizzata una separabilità tra la sanzione amministrativa e quella penale. Questa ipotesi meglio si adatta alle caratteristiche del sistema sanzionatorio italiano. Nel vigente sistema vi è, infatti, autonomia tra procedimento penale e processo tributario (principio di specialità). Ad esempio, il processo tributario non può essere sospeso per la pendenza del procedimento penale avente ad oggetto i medesimi fatti nè il giudicato penale per reati tributari rileva ai fini del contenzioso tributario. Per le sanzioni amministrative, l'amministrazione finanziaria può quindi irrogare subito le sanzioni a prescindere della rilevanza penale del fatto e viceversa.

In sintesi, l'aliquota sanzionatoria attesa π^e sarà funzione:

(1) della probabilità di accertamento $\left[\left(\frac{\partial \pi^e}{\partial p} \right) > 0 \right]$;

(2) dell'aliquota sanzionatoria amministrativa sull'imposta evasa $\left[\left(\frac{\partial \pi^e}{\partial \sigma} \right) > 0 \right]$;

(3) dell'efficacia del sistema sanzionatorio penale, misurato come la probabilità che dall'accertamento tributario derivi una denuncia all'autorità giudiziaria $\left[\left(\frac{\partial \pi^e}{\partial \lambda} \right) > 0 \right]$;

(4) dal costo dell'attività difensiva in sede giurisdizionale (o costo psicologico dell'essere sottoposto a procedimento giurisdizionale) $\left[\left(\frac{\partial \pi^e}{\partial f} \right) > 0 \right]$;

(5) della dimensione del settore informale⁸ $\left[\left(\frac{\partial \pi^e}{\partial S} \right) > 0 \right]$;

(6) dell'aliquota d'imposta $\left[\left(\frac{\partial \pi^e}{\partial \tau} \right) > 0 \right]$.

Valendo l'ipotesi neoclassica di perfetta mobilità dei fattori produttivi tra il settore formale e quello informale, avremo che in equilibrio il tasso di rendimento tra i due settori deve essere necessariamente uguale (eq. (9)). Da questa condizione si ottiene la quota di ES rispetto al Pil ufficiale d'equilibrio. Restringendo la soluzione del modello ad una soluzione interna, dove cioè coesistono ES ed economia ufficiale, sostituendo e semplificando la condizione di uguaglianza dei saggi di rendimento settoriali (2) e (3), otteniamo:

$$(9) \quad (1 - \pi^e(S, \sigma, \tau, p, \lambda, f)) \delta^\alpha = (1 - \tau)$$

Sostituendo la (8) nella (9) ed esprimendo in funzione della dimensione relativa dell'ES abbiamo:

⁸ Come sottolinea LOAYZA N.V. (1996), collegare la sanzione alla dimensione relativa dell'economia informale è un modo per rendere endogena le scelte di politica di contrasto all'economia informale rispetto al contesto in cui si opera. In altri termini si sta ipotizzando che per i Paesi dove il fenomeno dell'ES assuma dimensioni maggiori vi sia un maggiore interesse del *policy maker* a contrastare il fenomeno che si sostanzia in politiche sanzionatorie più dure.

$$(10) \quad S = \frac{\delta^\alpha + \tau - 1}{\delta^\alpha p(\sigma\tau + \lambda f)}$$

La cui soluzione interna è garantita imponendo le seguenti restrizioni sui parametri:

$$(11) \quad S > 0 \Rightarrow \tau > 1 - \delta^\alpha$$

$$(12) \quad S < 1 \Rightarrow \delta^\alpha [1 - p(\sigma\tau + \lambda f)] + \tau < 1$$

Dalla (12) possiamo ricavare il vincolo sul parametro $\sigma(\tau, p, \lambda, \delta, \alpha, f)$:

$$(13) \quad \sigma > \frac{\delta^\alpha (1 - p\lambda f) + \tau - 1}{\tau \delta^\alpha p}$$

Dalla (13), per garantire che la sanzione amministrativa sia maggiore dell'imposta evasa, otteniamo il vincolo su $\tau(p, \lambda, \delta, \alpha, f)$:

$$(14) \quad \tau > \frac{1 + \delta^\alpha (p\lambda f - 1)}{1 - \delta^\alpha p}$$

Dalla (14), per garantire un'aliquota d'imposta positiva, possiamo determinare il limite inferiore della probabilità di accertamento $p(\lambda, \delta, \alpha, f)$:

$$(15) \quad p > \frac{\delta^\alpha - 1}{\delta^\alpha \lambda f}$$

Considerando che $p < 1$, dalla (15) deriviamo che il limite inferiore al parametro di *enforcement* giudiziario $\lambda(\delta, \alpha, f)$:

$$(16) \quad \lambda > \frac{\delta^\alpha - 1}{f \delta^\alpha}$$

Vincolo che risulta sempre verificato per valori di $\lambda > 0$, in quanto il membro di destra della disequazione è sempre negativo essendo sia la produttività del settore pubblico rispetto al settore privato (α) che la quota di servizi pubblici non

escludibili agli operatori dell'ES (δ) compresi nell'intervallo unitario e il costo della difesa nel procedimento giudiziario sempre positivo ($f > 0$).

Dalla (10), e rispettando i vincoli sui parametri che garantiscono soluzione interna all'ES (11-15), possiamo rilevare in che modo variazioni nelle variabili influiscono sulla dimensione dell'ES:

1. Quando la probabilità di accertamento aumenta allora l'ES si riduce $[(\partial S/\partial p) < 0]$
2. Quando l'aliquota della sanzione amministrativa aumenta allora l'ES si riduce $[(\partial S/\partial \sigma) < 0]$;
3. Quando l'aliquota fiscale aumenta allora l'ES aumenta $[(\partial S/\partial \tau) > 0]$;
4. Quando la rilevanza penale dei reati tributari aumenta allora l'ES si riduce $[(\partial S/\partial \lambda) < 0]$;
5. Quando il costo del procedimento penale (tutela legale, spese processuali, costo psicologico) aumenta allora l'ES si riduce $[(\partial S/\partial f) < 0]$;
6. Quando la produttività del settore pubblico, relativamente a quello privato, aumenta allora l'ES si riduce $[(\partial S/\partial \alpha) < 0]$;
7. Quando la quota di servizi pubblici non escludibili agli operatori del sommerso (possibilità di *free riding*) aumenta allora l'ES aumenta $[(\partial S/\partial \delta) > 0]$.

Sostituendo la (7) nella (3) abbiamo che in condizioni di equilibrio tra ES ed economia ufficiale il rendimento netto del capitale (r) è dato da:

$$(17) \quad r = [A(1 - \tau)\tau^\alpha] \left\{ \eta(q, \lambda, p) [1 - S(\sigma, \tau, p, \lambda, \delta, \alpha, f)] \right\}^\alpha$$

Il primo addendo $[A(1 - \tau)\tau^\alpha]$ corrisponde al caso in cui non vi è ES. In questa ipotesi il tasso di rendimento del capitale è una funzione prima crescente e poi decrescente rispetto all'aliquota fiscale τ .

Per quanto concerne la scelta ottima del consumatore, Loayza (1996) utilizza una funzione di utilità che ha un'elasticità di sostituzione intertemporale (ESI) costante.

$$(18) \quad \text{Max } U = \int_0^\infty \frac{c_i^{1-\theta} - 1}{1-\theta} e^{-\rho t} dt$$

Dove ρ è il tasso (costante) di preferenza intertemporale del consumo e la cui ESI è data dal reciproco dell'avversione relativa al rischio (θ).

Il vincolo di bilancio dinamico indica che la variazione della ricchezza del consumatore rappresentativo è pari alla differenza tra reddito da capitale al netto dell'imposta ed il consumo $[c_i(t)]$.

$$(19) \quad \dot{k}_i(t) = y_i(t) - c_i(t) = rk_i(t) - c_i(t)$$

L'Hamiltoniano del problema è:

$$(20) \quad H = \frac{c_i^{1-\theta} - 1}{1-\theta} e^{-\rho t} + \mu [rk_i - c_i]$$

Le condizioni per la massimizzazione sono:

$$H_c = 0 \rightarrow e^{-\rho t} c_i^{-\theta} = \mu \quad H_k = -\dot{\mu} \rightarrow r = -\frac{\dot{\mu}}{\mu}$$

e la condizione di trasversalità, che impone al consumatore di non indebitarsi senza limiti, è $\mu_T k_T = 0$. Dalla soluzione del problema di ottimo, abbiamo il tasso di crescita del consumo che è costante ed è pari a γ :

$$(21) \quad \frac{\dot{c}_i(t)}{c_i(t)} = \gamma = \frac{1}{\theta} [r - \rho]$$

In stato stazionario la soluzione del modello implica che i tassi di crescita del capitale, della produzione aggregata (Y) sia formale (Y^F) che informale (Y^I) siano costanti ed uguali al tasso di crescita del consumo (γ). Possiamo scrivere quindi che:

$$(22) \quad \gamma = \frac{1}{\theta} \left\{ \left[A(1-\tau)\tau^\alpha \right] \left\{ \eta(q, \lambda, p) \left[1 - S(\sigma, \tau, p, \lambda, \delta, \alpha, f) \right] \right\}^\alpha - \rho \right\}$$

Che in forma esplicita diventa:

$$(23) \quad \gamma = \frac{1}{\theta} \left\{ \left[A(1-\tau)\tau^\alpha \right] \left\{ \lambda p (q - \lambda) \left[1 - \frac{\delta^\alpha + \tau - 1}{\delta^\alpha p (\sigma\tau + \lambda f)} \right] \right\}^\alpha - \rho \right\}$$

3. - Calibrazione per l'economia Italiana

In questo paragrafo si calibrerà il modello di crescita endogena nella sua condizione di equilibrio di lungo periodo come espresso dall'equazione (23) dove, accanto ai parametri di *policy* e di preferenze del consumatore, anche la qualità delle istituzioni è trattata come esogena. Questa ipotesi, se per un verso semplifica l'utilizzo del modello per esercizi di statica comparata, dall'altro ne costituisce un limite alla capacità di interpretare compiutamente la relazione endogena tra istituzioni e crescita di stato stazionario. In termini economici tale assunto sottintende l'irrelevanza di eventuali retroazioni tra variazioni del livello di sviluppo economico o di dimensione dell'ES e qualità del contesto istituzionale.

Per analizzare con una simulazione per equazioni, *ceteris paribus*, gli effetti di variazioni nei parametri di *policy* sia sul tasso di crescita del PIL che sulla quota di ES di lungo periodo è necessario calibrare 11 parametri. Tre parametri connessi a caratteristiche tecnologiche e produttive (produttività settore pubblico rispetto al settore privato, quota di servizi pubblici produttivi non escludibili agli operatori dell'ES e produttività esogena del capitale). Sei coefficienti riferibili al contesto istituzionale e di politica fiscale (qualità istituzioni, probabilità di accertamento, aliquota della sanzione amministrativa, probabilità che dal controllo fiscale derivi un procedimento penale, costo difesa legale/costo psicologico/spese processuali per una denuncia in sede penale, aliquota media d'imposizione fiscale) e due connessi alle preferenze del consumatore (tasso di sconto della preferenza intertemporale al consumo e avversione relativa al rischio).

Busato e Chiarini (2003) sintetizzano tre possibili metodologie per la calibrazione dei modelli macroeconomici. Una prima fonte di informazioni per la calibrazione può derivare dalla struttura del modello nelle funzioni in forma ristretta. Una seconda strategia che fissa i parametri al fine di riprodurre, attraverso il modello, alcune grandezze macroeconomiche misurabili di lungo periodo. Infine, per i parametri delle preferenze e la propensione al rischio degli agenti, si può utilizzare come fonte di informazioni le stime econometriche ottenute in letteratura. Come è prassi nelle calibrazioni, anche in questo lavoro si seguiranno, in funzione della disponibilità di dati, le precedenti metodologie di calibrazione. In relazione ai tre parametri connessi alle caratteristiche tecnologiche e produttive:

1. Per la produttività del settore pubblico rispetto al settore privato non esistono specifiche stime in letteratura, si è quindi utilizzato un valore di calibrazione pari al valore centrale dell'intervallo di oscillazione consentito ($\alpha = 0,5$).

2. Con riferimento alla quota di spesa pubblica non escludibile sul totale della spesa pubblica (δ), si propone una stima per l'economia italiana. Dalle stime dell'ISTAT⁹ sulla spesa a prezzi correnti delle Amministrazioni pubbliche (AP) suddivisa per funzione calcoliamo, per tutte le AP, la percentuale di spesa pubblica, a cui possono realisticamente accedere esclusivamente i produttori o (ex produttori) formali, sul totale delle uscite delle AP. In particolare si ipotizza che il rapporto tra le spese delle AP per consumi intermedi, per acquisto di beni e servizi prodotti da produttori *market*, per contributi alla produzione e per le prestazioni sociali in denaro sul totale delle uscite possa essere considerata come la percentuale di spesa a cui possono beneficiare solo i produttori formali (quindi escludibile). Di conseguenza il suo complemento sarà una *proxy* della "spesa pubblica non escludibile" richiamata dal modello. Stimata tale percentuale per ciascuno degli anni in cui sono disponibili i dati (1990-2006), la media aritmetica è il valore che assegniamo al parametro ($\delta = 0,5$).

3. Per il parametro di produttività esogena del capitale (A), anche in questo caso non vi sono indicazioni precise in letteratura, il valore più neutrale rispetto a valori alternativi è 1.

I cinque coefficienti della politica fiscale e il parametro della qualità delle istituzioni sono stati calibrati in base alle seguenti ipotesi:

4. Per la probabilità di subire un accertamento (p), si parte dall'assunto che essa debba essere calcolata non in relazione al totale dei contribuenti ma con riferimento a quegli operatori che possono effettivamente decidere se partecipare o meno al settore informale. Per fare ciò la probabilità è stata stimata come rapporto tra numero dei controlli da parte delle due principali AP titolari dell'accertamento fiscale ed il numero dei (potenziali) evasori¹⁰. Nello specifico

⁹ Dati ISTAT, Spesa delle Amministrazioni pubbliche per funzione (scaricabile da: http://www.istat.it/salastampa/comunicati/non_calendario/20100121_00/base_dati.zip).

¹⁰ Ad esclusione dell'accertamento formale (artt. 36-*bis* e 36-*ter* del D.p.r. n. 600 del 1973) che è un mero riscontro cartolare delle dichiarazioni ed è finalizzato alla verifica della correttezza degli adempimenti in sede di dichiarazione a carico dei contribuenti ed è effettuato sul 100% dei contribuenti che presentano una dichiarazione, l'accertamento sia analitico che induttivo è rivolto quasi esclusivamente ai lavoratori autonomi ed imprese. In questo senso dividere il numero dei controlli sul totale dei contribuenti (incluso quindi anche i lavoratori dipendenti) non è coerente con la percezione degli agenti che sono consapevoli che la probabilità di subire un accertamento dipende dalla tipologia di contribuente a cui si appartiene ed anche dallo specifico settore produttivo in cui si opera. Con riferimento alla *proxy* qui proposta, si sta assumendo che la probabilità attesa del contribuente-evasore di essere sottoposto ad accertamento sia prossima alla percentuale "minacciata" dall'amministrazione finanziaria per chi non si adegua agli studi di settore.

si sono calcolate le percentuali di contribuenti che l'Agenzia delle Entrate avrebbe dovuto/potuto sottoporre ad accertamento induttivo (in quanto costoro hanno presentato dichiarazioni fiscali non congrue rispetto agli studi di settore)¹¹ sul totale dei lavoratori autonomi ed imprese (circa 5 milioni di contribuenti)¹² per gli anni dal 1998 al 2004. La media di queste percentuali costituisce la prima componente della probabilità di essere sottoposti ad accertamento fiscale ($p_1 = 0,38$). La seconda componente della probabilità di subire un accertamento è legata alla possibilità di essere sottoposti a controllo da parte della Guardia di Finanza. In particolare, dai dati dell'annuario della Guardia di Finanza (2006) si è calcolata una misura approssimativa della probabilità di subire una verifica fiscale. Essa è stimata dividendo il numero totale di controlli eseguiti (verifiche generali, verifiche specifiche, controlli strumentali in materia di scontrini e ricevute fiscali sugli esercenti, controlli sulle accise e controlli intracomunitari)¹³ sul totale dei contribuenti (autonomi ed imprese) a cui sono generalmente indirizzate queste forme di verifiche. Per il 2005, che è l'annualità con disponibilità di dati più prossima al periodo di riferimento della nostra analisi, essa è pari a $p_2 = 0,09$. Poiché Agenzia delle Entrate e Guardia di Finanza individuano e controllano i contribuenti in modo indipendente, il valore utilizzato per la calibrazione è dato dalla somma di p_1 e p_2 ($p = 0,47$). Per quanto l'approccio di stima sia abbastanza approssimativo, l'assenza di statistiche ufficiali rende il valore stimato una *proxy* indicativa del rischio percepito dai contribuenti di essere accertati, e quindi sanzionati, se si sceglie di operare nell'ES.

5. L'aliquota della sanzione amministrativa (σ), è calibrata attribuendo al parametro il valore minimo (1,2) previsto dalle norme tributarie italiane. Sia gli uffici finanziari che le commissioni tributarie applicano infatti usualmente il valore minimo dell'intervallo di oscillazione possibile (dal 120% al 240%). Sommando

¹¹ Lo scostamento dei ricavi o compensi dichiarati, rispetto a quelli attribuibili in base allo studio di settore, costituisce infatti una presunzione relativa grave, precisa e concordante su cui l'Agenzia delle Entrate ha la possibilità di fondare un eventuale "accertamento induttivo".

¹² Questi dati sono forniti dall'Agenzia delle Entrate e rielaborati da ARACHI G. e SANTORO A. (2007, p. 239).

¹³ Annuario GUARDIA DI FINANZA (2006) dati per l'anno 2005: verifiche generali (32.300); verifiche specifiche (66.101); controlli intracomunitari (2.423); controlli sui misuratori fiscali a carico degli esercenti (303.269) controlli sull'emissione degli scontrini fiscali a carico degli esercenti (54.261) controlli sulle accise (5.276) per un totale di 463.630 diviso i circa 5.150.000 lavoratori autonomi ed imprese.

- alla aliquota sanzionatoria anche il pagamento dell'imposta dovuta ma non versata il valore di σ con riferimento al sistema tributario italiano è pari a 2,2.
6. La probabilità che dall'accertamento fiscale derivi, oltre che la sanzione amministrativa, anche un procedimento penale (λ) è stimata sulla base dei dati forniti dalla Guardia di Finanza (2006). Come in precedenza, si utilizza il dato del 2005 per calcolare la probabilità che dal controllo della polizia tributaria derivi una denuncia all'autorità giudiziaria. In ragione della riforma delle sanzioni tributarie penali (D.lgs n. 74 del 2000 e successive modifiche), si è ristretto il caso alle sole tipologie di controlli che potrebbero dare luogo a reati tributari con dolo specifico e potenzialmente idonee a superare le soglie di impunitività individuate dalla norma quali requisiti per rendere l'illecito fiscale anche penalmente perseguibile. In particolare la probabilità è calcolata come rapporto tra il numero totale di denunce (6.120) ed il totale delle verifiche generali, verifiche specifiche, controlli sulle accise e sugli scambi intracomunitari. ($\lambda = 6.120/106.100 = 5,8\%$).
 7. In caso di denuncia penale l'evasore dovrà sostenere l'onere (psicologico e monetario) della tutela giudiziaria e dell'eventuale condanna (detentiva e pagamento spese processuali). Una stima di tale costo non è quantificabile data la variabilità dei fattori che influiscono sul *quantum* e le componenti soggettive in gioco (reddito del contribuente, tipo di condanna, precedenti penali, grado di moralità fiscale, ecc.). Con l'obiettivo di rendere la scelta di tale parametro il più neutrale possibile, si assegna ad f un valore unitario. In termini economici questo valore potrebbe indicare che, se l'operatore dell'ES è incriminato davanti alla giustizia penale, egli debba rinunciare a tutto il suo reddito (ad es. perché recluso)¹⁴.
 8. L'aliquota media d'imposizione fiscale τ è fissata pari al valore medio della pressione fiscale in Italia nel periodo dal 1982 al 2006 ($\tau = 0,396$).
 9. Il parametro che misura la qualità delle istituzioni (q) è una grandezza non osservabile e senza unità di misura. In ragione dell'intervallo di oscillazione tra 0 ed 1, la scelta più neutrale è stata, in analogia con gli altri parametri di questo tipo, quella di fissare questo indice uguale al valore mediano ($q = 0,5$).

¹⁴ Data la bassa probabilità stimata per l'Italia di essere sottoposto a procedimento penale per illecito tributario, si evidenzia come l'elasticità del tasso di crescita di lungo periodo a variazioni di f è molto limitata e non incide in modo sostanziale sulle soluzioni del modello. Ad esempio, nella calibrazione di *benchmark*, un incremento del 100% di f ($f=2$) aumenta il tasso di crescita di 0,06 punti percentuali e riduce l'ES di 1,96 punti percentuali. In questo senso l'assegnazione di valori alternativi ad f non è particolarmente rilevante per gli obiettivi di questa analisi.

Per quanto concerne la calibrazione dei due coefficienti connessi alle preferenze del consumatore inclusi nella funzione di utilità, la letteratura non è univoca sul valore da attribuire a ρ e θ (Browning e Lusardi, 1996). Ad esempio, Hansen e Singleton (1982) ottengono una stima dell'avversione relativa al rischio (θ) nell'intervallo 0,52-0,97. Epstein e Zin (1991) e Gourinchas and Parker (2002) stimano un θ rispettivamente negli intervalli 0,4-1,4 e 0,5-1,4. Hamori (1996), con riferimento al Giappone, stima un ESI ($1/\theta$) tra 1,0 e 1,4. Recentemente Engelhardt e Kumar (2009) fissano l'ESI nell'intervallo 0,37 e 1,21 con una probabilità del 95%. Per quanto riguarda invece il fattore di sconto intertemporale del consumo, esso assume valori generalmente tra il 2% ed il 4%. Turnovsky (2000), ad esempio, parametrizza il suo modello con $\rho = 0,04$. Data la molteplicità di calibrizioni proposte in letteratura, in questo lavoro si utilizzeranno come riferimento per la parametrizzazione della funzione di utilità i valori utilizzati da Bambi (2008). La preferenza per questa parametrizzazione è motivata dal fatto che il modello calibrato da Bambi (2008) presenta significative similitudini in termini sia di struttura della funzione di utilità che di approccio alla crescita endogena (modello Ak) con il modello qui proposto. Nello specifico, da Bambi (2008, p. 1029) avremo:

10. Tasso di sconto della preferenza intertemporale al consumo $\rho = 0,02$;

11. Avversione relativa al rischio $\theta = 0,80$.

La Tavola 1, sintetizza i valori di calibrazione.

TAV. 1

CALIBRAZIONE DEI PARAMETRI DEL MODELLO

PARAMETRO	SIGNIFICATO	VALORE
A	Produttività esogena del capitale	1,00
α	Produttività settore pubblico rispetto al settore privato	0,50
δ	Quota dei servizi pubblici (non escludibili agli operatori dell'ES)	0,50
q	Indice della qualità delle istituzioni	0,50
p	Probabilità di accertamento	0,50
σ	Sanzione tributaria amministrativa	2,20
λ	Probabilità di subire anche procedimanto tributario penale	0,058
f	Sanzione tributaria penale (costo psicologico/spese/lucro cessante)	1,00
τ	Aliquota fiscale (media) d'imposizione	0,396
ρ	Tasso di sconto preferenza intertemporale al consumo	0,02
θ	Avversione Relativa al Rischio (Reciproco dell'ESI del consumo)	0,80

In base a questa calibrazione l'*output* del modello (equilibrio di lungo periodo) in termini di tasso di crescita del PIL (γ) e dimensione dell'ES (S) sono rispettivamente dell'1,7% e del 33,3%, valori che sono consistenti con le medie di lungo periodo (1982-2004)¹⁵ osservate in Italia.

4. - Simulazione

Il modello calibrato può essere utilizzato per simulare gli effetti sul tasso di crescita del PIL e sulla dimensione relativa dell'ES nel lungo periodo. Simuleremo in particolare gli effetti di variazioni della qualità istituzionale, dell'efficacia dell'*enforcement* (probabilità di accertamento, aliquota sanzione tributaria amministrativa, probabilità di essere sottoposti a procedimento penale) della produttività relativa del settore pubblico rispetto il settore privato e della dimensione del *free riding* degli operatori dell'ES nell'usufruire dei servizi pubblici.

La rappresentazione grafica dell'*output* della simulazione sarà effettuata in uno spazio bidimensionale (crescita economica *vs* pressione fiscale e ES *vs* pressione fiscale) in quanto tale esposizione grafica rende immediata la relazione delle sopracitate variazioni dei parametri di *policy* con la pressione fiscale che è tra i fattori più rilevanti nella determinazione sia della crescita economica che dell'ES.

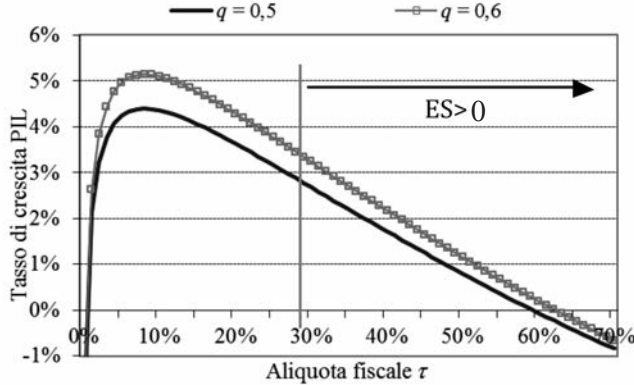
Nei grafici 1-6 si indicherà, con una barra verticale, il limite inferiore dell'aliquota fiscale (eq. (14)), che garantisce una quota di ES sul PIL interna all'intervallo]0,1[. In base alle simulazioni, ad eccezione dell'aumento di produttività relativa del settore pubblico (α) e dell'aumento della spesa pubblica non escludibile (δ), rileviamo che l'ES è presente in tutti i sistemi economici con pressione fiscale (aliquota proporzionale nominale) maggiore del 30% (35% se $\alpha = 0,6$; 23% se $\delta = 0,6$).

La prima simulazione (Grafico 1) analizza l'effetto che un miglioramento istituzionale ha in termini di variazione dei tassi di crescita economica per ciascun livello di aliquota fiscale.

¹⁵ Il periodo temporale 1982-2004 è utilizzato per rendere tale simulazione consistente con l'analisi empirica presentata nel paragrafo 5.

GRAF. 1

EFFETTO DI UN MIGLIORAMENTO ISTITUZIONALE (q)

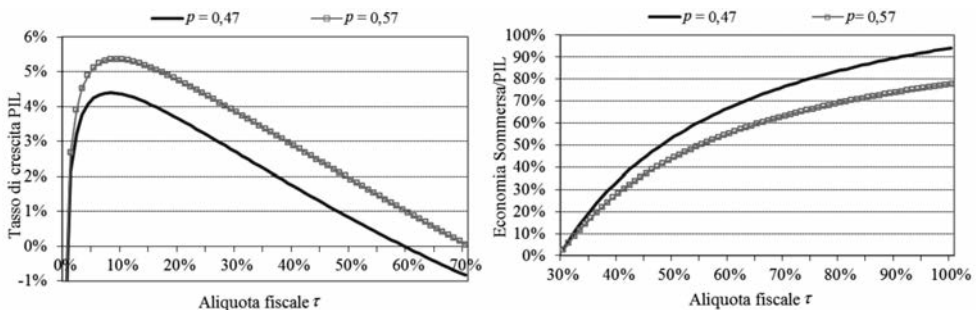


Come mostra il Grafico 1, un miglioramento del contesto istituzionale (incremento del parametro (q)) implica un maggiore tasso di crescita dell'economia ufficiale per ogni livello di aliquota fiscale. In termini numerici, significa che assumendo una pressione fiscale di circa il 40%, a cui corrisponde un livello di ES del 33%, un miglioramento del contesto istituzionale del 20% ($\Delta q = +0,1$) aumenta il tasso di crescita in stato stazionario dal 1,7% al 2,1%.

La seconda simulazione (Grafico 2) mostra gli effetti di una maggiore frequenza nei controlli sia sul tasso di crescita del PIL che sull'ES.

GRAF. 2

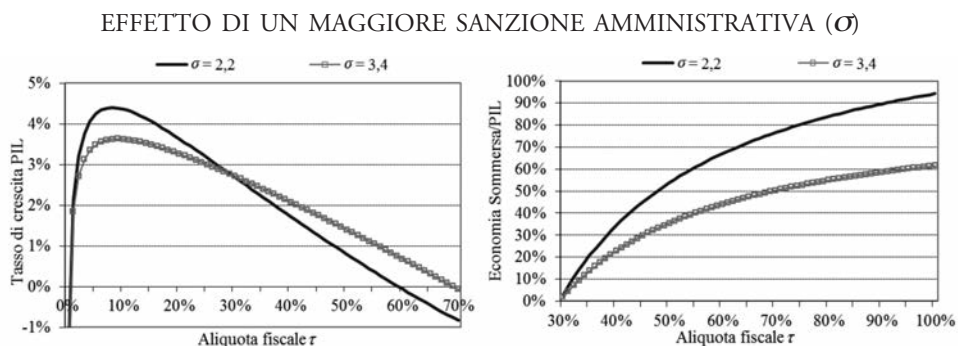
EFFETTO DI UN MAGGIORE PROBABILITÀ DI ACCERTAMENTO (p)



Partendo dalla stessa situazione ante-manovra del caso precedente – pressione fiscale del 40% e sommerso del 33% – una maggiore frequenza nei controlli di circa il 20% ($\Delta p = +0,1$) aumenta il tasso di crescita in stato stazionario dal 1,7% al 2,9%. L'effetto del maggior *enforcement* modifica anche il livello di ES in relazione al PIL di lungo periodo (eq. (10)) in misura crescente all'aumentare dell'aliquota. Con riferimento a $\tau = 0,4$ l'aumento della probabilità (anche solo percepita) di essere accertati di 10 punti percentuali riduce l'ES di circa 5 punti percentuali (ES = 28,3%).

La terza simulazione (Grafico 3) mostra gli effetti di una più severa sanzione amministrativa per punire gli illeciti tributari.

GRAF. 3

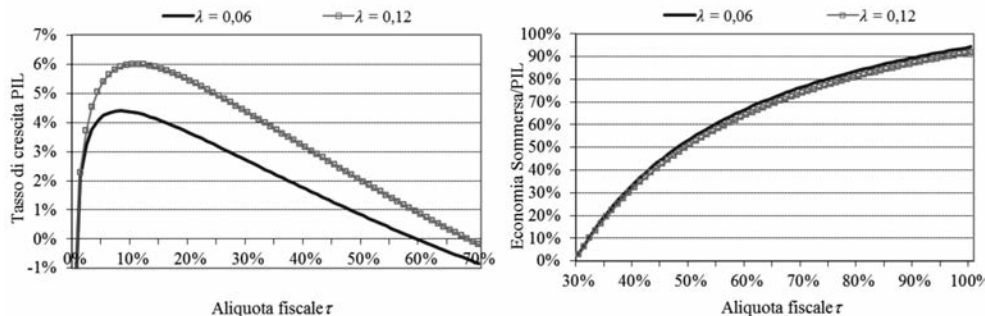


Ipotizzando l'applicazione della sanzione al livello massimo consentito dall'attuale ordinamento ($\Delta\sigma = +1,2$), l'effetto sul tasso di crescita in stato stazionario sarebbe positivo ma di limitata entità (dal 1,7% al 2,1%) mentre molto più rilevante sarebbe l'effetto deterrenza con una riduzione, per $\tau = 0,4$, dell'ES di circa 11 punti percentuali (ES = 22,0%).

Il grafico 4 riporta l'*output* della simulazione degli effetti di un'estensione della punibilità penale dei reati tributari sulla crescita del PIL e sull'ES.

GRAF. 4

EFFETTO DI UN AUMENTO DELLE SANZIONI PENALI TRIBUTARIE (λ)



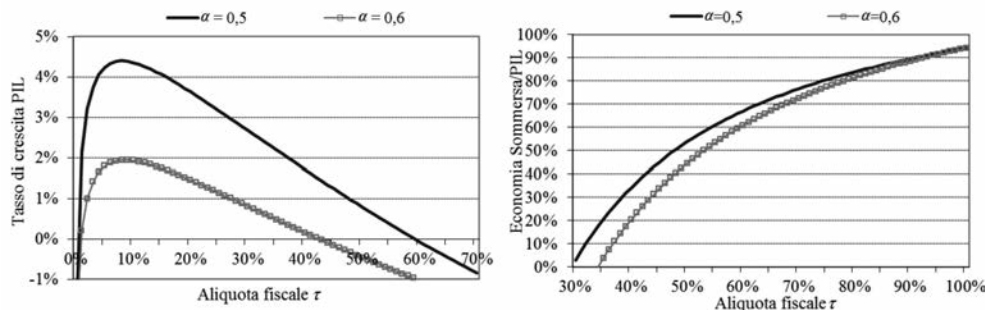
Partendo dalla situazione di *benchmark*, una maggiore punibilità dei reati fiscali dal 6% al 12% ha un notevole effetto sul tasso di crescita di lungo periodo (dal 1,7% al 3,1%), mentre più lieve è l'effetto sull'ES. Con riferimento ad un'aliquota d'imposta del 40%, l'aumento di λ di 6 punti percentuali riduce infatti l'ES di solo di 1 punto percentuale (ES = 32,3%).

In base alle precedenti simulazioni, si rileva come un aumento dell'*enforcement* ha sempre un effetto espansivo sul tasso di crescita dell'economia ufficiale mentre determina una riduzione dell'ES.

Le successive simulazioni si concentrano sulla produzione del settore pubblico. Il primo esercizio di simulazione considera gli effetti di una variazione della produttività relativa del settore pubblico sulla crescita ed ES (Grafico 5). L'*output* della simulazione mostra che, quando la produzione nel settore privato si riduce rispetto a quella del settore pubblico ovvero il settore pubblico recupera efficienza produttiva, sia il tasso di crescita del PIL ufficiale che l'ES decrescono.

GRAF. 5

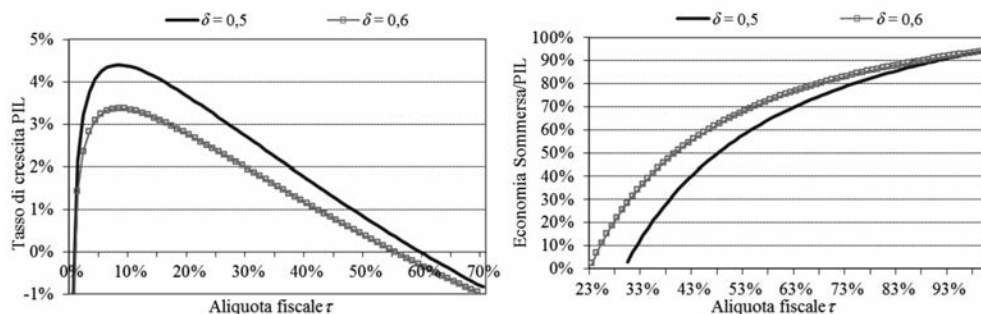
EFFETTO DI UN AUMENTO DELL'ELASTICITÀ DELL'OUTPUT RISPETTO A G/Y (α)



In termini numerici, un aumento dell'elasticità dell'*output* rispetto a G/Y del 20% ($\Delta\alpha = +0,1$) contrae il tasso di crescita del PIL di lungo periodo (dall'1,7% allo 0,2%) e riduce l'ES di circa 13 punti percentuali (dal 33,3% al 20,5%).

L'ultima simulazione analizza gli effetti di una variazione della quota di spesa pubblica utilizzata per finanziare beni e servizi non escludibili (Grafico 6).

GRAF. 6

EFFETTO DI UN AUMENTO DELLA SPESA PUBBLICA NON ESCLUDIBILE (δ)

Dalla simulazione si evince che, quando la produzione di beni pubblici non escludibili agli operatori dell'ES aumenta (aumento delle possibilità di *free riding*) del 20% ($\Delta\delta = +0,1$) il tasso di crescita del PIL di lungo periodo si riduce di mezzo punto percentuale (dal 1,7% al 1,2%). Parallelamente si registra un'espansione dell'ES, il cui valore assoluto decresce all'aumentare dell'aliquota di imposizione. In particolare con riferimento a $\tau=0,4$ l'ES aumenta dal 33,3% al 50,4%.

5. - Una verifica empirica per l'economia italiana

In questo paragrafo si propone una verifica empirica delle ipotesi utilizzate per la costruzione del modello teorico con riferimento all'economia italiana dell'ultimo trentennio. In generale, le indagini empiriche che analizzano le interazioni tra economia formale ed informale incontrano due problemi fondamentali a cui anche questa analisi non si sottrae. Il primo deriva dall'inevitabile inclusione nella regressione di variabili con evidenti errori di misurazione (ad esempio stime dell'evasione fiscale/economia sommersa, indici di *enforcement*). La seconda limitazione deriva dal fatto che, se si volesse utilizzare un modello esaustivo delle interazioni tra economia ufficiale ed ES, il numero di variabili da includere sarebbe sproporzionato rispetto alla (ridotta) dimensione campionaria condizionata

dalla scarsa quantità/qualità/estensione dei dati disponibili. Un'esaustiva rappresentazione del fenomeno comporterebbe quindi una riduzione di gradi di libertà che comprometterebbe, ulteriormente, l'attendibilità delle stime. In definitiva la presenza di errori di misurazione, la necessaria omissione di alcune variabili rilevanti e la dimensione campionaria ridotta ci impongono di valutare con cautela i risultati dell'analisi empirica.

La verifica empirica che si propone è volta a testare essenzialmente se, come ipotizzato nel modello teorico, il tasso di crescita del PIL reale *pro capite* italiano è influenzato da variazioni della pressione fiscale, del grado di *enforcement* e della dimensione dell'ES. In questo senso una conferma empirica alle ipotesi del modello teorico potrebbe, implicitamente, avvalorare le simulazioni presentate nel paragrafo precedente.

5.1 Database

Il database è costruito con riferimento all'Italia e copre un periodo che va, per le variabili senza valori mancanti, dal 1980 al 2007. La presenza di *missing values* e l'utilizzo della serie in differenze prime, riduce infatti sensibilmente il campione delle regressioni stimate che oscilleranno infatti dal 1981 al 2004 o 1991-2004 a seconda della specificazione del modello. Rinviando all'appendice 1 per il dettaglio delle fonti dei dati e le relative definizioni, in questa sezione si esplicherà il significato delle variabili utilizzate evidenziando le principali problematiche connesse alla costruzione del campione.

La variabile dipendente è rappresentata dal tasso di crescita del PIL reale *pro capite*, pubblicato nella banca dati della *World Bank*. Come *proxy* dell'aliquota fiscale proporzionale si è utilizzato l'indice di pressione fiscale. Per controllare il problema della multicollinearità con la *proxy* dell'ES e il fenomeno dell'illusione fiscale, si è utilizzato anche un indice di pressione tributaria diretta.

Per l'ES si sono utilizzate le stime di Pisani e Marigliani (2007) della base imponibile IVA evasa in proporzione alla base imponibile (teorica) complessiva. Sebbene la dimensione del valore aggiunto non dichiarato non sia perfettamente coincidente con la dimensione dell'ES, essa ne rappresenta una misura attendibile della stessa (Chiarini, Marzano e Schneider, 2011)¹⁶. Le statistiche sul grado di *enforcement* sono un dato di difficile reperimento. Per l'Italia non sono, infatti,

¹⁶ Questa stessa serie è utilizzata da CHIARINI B., MARZANO E. e SCHNEIDER F. (2011) come *proxy* dell'ES. Gli autori sostengono infatti che queste due misure potrebbero avere molto in comune, poiché l'evasione dell'IVA potrebbe essere considerata un importante strumento che agevola l'occultamento di base imponibile di altre imposte e, quindi, attività sommerse.

disponibili dati ufficiali sul numero dei controlli o sugli esiti degli stessi, né le AP preposte pubblicano in modo strutturato e con metodi omogenei informazioni connesse alla loro attività ispettiva. In questo lavoro si è tentato di aggirare il problema creando una *proxy* dell'*enforcement penale* basato sulle statistiche giudiziarie pubblicate dall'ISTAT¹⁷. Esso è calcolato come il rapporto tra il numero di soggetti denunciati dall'autorità giudiziaria per illeciti tributari (evasione fiscale e contributiva, omesso versamento di ritenute previdenziali, evasione imposte di fabbricazione sugli olii minerali e frodi su olii minerali) e la dimensione (in milioni di euro) della base imponibile IVA evasa stimata da Pisani e Marigliani (2007). L'ipotesi che sta alla base dell'uso di questo indicatore è che un incremento del numero di soggetti denunciati per reati tributari per milione di euro di imposta evasa dovrebbe indicare una maggiore efficacia dell'*enforcement*¹⁸.

Per tenere conto dei problemi connessi all'omissione di variabili rilevanti sono state incluse tre variabili di controllo volte a spiegare alcune delle determinanti della crescita che il modello teorico trascura per semplificazione. In particolare si include un indice di apertura dell'economia al commercio internazionale¹⁹, un indicatore della qualità del capitale umano ed il tasso di partecipazione alla forza lavoro. Al fine di controllare la robustezza dei risultati si sono utilizzate specificazioni alternative del modello di regressione.

5.2 Metodologia econometrica e risultati dell'analisi empirica

Dal lavoro di Nelson e Plosser (1982) è cresciuta l'attenzione nei confronti dell'eventuale presenza di *trend* stocastici nelle serie storiche. In caso di regressione statica con stimatori OLS che utilizzano variabili in livelli, l'eventuale presenza di processi integrati dello stesso ordine può far incorrere in una regressione spuria, con la conseguenza di avvalorare l'esistenza di relazioni tra variabili anche quando queste sono stocasticamente indipendenti. Una soluzione per prevenire tale problema è differenziare le variabili in modo da ottenere stazionarietà. Engle e Granger (1987) dimostrano però che non necessariamente le regressioni statiche ai livelli con variabili integrate sono regressioni spurie. Esiste, infatti, la possibilità che una combinazione lineare di processi I(1) risulti essere un processo stazionario,

¹⁷ Vari anni, vedi APPENDICE 1.

¹⁸ Questo indicatore ha un *break* strutturale nel 2001 che potrebbe minarne l'attendibilità. Il D.lgs. n. 74 del 2000 (e successive modifiche) ha infatti parzialmente depenalizzato la delittuosità connessa all'evasione fiscale e contributiva.

¹⁹ L'indice è calcolato come la somma delle importazioni ed esportazioni di beni diviso il valore del PIL a prezzi correnti.

$I(0)$ (c.d. cointegrazione). In questo caso, la relazione di cointegrazione è interpretata come la relazione statica di equilibrio di lungo periodo. Due approcci sono adottati in letteratura per verificare la presenza di cointegrazione: la procedura a due passi basata sui residui per testare l'ipotesi nulla di assenza di cointegrazione (Engle e Granger, 1987; Phillips e Ouliaris, 1990) e l'approccio sulle regressioni di rango ridotto basato su stime del sistema multivariato dovuto a Johansen (1991, 1995). La ridotta dimensione campionaria limita, ancora una volta, l'attendibilità di questi test. In particolare risulta del tutto inapplicabile al nostro campione l'approccio sistemico di Johansen.

In appendice 2 si riportano i risultati dei test univariati per la ricerca dell'ordine di integrazione e dei test di Engle-Granger e Phillips-Ouliaris per la presenza di cointegrazione tra le variabili incluse nel modello. Dall'analisi univariata si evince che tutte le variabili sono $I(1)$ mentre dai test di cointegrazione rigettiamo l'ipotesi di cointegrazione tra le variabili del modello. In base a questi risultati l'approccio più adeguato per evitare il rischio di regressione spuria è la trasformazione in differenze prime e la successiva stima con stimatori OLS.

Evitato il problema della regressione spuria, si stima un modello econometrico al fine di verificare se l'effetto di variazioni della pressione fiscale o della pressione tributaria diretta, del livello di *enforcement* penale e dell'ES sul tasso di crescita reale *pro capite* confermano, per l'economia italiana, le previsioni del modello. La non disponibilità di dati con adeguata estensione non ha reso possibile una verifica empirica anche degli altri parametri considerati nell'esercizio di simulazione sul tasso di crescita del PIL (ad es. qualità istituzioni, probabilità di accertamento, sanzione amministrativa, efficienza del settore pubblico; spesa pubblica non escludibile).

I modelli specificati sono volutamente parsimoniosi in termini di variabili incluse nella specificazione a causa della ridotta dimensione campionaria. Al fine di contemperare nella valutazione dell'*output* gli effetti della multicollinearità ed endogenità tra pressione fiscale e base imponibile IVA evasa (*proxy* dell'ES) si considera preferibile utilizzare il rapporto tra gettito delle sole imposte dirette sul PIL (modelli da IV a XII)²⁰. Una sintesi delle regressioni più significative è riportata in Tavola 2.

²⁰ In questo caso si ipotizza non esista un legame tra il livello di evasione della principale imposta indiretta (IVA) e l'aliquota media delle imposte dirette.

VARIABILE DIPENDENTE: TASSO DI CRESCITA PROCAPITE DEL PIL

	I	II	III	IV	V	VI	VII	VIII
Costante	-0,214 (-0,411)	-1,349** (-2,584)	-0,787** (-2,563)	-0,806 (-1,552)	-	-0,227 (-0,724)	-0,059 (-0,113)	0,242 (0,808)
Pressione Fiscale (τ_1)	-0,240 (-0,788)	-0,395 (-1,464)	-0,481 (-1,595)	-	-	-	-	-
Pressione Fiscale ² (τ_1) ²	-	-	0,225 (1,660)	-	-	-	-	-
Pressione imp. Dir. (τ_2)	-	-	-	0,533 (-1,902)	-0,611* (-2,022)	-0,641** (-2,404)	-0,634 (-1,185)	-0,536 (-1,442)
Pressione imp. Dir. ² (τ_2) ²	-	-	-	-0,604** (-3,169)	-0,802*** (-5,200)	-0,654** (-3,067)	-	-0,613* (-1,924)
Economia sommersa (ES)	-0,202** (-2,176)	-0,158* (-2,159)	-0,090 (-1,299)	-0,212* (-2,383)	-0,218* (-2,218)	-0,206** (-2,491)	-0,186** (-2,605)	-0,197** (-2,407)
Indice <i>enforcement</i> (λ)	-	-1,525* (-2,067)	-1,107 (-1,842)	-0,878 (-1,423)	-0,489 (-0,611)	-0,666 (-0,950)	-	-
Variabili di controllo								
Apertura Commercio Internazionale	0,285* (1,938)	0,360** (2,673)	0,338** (3,054)	0,466*** (3,805)	0,470*** (3,860)	0,457*** (3,997)	0,310** (2,468)	0,248** (2,251)
Istruzione Universitaria	0,022 (0,098)	0,339* (1,945)	-	0,224 (1,301)	0,000 (0,002)	-	-0,063 (-0,302)	-
Partecipazione Forza lavoro	0,304 (0,632)	-0,152 (0,707)	-	0,198 (0,714)	0,238 (0,658)	-	0,362 (1,127)	-
Osservazioni	22 '81-'04	14 '91-'04	14 '91-'04	14 '91-'04	14 '91-'04	14 '91-'04	22 '81-'04	24 '81-'04
R^2 corretto	0,190	0,604	0,623	0,665	0,631	0,685	0,253	0,253
Durbin-Watson statist.	1,607	2,060	2,007	1,971	2,980	1,930	1,528	1,602
LM Test Auto-correl. ^a	0,950	0,717	0,432	0,803	0,989	0,878	0,988	0,693
BPG Test Eterosched. ^b	0,726	0,998	0,997	0,978	0,603	0,673	0,674	0,979
JB Test Normalità ^c	0,648	0,385	0,667	0,778	0,744	0,560	0,764	0,652

Nota: ***indica significatività all'1%; **indica significatività al 5%; *indica significatività al 10%.

Tutte le variabili sono in differenze prime. Le *t*-statistiche sono calcolate usando *standard errors* consistenti in presenza di eteroschedasticità (WHITE H., 1980).

^a Breusch-Godfrey LM Test per correlazione seriale.

^b BPG: Breusch-Pagan-Godfrey Test per eteroschedasticità.

^c Jarque-Bera test per normalità dei residui.

Per i tre test (*a*, *b*, *c*) si riportano i *p*-value dove le ipotesi nulle sono rispettivamente: assenza di correlazione seriale fino al 2° ordine, assenza di eteroschedasticità e assenza di asimmetria e curtosi nei residui.

In base ai risultati ottenuti rileviamo come sia non statisticamente diverso da zero il coefficiente che lega la pressione fiscale al tasso di crescita del PIL (modelli I, II e III). È statisticamente significativo, invece, il legame di tipo quadratico tra crescita del PIL e pressione tributaria diretta. La relazione concava verso il basso tra queste

due variabili, conferma la natura non lineare tra crescita economica e aliquota d'imposizione prevista dalla relazione di equilibrio di stato stazionario (eq. (22)).

Il legame esistente tra tasso di crescita del PIL ed ES è negativo. Tale risultato conferma l'analisi di simulazione e supporta l'ipotesi utilizzata da Loayza (1996) per la costruzione del modello. In particolare l'autore stima, per i Paesi dell'America Latina, che ad un aumento di 1 punto percentuale dell'ES rispetto al PIL – *ceteris paribus* – corrisponda una riduzione del PIL reale ufficiale *pro capite* di 1,2 punti percentuali. In base alle nostre stime per l'Italia, nel periodo 1981-2004, un incremento dell'ES di 1 punto percentuale riduce la crescita del PIL reale *pro capite* dello 0,2% circa.

In base alle ipotesi del modello teorico, la correlazione negativa tra dimensione dell'ES e crescita economica è motivata dall'assunto che la produzione ufficiale dipende anche dai servizi pubblici soggetti a congestione a cui il settore informale, pur non contribuendo al finanziamento, può accedere per la parte di servizi non escludibili. L'ipotesi di beni e servizi pubblici "comuni" (non escludibili e rivali) utilizzata nel modello implica quindi che una maggiore ES, limitando la capacità del *policy maker* di finanziare quei beni e servizi pubblici necessari per lo sviluppo del Paese, produca una riduzione del tasso di crescita di lungo periodo.

L'indicatore di *enforcement* della giustizia penale tributaria non è statisticamente significativo (ad esclusione del modello II ma solo ad un livello di significatività del 10%). Questo risultato non supporta quindi l'ipotesi del modello teorico proposto secondo cui un maggior *enforcement* (giudiziario) aumenta il tasso di crescita del PIL. La non robustezza dei risultati a specificazioni alternative del modello di regressione, gli errori di misura nella variabile λ , ed il *break* strutturale del 2001, suggeriscono però di valutare come non conclusivo il risultato di non significatività del parametro sul tasso di crescita. Purtroppo la mancanza di dati sugli altri, e più rilevanti, parametri di *enforcement* (probabilità di accertamento, livello delle sanzioni amministrative e penali) e qualità delle istituzioni non ci consentono di testare empiricamente l'attendibilità della rilevanza dell'*enforcement* sulla crescita economica.

In conclusione l'analisi empirica presentata, sebbene condizionata dalla scarsa qualità di alcune serie di dati, corrobora le previsioni del modello teorico con riferimento alle relazioni di lungo periodo tra tasso di crescita del PIL, ES e pressione tributaria diretta.

6. - Conclusioni

Questo lavoro analizza la relazione esistente tra crescita economica ed economia sommersa in un contesto dove un ruolo rilevante è svolto dalla qualità delle istituzioni. Si è proposto un adattamento per fini simulativi del modello di crescita endogena proposto da Loayza (1996). Esso è stato integrato da una più realistica formalizzazione del sistema sanzionatorio tributario strutturato e calibrato con riferimento all'economia italiana.

L'analisi di simulazione ha consentito di valutare gli effetti che un miglioramento istituzionale, una serie di variazioni delle variabili connesse all'attività di contrasto all'ES ed alla produzione di servizi pubblici possono avere in termini di variazione dei tassi di crescita economica e di dimensione relativa dell'ES di stato stazionario.

Si è evidenziato come un miglioramento del contesto istituzionale, delle politiche di contrasto all'ES (maggiori probabilità di accertamento, maggiori sanzioni amministrative e penali, estensione della procedibilità penale per i reati tributari) determini un incremento del tasso di crescita del PIL per ogni livello di aliquota fiscale. Al contrario, un aumento della pressione fiscale, una maggiore efficienza del settore pubblico (ovvero minore efficienza del settore privato) e una maggiore possibilità di *free riding* da parte degli operatori dell'ES comportano una contrazione del tasso di crescita del PIL di lungo periodo.

Si è proposta anche una verifica empirica del modello presentato. L'analisi econometrica è stata effettuata con dati relativi all'economia italiana nel periodo 1981-2004. I risultati di tale indagine dovrebbero comunque essere valutati con cautela per la ridotta dimensione campionaria e la caratteristica di non diretta osservabilità di alcune delle variabili incluse nella regressione (ES, variabili di *enforcement*, qualità delle istituzioni), con la conseguenza che tali dati inevitabilmente presentano errori di misurazione. Ciò premesso, i risultati dell'analisi empirica evidenziano che:

- (a) il legame tra pressione fiscale complessiva e il tasso di crescita del PIL è statisticamente debole mentre statisticamente significativa è la correlazione con la pressione tributaria diretta. In linea con il modello teorico è presente un legame di tipo quadratico (concavo) tra crescita del PIL *pro capite* e aliquota media di imposizione diretta.
- (b) In relazione agli effetti di un più efficace *enforcement*, a causa dell'indisponibilità di dati, si è potuto verificare empiricamente solo l'effetto di una maggiore procedibilità penale dei reati tributari sul tasso di crescita. L'analisi econome-

trica ha mostrato una incorrelazione tra le due variabili. Si ritiene comunque che tale risultato debba essere valutato con cautela anche per la presenza di un *break* strutturale nella serie esaminata (D.lgs. n. 74 del 2000 e successive modifiche). Per l'impossibilità di testare empiricamente gli effetti delle altre, e più importanti, modalità con cui si aumenta l'efficacia delle politiche di contrasto all'ES (ad es. aumento probabilità dei controlli e inasprimento sanzioni), la valutazione del legame tra crescita del PIL ed *enforcement* deve comunque essere rinviato ad altre analisi che saranno possibili solo quando tali statistiche saranno rese disponibili dalle autorità competenti.

- (c) La relazione tra tasso di crescita del PIL ed ES è stimata con segno negativo confermando le ipotesi del modello teorico. Il legame inverso tra dimensione dell'ES e crescita economica derivava, nel modello teorico, dall'esistenza di beni e servizi pubblici non escludibili e rivali per cui, la maggiore ES, riducendo la capacità di finanziamento del *policy maker* di quei beni e servizi pubblici propedeutici allo sviluppo economico, limitava il tasso di crescita di stato stazionario. Dal modello empirico stimato risulta che per ogni incremento di un punto percentuale dell'ES vi sia una riduzione di circa 0,2 punti percentuali del tasso di crescita del PIL reale *pro capite*.

Questo risultato si inserisce in un filone della letteratura che analizza la relazione tra ES ed economia regolare. Il risultato qui ottenuto non conferma Schneider (2005) secondo cui la dipendenza del segno di correlazione tra ES e PIL ufficiale è funzione del grado di sviluppo economico. Per Schneider (2005), in presenza di un'eccessiva pressione fiscale (e regolamentazione), i benefici dell'ES sul tasso di crescita del PIL dovrebbero controbilanciare i danni derivanti dal sottodimensionamento (per gettito insufficiente) degli investimenti pubblici propedeutici alla crescita. La nostra interpretazione del risultato empirico per l'Italia è che, probabilmente a causa di un *gap* infrastrutturale maggiore rispetto agli altri Paesi sviluppati, il secondo effetto prevalga sul primo, per cui una maggiore ES riduce il tasso di crescita del PIL di lungo periodo.

In base alla nostra analisi, considerando che la dimensione dell'ES in Italia è circa del 25%-30% del PIL, l'attuazione di politiche di contrasto che la conducano a livelli medi europei (15%-20%) potrebbero incrementare il tasso di crescita di circa 2 punti percentuali.

APPENDICE 1

TAV. 3

DATABASE						
	Fonte	Media	MAX	MIN	# obs.	
Tasso di crescita del PIL <i>pro capite</i> (%)	Banca mondiale – “GDP <i>per capita</i> growth (annual %)”. World Development Indicators Online. Washington, DC. Codice: NY.GDP.PCAP.KD.ZG	1.6	4.1	-1.0	28	(‘80-‘07)
Pressione Fiscale (%) (τ_1)	Banca d’Italia - Base informativa pubblica della Banca d’Italia: tabella TUEE0170.	39.0	43.7	31.1	28	(‘80-‘07)
Imp. Dirette/PIL (%) (τ_2)	Banca d’Italia Base informativa pubblica della Banca d’Italia: tabella TUEE0210.	13.5	15.8	9.3	28	(‘80-‘07)
Indice <i>enforcement</i> giustizia penale (λ)	ISTAT (vari anni). Statistiche giudiziarie penali. Tab. 10.1 (1990, 1991) e Tab. 3.1 (1992-2004) – [numero persone denunciate per i quali l’Autorità giudiziaria ha iniziato l’azione penale per Evasione fiscale e contributiva + Omesso versamento ritenute previdenziali + Imposte di fabbricazione olii minerali (L. 739/39) + Frodi su olii minerali (L. 474/57)] / [IVA evasa in milioni di euro (Tab. 1, pag. 11 - colonna “IVA EV” (Pisani e Marigliani, 2007)]. La serie è trasformata in logaritmi.	6.6	7.7	5.4	15	(‘90-‘04)
Economia sommersa (%)	Pisani e Marigliani (2007). Tab. 1, pag. 11 - colonna “rapporto BIND su BIT”. Base imponibile evasa / Base imp. Teorica.	33.3	37.1	26.9	25	(‘80-‘04)
Apertura commercio internazionale (%)	Banca Mondiale – “Merchandise trade as a share of GDP”. L’indice è calcolato come la somma delle importazioni ed esportazioni di beni diviso il valore del PIL a prezzi correnti. World Development Indicators Online. Washington, DC.	37.4	47.8	29.0	28	(‘80-‘07)
Istruzione Universitaria (%)	Banca Mondiale – “School enrollment, tertiary (% gross)”, World Development Indicators Online. Washington, DC. Codice: SE.TER.ENRR (manca il valore per il 1982)	40.8	67.1	24.8	27	(‘80-‘07)
Partecipazione forza lavoro (%)	Banca Mondiale – “Labor participation rate, total (% of total population ages 15+)”. World Development Indicators Online. Washington, DC. Codice: SL.TLF.CACT.ZS	48.9	50.4	46.9	28	(‘80-‘07)

APPENDICE 2

Test di Radice unitaria ed analisi di Cointegrazione

La tabella 4 riporta alcuni dei test più utilizzati per stabilire la presenza di radice unitaria nelle serie univariate. In letteratura è noto che la presenza di *break* strutturali, la ridotta dimensione campionaria, la molteplicità dei *trend*, potrebbe compromettere l'attendibilità di test classici quali l'Augmented Dickey-Fuller (ADF) ed il Phillips-Perron (PP) test. Per supportare la scelta dell'ordine di differenziazione si eseguono quindi anche i test proposti da Andrews e Zivot (1992) – ZA: suggerito in caso di presenza di *break* strutturale; Kwiatkowski, Phillips, Schmidt e Shin (1992) – KPSS: che a differenza dei precedenti è un test di stazionarietà e non di radice unitaria per cui è utile per verificare la robustezza dei precedenti *output*²¹; i quattro test proposti da Ng e Perron (2001) - NG-P: che rappresentano una versione più efficiente del Phillips-Perron test modificato da Perron e Ng (1996).

In caso di risultati divergenti circa l'ordine di integrazione, si è stabilito l'ordine di integrazione in base a quanto suggerito dalla maggioranza dei test con riferimento al valore soglia del 5%.

²¹ È possibile, infatti, che l'ipotesi nulla di un test di radice unitaria non venga rifiutata per scarsa potenza del test. Se il test KPSS conferma la non stazionarietà del processo, la nostra fiducia nella presenza di una radice unitaria cresce.

ADF, PP, ZA; KPSS, NG-P TEST DI RADICE UNITARIA

Variabili	Specific. test	Trasf. serie	ADF p-v.*	PP p-v.*	ZA p-v.	KPSS (c. v. 5%)*	NG-P (critical value 5%)°			
							MZa	MZt	MSB	MPT
						Int.=0,463	Int.=	Int.=	Int.=	
						I e T=	I e T=	I e T=	I e T=	
						0,146	-8,1	-1,98	0,233	3,17
							-17,3	-2,9	0,168	5,48
Tasso di crescita PIL <i>pro capite</i>	Intercetta	Livelli	0,078	0,041	0,302	0,739	-8,96	-2,06	0,231	2,92
		Differn.	0,000	0,001	0,053	0,038	-8,91	-2,11	0,237	2,75
Pressione Fiscale	Trend e intercetta	Livelli	0,357	0,557	0,065	0,147	-4,28	-1,43	0,335	21,01
		Differn.	0,016	0,003	0,447	0,054	-12,54	-2,45	0,200	7,57
Imp. Dirette/PIL	Trend e intercetta	Livelli	0,395	0,332	0,217	1,998	-3,30	-1,27	0,384	27,26
		Differn.	0,069	0,001	0,035	0,077	-12,69	-2,46	0,195	7,58
Enforcement giustizia penale	Trend e intercetta	Livelli	0,281	0,363		0,094	-1,46	-0,78	0,532	53,95
		Differn.	0,019	0,019	-	0,052	-5,71	-1,69	0,296	15,97
Evasione fiscale	Intercetta	Livelli	0,119	0,071	0,053	0,510	-8,42	-2,05	0,244	2,91
		Differn.	0,003	0,000	0,262	0,201	-10,23	-2,26	0,220	2,42
Apertura commercio Inter.	Intercetta	Livelli	0,809	0,782	0,003	4,972	-1,74	-0,65	0,376	10,374
		Differn.	0,010	0,010	0,114	0,322	-11,77	-2,42	0,206	2,095
Istruzione terziaria	Trend e intercetta	Livelli	0,239	0,511	0,095	7,609	-0,16	-0,13	0,806	129,41
		Differn.	0,109	0,020	0,159	0,117	-11,35	-2,29	0,202	8,49
Partecipazione Forza lavoro	Intercetta	Livelli	0,566	0,516	0,000	3,029	-5,57	-1,65	0,297	4,44
		Differn.	0,179	0,004	0,039	0,238	-12,47	-2,49	0,200	1,99

Nota: *MACKINNON J.G. (1996); °NG S., PERRON P. (2001, Table 1); *KWIATKOWSKI D., PHILLIPS P., SCHMIDT J., SHIN P.C.B. (1992, Tabella 1).

Ipotesi nulla: la serie ha radice unitaria per ADF, PP, ZA, NG-P, Ipotesi nulla: la serie è stazionaria per KPSS. Per ADF e PP test inclusivo di 1 ritardo ad esclusione della variabile di *enforcement* che per la dimensione ridotta della serie ($lag=0$). Per ZA il massimo numero di *lag* incluso è 1. Per l'indice di *enforcement* la dimensione campionaria limitata non consente al *software* (Eviews 7.2) di stimare il *p*-value. Per KPSS e NG-P si utilizzano i criteri automatici per stabilire il numero di ritardi ed il metodo non parametrico di stima dello spettro, rispettivamente, Akaike info Criterion e AR-spectral OLS.

Per testare la presenza di cointegrazione sono riportate in Tavola 5 le stime OLS con le variabili in livelli dei modelli specificati in Tavola 2. Si riportano i risultati dei test basati sui residui proposto da Engle e Granger (1987, EG) e Phillips

e Ouliaris (1990, PO). Essi sono test di radice unitaria applicati ai residui ottenuti dalla stima OLS statica di variabili non stazionarie (per il test di EG il test di non stazionarietà è l'ADF, per il test di PO è il PP). Le serie storiche saranno cointegrate se i residui della regressione OLS statica sono stazionari. Il risultato riportato in tabella 5 è robusto a modifiche di specificazione del *trend* (intercetta, *trend* lineare ed intercetta, né intercetta né *trend*). Si è riportato, solo per completezza nell'esposizione, anche due (II e IV) dei cinque modelli con numero di osservazioni inferiore al valore soglia di 20 osservazioni. Per serie con numerosità così ridotta (14) l'attendibilità dei valori critici stimati non è infatti statisticamente attendibile.

TAV. 5

TEST DI COINTEGRAZIONE (ENGLE-GRANGER E PHILLIPS-OULIARIS)

	I	I	II	II	IV	IV	VII	VII	VIII	VIII
Osservazioni	22	22	14	14	14	14	22	22	24	24
Trend sp.	I	I,T	I	I,T	I	I,T	I	I,T	I	I,T
EG Test (tau-stat)	0,43	0,06	0,47	0,78	0,55	0,86	0,21	0,07	0,62	0,60
PO Test (tau-stat)	0,43	0,07	0,44	0,78	0,53	0,86	0,21	0,07	0,58	0,56

Nota: I: Intercetta; T: *Trend* lineare. Si riporta per EG e PO, i *p*-value associati alla τ -statistica dei test.

Dalla Tavola 5 si evince che i modelli stimati con OLS in livelli presentano sempre residui stazionari. Non è possibile quindi rifiutare l'ipotesi nulla di assenza di cointegrazione e l'utilizzo delle differenze prime utilizzate per le stime riportate in Tavola 2 risulta quindi appropriato.

Si evidenzia che le conclusioni a cui si perviene sia sulla base dei test di radice unitarie che, soprattutto, dall'analisi di cointegrazione debbano essere valutate con cautela per la ridotta dimensione campionaria.

BIBLIOGRAFIA

- ANDREWS D. - ZIVOT. E., «Further Evidence on the Great Crash, the Oil Price Shock, and the Unit-Root Hypothesis», *Journal of Business and Economic Statistics*, vol. 10, 1992, pages 251-270.
- ARACHI G. - SANTORO A., «Tax Enforcement for SMEs: Lessons from the Italian Experience?», *E-journal of Tax Research*, vol. 5, no. 2, 2007, pages 224-242.
- BAMBI M., «Endogenous Growth and Time-To-Build: The Ak Case», *Journal of Economic Dynamics & Control*, vol. 32, no. 4, 2008, pages 1015-1040.
- BANERJEE A.V. - DUFLO E., «Growth Theory through the Lens of Development Economics», in DURLAUF S. - AGHION P. *Handbook of Economic Growth*, Holland, Elsevier Science, 2005, pages 473-552.
- BARRO R.J. - SALA-I-MARTIN X., *Economic Growth*, New York, McGraw-Hill, 1995.
- BOVI M., «Evasione e sommerso nella contabilità nazionale», in GUERRA M.C. - ZANARDI A. (a cura di), *Rapporto 2006, La finanza pubblica italiana*, Il Mulino, 2006, pag. 241-265.
- BOVI M. - DELL'ANNO R., «The Changing Nature of the OECD Shadow Economy», *Journal of Evolutionary Economics*, vol. 20, no. 1, 2010, pages 19-48.
- BROWNING M. - LUSARDI A., «Household Saving: Micro Theories and Micro Facts», *Journal of Economic Literature*, vol. 34, 1996, pages 1797-1855.
- BUSATO F. - CHIARINI B., «Il mercato del lavoro nei modelli stocastici di crescita: una rassegna di applicazioni per l'economia italiana», *Economia Politica*, vol. 20, 2003, pages 101-141.
- CARILLO M.R. - PUGNO M., «The Underground Economy and Underdevelopment», *Economic Systems*, vol. 28, no. 3, 2004, pages 257-279.
- CHIARINI B. - MARZANO E. - SCHNEIDER F. «Tax Rates and Tax Evasion: An Empirical Analysis of the Long-Run Aspects in Italy», *European Journal of Law and Economics* (22 April 2011), pages 1-21, doi:10.1007/s10657-011-9247-6.
- CIMOLI M. - PRIMI A. - PUGNO M., «A Low-Growth Model: Informality as a Structural Constraint», *Cepal Review*, vol. 88, 2006, pages 85-102.
- DE SOTO H., *The Other Path: The Invisible Revolution in the Third Worlds*, New York, Harper and Row Publishers, 1989.
- DELL'ANNO R. «What is the Relationship between Unofficial and Official Economy? An Analysis in Latin American Countries», *European Journal of Economics, Finance and Administrative Sciences*, vol. 12, 2008, pages 185-203
- ENGELHARDT G.V. - KUMAR A., «The Elasticity of Intertemporal Substitution: New Evidence from 401(k) Participation», *Economics Letters*, vol. 103, no. 1, 2009, pages 15-17.
- ENGLE R.F. - GRANGER C.W.J., «Co-integration and Error Correction: Representation, Estimation, and Testing», *Econometrica*, vol. 55, 1987, pages 251-276.

- EPSTEIN L.G. - ZIN S.E., «Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis», *The Journal of Political Economy*, vol. 99, no. 2, 1991, pages 263-286.
- FRIEDMAN E. - JOHNSON S. - KAUFMANN D. - ZOIDO-LOBATON P., «Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries», *Journal of Public Economics*, vol. 76, no. 3, 2000, pages 459-493.
- GILES D.E.A. - TEDDS L.M., «Taxes and the Canadian Underground Economy», *Canadian Tax paper*, no.106, Canadian Tax Foundation, Toronto, 2002.
- GOURINCHAS P.O. - PARKER J.A., «Consumption over the Life Cycle», *Econometrica*, vol. 70, 2002, pages 47-89.
- GUARDIA DI FINANZA, *Rapporto Annuale Guardia di Finanza*. Roma (scaricabile dal sito <http://www.gdf.gov.it/>), 2006.
- HAMORI S., «Consumption Growth and the Intertemporal Elasticity of Substitution: Some Evidence from Income Quintile Groups in Japan», *Applied Economics Letters*, vol. 3, 1996, pages 529-532.
- HANSEN L.P. - SINGLETON K.J. «Generalized Instrumental Variables Estimation of Non-Linear Rational Expectation Models», *Econometrica*, vol. 50, 1982, pages 1269-1286.
- JOHANSEN S. «Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models», *Econometrica*, vol. 59, 1991, pages 1551-1580.
- .-, *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, 1995.
- JOHNSON S. - KAUFMANN D. - MCMILLAN J. - WOODRUFF C., «Why do Firms Hide? Bribes and Unofficial Activity after Communism», *Journal of Public Economics*, vol. 76, no. 3, 2000, pages 495-520.
- KWIATKOWSKI D. - PHILLIPS P.C.B. - SCHMIDT P. - SHIN Y., «Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root», *Journal of Econometrics*, vol. 54, 1992, pages 159-178.
- LOAYZA N.V., «The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America», *Carnegie-Rochester Conference Series on Public Policy*, vol. 45, 1996, pages 129-162.
- MACKINNON J.G., «Numerical Distribution Functions for Unit Root and Cointegration Tests», *Journal of Applied Econometrics*, vol. 11, 1996, pages 601-618.
- NELSON C.R. - PLOSSER C.I., «Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications», *Journal of Monetary Economics*, vol. 10, 1982, pages 139-162.
- NG S. - PERRON P., «Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power», *Econometrica*, vol. 69, 2001, pages 1519-1554.
- OECD, *Measuring the Non-Observed Economy: A Handbook*, OECD, Paris, 2002.

- PERRON P. - NG S., «Useful Modifications to Some Unit Root Tests with Dependent Errors and their Local Asymptotic Properties», *Review of Economic Studies*, vol. 63, 1996, pages 435-463.
- PHILLIPS P.C.B. - OULIARIS S., «Asymptotic Properties of Residual Based Tests for Cointegration», *Econometrica*, vol. 58, no. 1, 1990, pages 165-193.
- PISANI S. - MARIGLIANI M., «Le basi imponibili Iva. Aspetti generali e principali risultati per il periodo 1980-2004», in *Documenti di lavoro dell'ufficio studi dell'Agenzia delle Entrate*, no. 7, 2007 (<http://www.agenziaentrate.it/ufficiostudi>).
- SCHNEIDER F. - ENSTE D.H., «Shadow Economies: Size, Causes, and Consequences», *Journal of Economic Literature*, vol. 38, 2000, pages 77-114.
- SCHNEIDER F., «Shadow Economies Around the World: What do We Really Know?», *European Journal of Political Economy*, vol. 21, 2005, pages 598-642.
- TURNOVSKY S.J., «Fiscal Policy, Elastic Labor Supply, and Endogenous Growth», *Journal of Monetary Economics*, vol. 45, 2000, pages 185-210.
- WHITE H., «A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity», *Econometrica*, vol. 48, 1980, pages 817-838.

Il controllo dell'economia: bilancio e prospettive

Marisa Faggini*

Università degli Studi di Salerno

Governments are constantly forced to take decisions based on estimates and forecasts derived from models whose assumptions are not reflected in the daily work of economic agents. But if the theoretical construct on which they are based is at least doubtful, how effective can an indication of policy be?

The aim of this work is a synthesis of the literature on control theory with particular emphasis on the control of chaotic systems. Synthesis seems to be lacking in the literature or heavily dated not allowing to consider the opportunities offered by the new methodologies and tools.

[JEL Classification: A11; E6; E61].

Keywords: teoria dei controlli; sistemi economici; caos, non-linearità.

* <mfaggini@unisa.it>, Dipartimento di Scienze Economiche e Statistiche. L'autore ringrazia i due anonimi referee per gli utili commenti e suggerimenti che hanno contribuito a migliorare il lavoro. La responsabilità di eventuali errori ed omissioni è da attribuire esclusivamente all'Autore.

1. - Il controllo dell'economia: bilancio e prospettive

1.1 *Introduzione*

Momento centrale in ogni processo di teorizzazione è la creazione di un modello, cioè di un costrutto matematico che “*con l'aggiunta di certe interpretazioni verbali, descrive fenomeni osservati*” (Gleick, 1997, p. 268). I processi caratteristici dei sistemi che si ritiene li descrivono e li interpretano vengono così tradotti in linguaggio matematico e utilizzati per spiegarne l'evoluzione e predirne gli stati futuri. Ciò permetterà la costruzione di un modello interpretativo, nel quale si individuano gli obiettivi e il modo attraverso cui conseguirli.

La realtà descritta dal modello, tuttavia, non è che una sua rappresentazione parziale e limitata e, soprattutto, nasce da una ben precisa visione del mondo, da un ben determinato punto di vista. Cosa, come e quanto descrivere dipende dall'impostazione del ricercatore e dal desiderio-necessità di realizzare un modello quanto più realistico possibile. La parzialità e la soggettività della rappresentazione racchiusa in un modello inducono a doverne valutare l'efficacia in termini di descrizione e comprensione della realtà. È necessario cioè verificare se il modello permette di acquisire informazioni sempre più dettagliate e realistiche del fenomeno che descrive. Nei modelli di politica economica, mentre obiettivi e strumenti sono manifestazioni delle preferenze dei *policy-maker*, la forma funzionale del modello è definita dall'economista, sulla base di quella che è la sua visione del fenomeno economico.

Non pochi economisti¹, anche alla luce dell'incapacità di prevedere l'attuale crisi, stanno puntato l'indice contro modelli sviluppati e utilizzati dalla teoria economica tradizionale, i cui assunti non hanno riscontro nell'operare quotidiano degli agenti economici - individui, imprese o Stato che siano. Il fallimento delle previsioni macroeconomiche e l'inadeguatezza delle politiche economiche teorizzate, una volta applicate ai problemi reali, sono da imputare al fatto che i modelli utilizzati sono una rappresentazione troppo semplificata e semplicistica della realtà e che le ipotesi su cui si basano sono spesso non corrette. I *policy-maker* nell'implementare gli interventi di politica, come accade per i teorici della fisica classica, ritengono di poter operare su un ben definito sistema. Il sistema oggetto di studio può essere concettualmente isolato per cui nel corso del tempo non subisce sostanziali cambiamenti e attraverso una completa specificazione del sistema è possibile prevederne il comportamento futuro. In tale contesto una forza esterna produrrà un cambiamento prevedibile, mentre una piccola perturbazione non

¹ V. LUX T. - WESTERHOFF F. (2009); SCHNEIDER F. - KIRCHGÄSSNER G. (2009); COLANDER D. *et AL.* (2009) solo per citarne alcuni.

produce un effetto rilevante. Il comportamento del sistema non manifesta erraticità né cambia totalmente la sua natura qualitativa. Le modificazioni del mondo esterno e le loro conseguenze sono prese in considerazione, ma unicamente come semplici perturbazioni, temporaneamente destabilizzanti, da poter smorzare attraverso l'azione di controllo.

I governi sono costantemente spinti a prendere delle decisioni sulla base di stime e previsioni che derivano da questi modelli. Ma se il costrutto teorico su cui questi si basano è quantomeno dubbio, quale efficacia può avere l'indicazione di policy? Da ciò deriva che un intervento di politica economica produrrà un previsto e ben definito risultato. Quindi, se il sistema manifesta un comportamento diverso da quello previsto dal modello è possibile esercitare un controllo così da smorzare l'indesiderata oscillazione. Questo è quanto perseguito dall'applicazione della teoria dei controlli all'economia.

Il primo tentativo in tale direzione non fu realizzato da un economista, bensì da Tustin (1953), professore di ingegneria elettrica all'Università di Birmingham con il suo libro *The Mechanisms of Economic Systems*². L'indirizzo dato da Tustin non trovò molti seguaci tra gli economisti, ancora legati a modelli di equilibrio economico generale, in cui situazioni di disequilibrio, rispetto alle quali implementare un'azione di controllo, erano evitati evocando il processo di *tâtonnement* di Walras (1874).

Una prima applicazione della teoria dei controlli, di derivazione ingegneristica, all'economia ad opera di un economista risale al 1954 anno in cui Phillips, attraverso un modello moltiplicatore-acceleratore con *delay-time* fisso, dimostrò che l'uso di un controllo di *feed-back* determinava buoni risultati in termini di stabilizzazione dell'economia³.

Da questo primo contributo è derivata tutta una letteratura sull'applicazione della teoria dei controlli⁴ e la formulazione di politiche economiche di stabilizzazione⁵, dando inizio a quella che fu definita «*the glory of modern control*»⁶ in cui

² AOKI M. (1978).

³ «[...] it is quite likely that a monetary policy based on the principles of automatic regulating systems would be adequate to deal with all but the most severe disturbances to economic system», PHILLIPS A.W. (1954, p. 315).

⁴ v. SIMON E. (1956); THEIL H. (1957); TINBERGEN J. (1952, 1956).

⁵ Esempi si possono trovare in CHOW G.C. (1970); PINDYCK R.S. (1973b); SENGUPTA J.K. (1970).

⁶ «According to Friedland, we may label the period before 1868 as prehistoric period of automatic control; from 1868 to early 1900s the primitive period. Usually, the control theory developed between 1900 and 1960s is the classical control, and from 1960s to the present times is the so-called modern control era. Again, people divide modern control into two periods: state-variable or time-domain period (1960-1975), and frequency-domain period (1975 to present)», SUBBARAM NAIDUS D. (2000, p. 1347).

la teoria dei controlli fu considerata «[...] a mathematical supertool that, when used together with econometric models, could be substantially advance the science of economic and financial management [...] it promised an improved ability to manage short-run economic stabilisation, long-run economic growth, investment portfolios, and corporate cash positions»⁷. L'ipotesi di base era che «[...] assuming a quadratic utility functional and a linear model it is possible to obtain the optimal policy as linear feedback law, which is a particularly convenient form in which obtain the solution [...] optimal stabilization policies»⁸. Si tratta, quindi, di risolvere problemi di ottimizzazione dinamica, attraverso cui si individuano le migliori possibilità di *trade-off* implicite nel modello, variando opportunamente i parametri della funzione obiettivo⁹. In tale contesto il comportamento del *policy-maker* è individuato quale scelta, in un ben definito orizzonte temporale, di una *time-path*¹⁰ di decisioni, tenuto conto di alcuni vincoli dinamici e di condizioni di fattibilità. La scelta è condizionata da un criterio di *performance*, il quale a sua volta viene definito sulla base delle decisioni e delle conseguenze associate alla scelta.

Nel corso del tempo i modelli economici hanno raggiunto livelli più elevati di sofisticazione passando da applicazioni in cui evidente era il riconoscimento che la teoria dei controlli, così come ereditata dall'ingegneria, potesse essere validamente applicata all'economia per determinare politiche ottime,¹¹ alla consapevolezza che non si poteva parlare di adattamento automatico di una tecnica da una disciplina all'altra. Infatti, l'introduzione nei modelli econometrici di caratteristiche come le aspettative razionali, l'informazione asimmetrica, il comportamento dei *policy-maker*, l'incertezza nella sua diversa natura, ha spinto gli economisti a definirne uno specifico utilizzo al fine di tener conto di questi elementi. Si consolida negli economisti l'idea che il sistema economico poteva essere controllato, e che ciò poteva realizzarsi fornendo ai *policy-maker* dei modelli che permettessero loro di definire la politica migliore da implementare per raggiungere gli obiettivi che ritenevano più opportuni.

All'inizio del XX secolo, tuttavia, si assiste ad un'inversione di tendenza che induce a questionare la validità dei modelli fino a questo momento utilizzati per descrivere, prevedere e controllare i fenomeni economici. Infatti, da modelli li-

⁷ Citato in ATHANS M. - KENDRICK D.A. (1974, p. 519).

⁸ v. TURNOVSKY S.J. (1973, p. 79).

⁹ v. CHIARINI B. (1990).

¹⁰ v. ERMINI L. (1988).

¹¹ v. Rassegne sulla teoria dei controlli applicata all'economia si possono trovare in ARROW K.J. (1968); DOBELL A.R. (1969); AOKI M. (1974); INTRILIGATOR M.D. (1975); ATHANS M. e KENDRICK D.A. (1974); KENDRICK D.A. (1976).

neari che si ritenevano contenere in sé tutta la realtà del fenomeno sotto analisi, capaci di prevedere con precisione l'evoluzione dello stesso, si è passati a modelli non-lineari, con cui ottenere una descrizione dei fenomeni con un maggiore grado di approssimazione. Si evidenzia che i sistemi non possono essere analizzati solo localmente, ma il risultato di un'analisi può definirsi soddisfacente solo se il sistema è studiato nella sua globalità e, dunque, considerando le interrelazioni che si costituiscono tra le parti. Il sistema non è semplicemente una somma delle parti, e le sue caratteristiche non sono quelle delle parti singolarmente considerate. Inoltre, poiché il sistema può manifestare una grande varietà di comportamenti l'attenzione dello studioso non è più incentrata sulla nozione di equilibrio che tutto al più è considerato un caso particolare della sua dinamica.

Lo stimolo in tal senso viene ancora una volta dalla fisica: la meccanica classica non è più un efficace strumento per la completa descrizione del mondo. In fisica, infatti, a livello microscopico, con il principio di indeterminazione di Heisenberg, si questiona la casualità operativa (come semplice disturbo esterno), sottolineando la presenza di casualità intrinseca al sistema e la complessità esistente tra realtà osservata ed osservatore. A livello macroscopico la teoria della relatività sottolinea la complessità della relazione tra spazio e tempo. A queste va aggiunta la teoria del caos che evidenzia la possibilità che comportamenti dinamici complessi possano manifestarsi in semplici sistemi deterministici non-lineari.

L'interesse per i modelli caotici in economia¹² è determinato, oltre che dalla possibilità di modellizzare in maniera più realistica i fenomeni economici, anche dal fatto che i molti economisti che si stanno interessando ad essi si attendono nuove ed importanti implicazioni di politica economica. La teoria del caos ha, infatti, aperto una nuova strada non solo all'interpretazione dei fenomeni ma anche alla definizione di tecniche di controllo attraverso cui implementare politiche che permettano di raggiungere dinamiche efficaci ed efficienti. L'utilizzo della teoria del caos ha ridotto l'importanza del ruolo giocato da eventi esterni per spiegare le fluttuazioni osservate nelle variabili economiche, ma in più ha permesso di ottenere differenti conclusioni di politica economica utilizzando lo stesso modello strutturale. Infatti, mentre per alcuni valori dei parametri del modello è possibile ottenere soluzioni stabili tipiche dell'economia tradizionale, per altri invece, si ottengono comportamenti instabili complessi. Situazione, dunque, ben diversa dai risultati che possiamo ottenere dai modelli macroeconomici lineari dove l'implementazione di diverse politiche richiede differenti strutture teoriche.

¹² V. FAGGINI M. (2005, 2007, 2009, 2011).

Ciò considerato l'obiettivo del presente lavoro è delineare una sintesi della letteratura sulla teoria dei controlli ponendo particolare enfasi sul controllo dei sistemi caotici. Sintesi che in letteratura sembra mancare o, laddove esista, essere fortemente datata non consentendo in tal modo di cogliere le opportunità offerte non solo dai nuovi approcci metodologici ma anche dai nuovi strumenti. Ciò che si vuole evidenziare è quale contributo essi possano offrire per lo studio dei fenomeni economici.

Il lavoro è così articolato. Nel secondo e nel terzo paragrafo si introdurrà ai diversi approcci utilizzati nella teoria dei controlli tradizionale, di cui nel quarto paragrafo se ne approfondiranno i limiti. Nel quinto saranno descritte le diverse tecniche di controllo del caos. Anche per queste si procederà, nel sesto paragrafo ad una disamina dei diversi approcci, mentre nel settimo e nell'ottavo saranno presentati alcuni esempi applicativi. L'ultimo paragrafo sarà dedicato a considerazioni di sintesi.

2. - Il controllo ottimo nei sistemi lineari deterministici e stocastici

Molti processi economici possono essere descritti in termini di sistemi dinamici la cui *performance* può essere migliorata attraverso l'implementazione di politiche. Tutti i problemi di scelta intertemporale sia che il soggetto sia il consumatore, l'impresa o il governo, sono espressi in termini di risoluzione di problemi di controllo ottimo. Inoltre, nel sistema economico i metodi di controllo stocastico o adattivo trovano specifica applicazione poiché tale sistema è caratterizzato da incertezza derivante da errori delle equazioni, da parametri non noti e da errori di misurazione. In microeconomia la teoria del controllo, ad esempio, è stata applicata al problema della scelta del consumatore fra differenti beni dato un vincolo di bilancio, al problema di determinazione di ottimi livelli di produzione e di scorta di magazzino con domanda incerta, al problema di decisione del prezzo da parte delle imprese che si confrontano con una risposta incerta del consumatore¹³. Viceversa, in macroeconomia tali applicazioni sono essenzialmente volte all'individuazione di politiche di stabilizzazione e di controllo delle variabili aggregate.

Le prime applicazioni della teoria dei controlli all'analisi macroeconomica sono dovute a Phillips¹⁴ (1954, 1957); Simon (1956); Theil (1957); Tinbergen (1952,

¹³ ATHANS M. - KENDRICK D. (1974).

¹⁴ «[...] He began the construction of a series of connected glass tube models where the flow represented GNP and the feedback system represented the use of monetary and fiscal policy», KENDRICK D.A. (2002b, p. 10).

1956). In particolare, Tinbergen¹⁵ nel 1939, fu uno dei primi ad applicare l'econometria, che proprio in quegli anni stava muovendo i suoi primi passi, ad aggregati macroeconomici, stimando i parametri del modello keynesiano. Sebbene, lo stesso Keynes, per certi versi, fu critico verso quella che egli definì *black magic*¹⁶, tuttavia, da questo momento, attraverso il contributo di Tinbergen (1952), l'econometria cominciò ad essere una disciplina cui gli economisti attingeranno a piene mani. I modelli macro-econometrici di controllo ottimo, sia in un contesto deterministico che stocastico, saranno successivamente ampiamente utilizzati in politica economica¹⁷. Nelle applicazioni della teoria dei controlli in economia il modello di riferimento è il modello lineare quadratico¹⁸ (QLP). Essendo il vettore del controllo generalmente espresso quale funzione del vettore degli stati, la politica corrispondente a tale controllo potrebbe difficilmente essere implementata. Se, invece, si impone che il sistema è lineare e si utilizza un'approssimazione quadratica della funzione obiettivo, è possibile formulare il problema del controllo con funzione quadratica lineare la cui risoluzione è facilmente ottenibile¹⁹.

Un importante elemento nella formulazione di un problema di controllo ottimo è dato quindi dalla definizione della funzione obiettivo. In qualunque forma espressa essa individua e specifica le preferenze del *policy-maker*. Pertanto, una ca-

¹⁵ «TINBERGEN J. (1952, 1956) and MEADE E.J. (1951) were among the first to establish the general guidelines for policy-execution, *i.e.* the choosing of "optimal" settings of policy instruments to achieve particular "targets"», <http://cepa.newschool.edu/het/essays/keynes/macropolicy.htm>, (giugno 2003).

¹⁶ <http://cepa.newschool.edu/het/essays/keynes/macropolicy.htm>, (giugno 2003).

¹⁷ «The recent resurgence of work in the field begins with a group of deterministic models. Livsey (1971) applied optimal control to non-quadratic nonlinear continuous-time model of British economy, Pindyck (1972) applied the quadratic-linear theory to a ten-equation model of US economy, Norman and Norman (1973) applied optimal control to the non-linear Klein model of US economy, Erickson *et al.* (1970) applied the optimal control to a model of the US economy, Holbrook (1973) worked with a large model of the Canadian economy, Friedman (1971) used a model of the US economy, and Fair (1973) solved a control version of his own nonlinear econometric model of the US economy. Shupp (1972) employed a small model of the US economy to study the optimal wage control policy», ATHANS M. - KENDRICK D.A. (1974, p. 521).

¹⁸ «In the so-called linear quadratic control, the term linear refers to the plant being linear and the term quadratic refers to the performance index that involves the square or quadratic of an error, and/or control. Originally, this problem was called mean-square control problem and the term linear quadratic did not appear in the literature until late 1950s», SUBBARAM NAIDUS D. (2000, p. 1348).

¹⁹ v. SENGUPTA J. K. (1997).

ratteristica richiesta per la sua determinazione è un algoritmo²⁰ di calcolo abbastanza *flessibile* che permetta di tener conto delle variazioni dei parametri e, quindi, della determinazione delle politiche ottimali. Punto ugualmente dibattuto in letteratura è la sua forma funzionale. Generalmente, come abbiamo già detto, essa è rappresentata da una funzione quadratica²¹. Tuttavia, sebbene questa rappresenti la forma matematica più semplice per determinare la funzione obiettivo²² e sia stata utilizzata anche per problemi di controllo stocastici non-lineari, essa non sembra la struttura più adeguata a rappresentare le preferenze dei *policy-maker*²³.

Considerata questa limitazione della funzione quadratica e l'ovvia considerazione che molti obiettivi in politica economica sono definiti in termini di intervalli anziché di singoli punti, la funzione *piecewise* sembra più adeguata a rappresentare una funzione obiettivo in modo più realistico. Inoltre, la funzione *piecewise* offre una struttura più generale²⁴ per la definizione di politiche ottime. Infatti, mentre la funzione quadratica rappresenta una forma speciale della *piecewise* che si ottiene se di questa consideriamo il caso simmetrico e strettamente convesso²⁵, la *piecewise* è asimmetrica e convessa. Rispetto alla funzione quadratica e alla *piecewise*, la *pe-*

²⁰ Per un approfondimento dei metodi di calcolo applicati nella teoria dei controlli si veda POLAK E. (1973); CHIARINI B. (1990); KENDRICK D.A. (2002a).

²¹ «The earliest and still rapidly developing strand of the dynamic stochastic general equilibrium literature makes use of simple linear-quadratic models, in which agents with quadratic preferences make optimizing decisions in environments with linear production technologies. [...] Linear-quadratic models are also convenient, because a large literature provides powerful methods for solving, analyzing and forecasting with them», DIEBOLD F.X., (1998, p. 184). Dimostrazioni dell'importanza dell'applicazione della funzione quadratica lineare si possono trovare nel libro HOLT C. *et al.* (1960), nonché in PHILLIPS A.W. (1954, 1957) e HOLT C. (1962).

²² v. THEIL H. (1964).

²³ «[...] In applications to economic policy, however, the quadratic function is not a very satisfactory representation of preferences likely to be pursued by policy makers. Exact desired values y^* and X^* for given policy targets or instruments in any period may not exist in these real preferences; often, policy makers see certain variables more as constraints, in the sense of bearing an implicit loss only for values outside some range. An even more unrealistic aspect of the quadratic function is the requirement that deviations of a target or instrument variable from its desired value bear the same loss regardless of the direction of the deviation», FRIEDMAN B.M. (1972, p. 1004).

²⁴ «The generalization achieved by this function relaxes the principal unattractive requirements associated with the quadratic form, and its inclusion within the framework of the Tinbergen-Theil approach therefore renders that approach more realistic and useful for examining economic policy problems», FRIEDMAN B.M. (1972, p. 1004).

²⁵ v. CHIARINI B. (1990).

*nalty function*²⁶ risulta essere più flessibile nel rappresentare le preferenze del *policy-maker*, in quanto permette di determinare una sequenza di problemi di ottimizzazione non vincolata, anziché un singolo problema di ottimizzazione vincolata come accade nella quadratica e nella *piecewise*. Tuttavia, Cullum J., (1969) «[...] showed that in the more general case the use of penalty functions to remove all state constraints leads to a solution of the relaxed optimal control problem, which may or may not be a solution of the original problem»²⁷.

Oltre la tipologia di funzione obiettivo altro elemento fondamentale per implementazione della *policy* è la tipologia di sistema che deve essere modellizzato. Generalmente si è soliti distinguere i sistemi dinamici in sistemi deterministici e sistemi stocastici. Un sistema dinamico è deterministico se esiste un'unica conseguenza (nota) associata a ciascun stato del sistema. Il sistema è, invece, stocastico se ci sono più conseguenze scelte da una distribuzione di probabilità nota associate ad ogni stato del sistema. Per i sistemi deterministici l'assunto di base: dato lo stato iniziale del sistema e la conoscenza che di esso si ha, è possibile prevederne l'evoluzione, cioè lo stato finale. Non è possibile che accadano eventi inattesi o anomali in quanto:

- si assume che gli effetti dell'incertezza nei sistemi oggetto di studio sono così piccoli da non avere una grande influenza sul risultato finale;
- si argomenta che se l'incertezza fosse considerata la politica ottimale derivata da un simile modello non sarebbe differente da quella ottenuta dallo stesso modello senza incertezza;
- si pensa che l'introduzione di tale variabile nell'analisi renderebbe il modello non trattabile analiticamente²⁸.

Nei modelli utilizzati per descrivere questi sistemi la soluzione al problema del controllo è data da un andamento futuro delle variabili predefinito. Infatti, sulla base del principio di certezza equivalente e attraverso la minimizzazione del valore atteso di una funzione quadratica per T periodi, si ottiene che la soluzione ottima per il primo periodo è identica a quella individuata nell'andamento predefinito della variabile²⁹. Un primissimo e basilare esempio di applicazione della teoria di

²⁶ Per approfondimenti si vedano CANNON M.D. *et al.* (1970); POLAK E. (1971); ZANGWILL W.I. (1967, 1969).

²⁷ POLAK E. (1973, p. 572).

²⁸ v. KENDRICK D.A. (2002a).

²⁹ v. CHOW G.C. (1976).

controllo ottimo ai sistemi deterministici è attribuibile a Pindyck (1972)³⁰. In un mondo di certezze il *policy-maker* che ha a disposizione più di uno strumento per raggiungere un determinato obiettivo sceglierà quello che riterrà più opportuno. Questa scelta, ovviamente, sarà determinata dal fatto che la risposta della variabile obiettivo rispetto allo strumento sarà certa.

I sistemi stocastici, invece, sono caratterizzati da incertezza, ciò implica che non è possibile conoscere *a priori* quale sarà il risultato della variazione di un parametro o di una variabile, ma solo che gli effetti di questa sono identificabili attraverso una legge di distribuzione. I valori delle variabili e dei parametri che caratterizzano il modello sono analizzati con un approccio probabilistico e stimati con i relativi metodi statistici³¹.

Le reazioni delle variabili economiche agli interventi di politica economica sono complicate e imprevedibili. Da ciò la necessità di definire degli strumenti che permettano di selezionare i diversi interventi tenendo conto della reazione delle variabili fondamentali del sistema. I *policy-maker*, operando in un contesto di incertezza, non definiscono i loro piani di intervento senza tener conto di quelle che potranno essere le modifiche che nel frattempo intervengono nel sistema economico. Da ciò l'irrelevanza e l'irrealismo di valutare le conseguenze derivanti da una sequenza fissa d'interventi da parte dello Stato. In un'ottica di controllo stocastico questo problema, in cui sono incorsi molti dei modelli econometrici³² degli anni '70, può essere superato identificando una specifica funzione obiettivo: *loss function*³³.

³⁰ «[...] applies deterministic control theory to study the optimal time path for the policy variables, using a linear econometric model [...]. The welfare function used is quadratic: it penalizes the performance of an economic variable [...] by the sum of squares of its deviations from the target values during the discrete time periods of a finite planning horizon. The derivation of the optimal paths in feedback form is based on a deterministic control theory – deterministic because the random disturbances in the econometric model were ignored (treated as having zero values) and because the parameters in the linear system are assumed to be given constants not subject to uncertainty», ATHANS M. - CHOW G.C. (1972, p. 375).

³¹ V. PRASADA RAO G. (luglio 2003), <http://www.eolss.com/eolss/47a.htm>.

³² «[...] Using an econometric model, one can make projections of the key economic variables in the future periods given a set of proposed future values for the policy variables or instruments. One may then examine the nature of these projections in order to evaluate a policy proposal. This approach is deficient for two reasons. First, the dynamic response of the economic variables to a particular course assigned to the policy variables is complicated and unpredictable. [...] The second, and perhaps more important reason is due to the uncertainty of economic projections. Given the proposed time paths for the policy variables, one cannot rely on an econometric model to make imperfect predictions of the important economic variables», CHOW G.C. (1976, p. 340).

³³ V. CHOW G.C. (1976).

Diversamente dai sistemi deterministici, nei sistemi stocastici la funzione obiettivo ha specifiche connotazioni. Precisamente: la funzione è sempre quadratica lineare, ma in questo caso si minimizza un valore atteso, considerata la presenza di elementi incerti attraverso cui è possibile confrontare le diverse proposte di politica economica. Infatti, il confronto tra le risposte delle variabili del sistema all'implementazione del controllo e i valori delle stesse che si vogliono ottenere e che si sono stabiliti in sede di determinazione dei *target* è ottenuto minimizzando i valori della funzione obiettivo rispetto agli strumenti di *policy*³⁴. Questa formulazione della funzione obiettivo è particolarmente utile nel misurare le possibilità di *trade-off* fra inflazione e disoccupazione. Diversamente dai modelli econometrici che non utilizzano le tecniche di controllo ottimo, per cui tale *trade-off* può essere ottenuto solo attraverso la sperimentazione con diverse politiche, sperando di individuarne quella migliore, attraverso una sistematica variazione dei parametri, nella *function loss* è possibile individuare le migliori combinazioni d'inflazione-disoccupazione raggiungibili sulla base di un definito modello econometrico³⁵.

Tenuto conto che quando si parla di incertezza dei parametri bisogna distinguere se la *noise* è nei coefficienti delle variabili di controllo³⁶ oppure se è nei coefficienti delle variabili di stato³⁷, l'incertezza nei modelli di sistemi stocastici è così classificata:

- errore additivo nelle equazioni del sistema che considera, ad esempio effetti di eventi futuri imprevedibili;
- parametri costanti il cui valore è, però, non conosciuto;
- parametri stocastici;
- errori di misurazione.

L'*additive noise* (1) è il più semplice tipo di incertezza da analizzare con i metodi di controllo. Il problema dell'*additive error terms*³⁸ viene risolto applicando

³⁴ v. CHOW G.C. (1976).

³⁵ v. CHOW G.C. (1976).

³⁶ In tal caso appropriati metodi di risoluzione del problema di controllo ottimale sono quelli proposti da TSE E. - ATHANS M. (1970).

³⁷ Per modelli di questo tipo cfr. AOKI M. (1967); WONHAM W. M. (1967) e BRYSON A.E. JR. - YU-CHI H. (1969).

³⁸ v. KENDRICK D.A. (2002a, p. 40). Per un approfondimento sull'*additive terms* nelle equazioni di sistemi lineari, si vedano CHOW G.C. (1972) e NORMAN A.L. (1974).

la *certainty-equivalence*³⁹, utilizzato nei modelli linearizzati con disturbi additivi e funzioni obiettivo quadratiche⁴⁰.

L'incertezza al punto (2) fa riferimento a parametri incerti ma costanti. In questo caso è, però, possibile acquisire informazioni su di essi attraverso una stima di dati. Ciò permetterà non di conoscerne il vero valore, ma di conoscerne media e co-varianza. In tal caso, non si conosce il vero valore dei parametri ma la distribuzione e la considerazione della costanza implica che durante l'orizzonte temporale questo valore non cambierà. Più difficile da considerare e da analizzare è il caso in cui i parametri sono stocastici (3). L'incertezza riguarda sia i veri valori dei parametri nonché la loro evoluzione. Per determinarli, ugualmente come nel caso di incertezza dei parametri costanti, si procede ad una loro stima attraverso l'analisi dei dati. Ma poiché si tratta di *time varying parameter* le procedure di controllo attivate sono quelle del *passive learning e dell'active learning*⁴¹. I metodi attraverso cui realizzare ciò sono essenzialmente due: il *Passive Learning* e l'*Active Learning*. Con la tecnica del *Passive Learning* i nuovi dati raccolti in ciascun periodo sono utilizzati per re-stimare i parametri nel modello economico. Tale ri-stima può essere eseguita anche nel caso di errori di misurazioni. L'*Active Learning*, detto anche controllo adattivo o controllo duale⁴², è applicato quando ci sono parametri sconosciuti sia costanti che *time-varying*. In questo caso non si procede solo ad una ri-stima, ma le variabili di controllo si scelgono tenendo conto dell'esistenza di future misurazioni. L'idea è che i valori dei parametri possono essere *learned over time*⁴³. Infine, abbiamo quelli che nella letteratura econometrica vengono definiti *errors-in-variables* (4). Questo tipo di incertezza è, tra l'altro, tipico dei sistemi economici in cui non è possibile avere delle misure esatte dei processi che, invece, sono sempre caratterizzati da *noise*. Anche qui la procedura con cui acquisire informazioni è ottenuta attraverso la stima della media e

³⁹ «[...] when the criterion function is quadratic, the planning problem for the case of uncertainty can be reduced to the problem for the case of certainty simply by replacing, in the computation of the optimal first-period action, the certain future values of variables by their unconditional expectations. In this sense, the unconditional expected values of these variables may be regarded as a set of sufficient statistics for the entire joint probability distribution, or alternatively, as a set of certainty equivalents», SIMON H. (1956, p. 74).

⁴⁰ v. CHIARINI B. (1990).

⁴¹ v. KENDRICK D.A. (2002a).

⁴² v. KENDRICK D.A. (2002a).

⁴³ v. KENDRICK D.A. (1981, p. 133).

della covarianza del vettore degli stati⁴⁴.

Tenuto conto di ciò, in letteratura i metodi per implementare il controllo sono solitamente distinti in *open-loop* e *closed-loop*. I metodi *open-loop*⁴⁵ sono utilizzati nei modelli deterministici, in cui cioè non sono contemplate incertezze di nessun tipo, né sui parametri, né sulla struttura dell'economia. Ciò che si considera sono solo gli effetti di disturbo noti. Quindi, data una conoscenza generale della risposta del sistema (variabile controllata) all'*input* (variabile di controllo), il valore desiderato in sede di determinazione del problema di controllo ottimo sarà ottenuto applicando lo specifico livello necessario della variabile. Tali metodi, però, non sono considerati efficaci per l'impossibilità di inglobare shock inattesi. Infatti, una volta che l'effetto dello *shock* (l'evento inatteso) è finito il sistema continua a muoversi lungo la traiettoria ottimale precedentemente calcolata⁴⁶.

I metodi *closed-loop*, invece sono applicati in un contesto stocastico, ossia in sistemi caratterizzati da incertezza. Se per questo sono da preferire agli *open-loop*, tuttavia, l'incertezza cui fanno riferimento è essenzialmente *additive noise*. Non è, cioè, indagata l'origine né la natura del disturbo, né questo è considerato nella *feed-back rule*. Diversamente dagli *open-loop*, però, i *closed-loop* trovano la soluzione ottimale in forma di *feed-back* che determina comportamenti auto-correttivi degli strumenti. Infatti la *feed-back rule* permette di confrontare l'effetto del controllo con quello che è il *target* e, laddove ci fossero delle divergenze tra ciò che si è ottenuto e quest'ultimo, sarà possibile modificare alcuni degli strumenti al fine di ottenere l'*output* desiderato.

In alcuni casi anche l'*open-loop* prevede una *feed-back rule* (OLOF)⁴⁷. Tuttavia, tra l'operatività della *feed-back rule* degli *open-loop* e dei *closed-loop* c'è una notevole differenza. Negli OLOF essa non anticipa le misurazioni future dei parametri o delle variabili di stato, ma individua unicamente le deviazioni verificatesi tra l'*output* e il *target*. L'operare della *feed-back rule* negli OLOF permetterà di tener conto

⁴⁴ «[...] thus the engineering models include estimates of the mean and covariance of the state vector. These notions are also being adopted in economics. Certainly measurements of economic systems are also noisy, so it is reasonable to assume that although the state variables are not known exactly, estimates of their means and co-variances can be made», KENDRICK D.A. (2002a, p. 41).

⁴⁵ Appartengono alla categoria dei metodi *open loop* gli algoritmi: Newton-Raphson o Gauss Newton method, quasi Newton e Feasible Directions. Nel primo la soluzione è ottenuta attraverso le derivate delle equazioni del sistema; il secondo è particolarmente adatto quando il modello e la funzione obiettivo sono non-lineari. CHIARINI B. (1990).

⁴⁶ v. CHIARINI B. (1990).

⁴⁷ Per un approfondimento si veda TSE E. - ATHANS M. (1972) e KU R. - ATHANS M. (1973).

delle variazioni intervenute nell'*output* nel periodo successivo e sarà in questo caso, infatti, che si attuerà un aggiustamento degli strumenti. Nei metodi *closed-loop*, invece, si assume che la scelta del controllo implementato può avere degli effetti e produrre delle incertezze sugli stati delle variabili e sul valore futuro dei parametri. L'aggiustamento degli strumenti rispetto alle variazioni intervenute è operato durante il periodo stesso. Da ciò, quindi, il fatto che i metodi *closed-loop* sono anche detti metodi adattivi o *active learning* e gli *open-loop*, invece, *passive learning*⁴⁸. Si aggiunga, inoltre, che negli *open-loop*⁴⁹ il valore corrente e futuro degli strumenti è definito all'inizio dell'orizzonte di pianificazione, per cui essi sono *state-independent*⁵⁰, cioè invarianti rispetto a qualsiasi nuova informazione che possa manifestarsi durante questo periodo. Nei *closed-loop*⁵¹, invece, lo strumento di controllo dipende dalle nuove informazioni che si manifesteranno durante l'orizzonte di pianificazione. Per questo motivo tali metodi sono *state-dependent*⁵².

C'è tuttavia da dire che, sebbene la presenza della *feed-back rule* renda i *closed loop* preferibili agli *open-loop*, essi risultano molto carenti se applicati a modelli non-lineari. Ciò è da imputare essenzialmente al fatto che le equazioni di *feed-back* sono basate sulla linearizzazione rispetto alla *path* ottimale⁵³.

3. - Il controllo ottimo nei sistemi non-lineari

Esistono due tipi di non-linearità: continua e discontinua. Un sistema è caratterizzato da non-linearità continua se può essere descritto da funzioni continue ed il set di controllo è *unconstrained*⁵⁴. L'approssimazione lineare⁵⁵ di tale sistema è efficace se, ovviamente, la non-linearità non è molto forte, se il sistema rimane nelle

⁴⁸ v. KENDRICK D.A. (1981).

⁴⁹ «In open loop control the controller is blind to what actually takes place at the output end and goes on driving the plant in a fixed and predetermined manner», PRASADA RAO G. (luglio 2003), <http://www.eolss.com/eolss/47a.htm>.

⁵⁰ v. BLACKBURN K. (1987).

⁵¹ «In closed loop or feedback control the controller notices what actually takes place at the output end and drives the plant in such a way as to obtain the desired output», PRASADA RAO G. (luglio 2003), <http://www.eolss.com/eolss/47a.htm>.

⁵² v. BLACKBURN K. (1987).

⁵³ v. CHIARINI B. (1990).

⁵⁴ v. SENGUPTA J.K. (1997, p. 87).

⁵⁵ «The approximation methods use a second-order approximation of the criterion function and a first-order approximation of the system equations. The approximation problem is then the form of the quadratic linear problems [...]. The approximation QLP [n.d.r. Quadratic Linear Problem] is the solved iteratively until the results converge», KENDRICK D. A. (2002a, p. 20).

vicinanze del punto iniziale e se, sulla base di quanto sancito nel noto teorema di Hartman-Grobman⁵⁶, permette di preservare le caratteristiche dinamiche del sistema non-lineare nell'ottenuto sistema lineare. Laddove il sistema evolva lontano da questo punto iniziale, la *performance* della legge di controllo si ridurrà a mano a mano⁵⁷.

Se un sistema è caratterizzato da non-linearità discontinua (*hard non-linearity*) esso può essere descritto da funzioni discontinue e il set di controllo è *constrained*. Questi sistemi, alcuni dei quali molto diffusi in economia e in finanza, non possono essere linearizzati.

Tuttavia, in economia, sebbene una serie di applicazioni siano state realizzate utilizzando i modelli non-lineari, in particolare in un contesto deterministico, considerando una funzione obiettivo sia quadratica che non⁵⁸, e tenuto conto che i modelli non-lineari in un contesto stocastico sono di più difficile formalizzazione, il risultato di queste applicazioni è sempre stato una linearizzazione del modello.

Il problema di controllo ottimo è stato generalmente risolto calcolando un andamento ottimale della variabile, attraverso la definizione di una *nominal path*⁵⁹, in cui i parametri e i termini di errore sono stati fissati al loro valore medio, e linearizzando il modello su questa *nominal path*. Questa linearizzazione è successivamente utilizzata per ottenere un *optimal feed-back control* attraverso cui misurare le deviazioni della variabile di controllo dalla *path* nominale.

Un problema di controllo non-lineare nei modelli stocastici, invece, viene risolto implementando il problema come se fosse problema deterministico⁶⁰. L'approssimazione lineare permetterà di applicare le tecniche di controllo *standard*. Tuttavia, con questo metodo non è possibile garantire la stabilità e la robustezza del controllo in regioni fuori dall'intorno del punto considerato.

I metodi di controllo non-lineare possono essere così sintetizzati. Un primo metodo di controllo non-lineare *for experienced control designers* è detto *trial and error*. Definito il controllo, prima di implementarlo, stabilità e robustezza sono provate attraverso delle simulazioni. Tale metodo ha fondamentalmente due limiti. Innanzitutto, può essere utilizzato solo se il sistema considerato è semplice; infatti con sistemi più complicati i risultati in termini di *performance* del controllo sono insoddisfacenti. Infine, è difficile da applicare a sistemi ampi e caratterizzati da incertezza.

⁵⁶ v. MEDIO A. - LINES M. (2001).

⁵⁷ «In such systems, nonlinearities cause of poor control performance, and might even cause system instability», SENGUPTA J.K. (1997, p. 87).

⁵⁸ v. HOLBROOK R.S. (1972).

⁵⁹ v. COOPER J.P. - FISCHER S. (1975).

⁶⁰ v. COOPER J.P. - FISCHER S. (1975).

Un secondo metodo è detto *feed-back linearization*. Con la *feed-back linearization*, dopo aver eliminato gli elementi non-lineari del modello, la minimizzazione della funzione obiettivo è ottenuta applicando le tecniche *standard* di controllo lineare. Tale metodo, però, è sensibile agli errori di modellizzazione e non è robusto rispetto all'incertezza parametrica. Può essere usato solo con modelli che descrivono completamente la dinamica del sistema.

Il *controllo robusto*, invece, consiste nel combinare una legge di controllo con cui si ottimizza la funzione obiettivo con una *tracking control law* con cui spingere la traiettoria del sistema su una *path* predefinita. Tale metodo implica un *trade-off* tra *accuratezza* della modellizzazione e *performance*, ed è particolarmente utile nel caso in cui i parametri del modello cambiano rapidamente e non sono re-stimati continuamente.

Il controllo robusto è un controllo sub-ottimale, in quanto non ottimizza pienamente la funzione obiettivo; tuttavia, diversamente dal controllo ottimo *closed-loop* le cui soluzioni talvolta sono instabili, produce soluzioni stabili anche nel caso di errori nei parametri. Il principale limite di questo metodo è che l'attività di controllo connessa al *tracking* è spesso eccessiva.

Infine, l'*adaptive control*, che consiste di due momenti. Primo, scegliere la legge di controllo utilizzando il sistema descritto dalle ultime stime disponibili dei parametri; secondo, scegliere un'*adaptation law* utilizzando una *tracking error dynamics*, ossia la differenza tra la *performance* attesa e quella attuale del sistema, per aggiornare e prevedere i valori dei parametri del sistema. Tale metodo, da poter utilizzare combinato con il controllo robusto, è particolarmente adatto ai sistemi economici, in cui i parametri del modello si modificano continuamente.

4. - I limiti della teoria dei controlli

La panoramica fin qui tracciata permette di evidenziare le caratteristiche fondamentali della teoria dei controlli ed alcuni dei problemi connessi con la sua applicazione in economia. Infatti, nonostante il miglioramento delle tecniche e i considerevoli risultati conseguiti in ingegneria, la teoria dei controlli non è riuscita a realizzare in economia quanto invece, in un'analisi iniziale si riteneva potesse fare⁶¹.

⁶¹ Ci fu infatti chi «[...] naively conveyed the impression that control techniques were the answer to policy-maker's dreams, that all one needed to do was to take the Treasury model, apply control methods, and after a process in which the preferences of policy makers would be established, the budget could be read straight from the computer printout», CURRIE D. (1985, p. 286).

Dopo un primo momento di automatico adeguamento di tecniche – quanto, in effetti, accadde nelle prime applicazioni economiche – si evidenziarono le notevoli differenze che esistono tra economia ed ingegneria. In ingegneria controllare⁶² significa essenzialmente spingere il sistema verso il raggiungimento di un determinato obiettivo, per cui particolare enfasi è data alla conoscenza del sistema, e il concetto di ottimalità, quale scelta più opportuna tra un insieme di alternative, è solo un aspetto secondario. In economia accade esattamente il contrario. Considerando l'ipotesi di razionalità del comportamento economico, quale capacità dell'agente economico «[...] *to establish a complete, transitive ranking among available alternatives*»⁶³, l'enfasi è posta sull'insieme di vincoli che permettono di ridurre le alternative disponibili. Inoltre, i sistemi economici sono caratterizzati da *forward-looking behaviour*: i valori correnti delle variabili endogene sono frutto anche dei valori futuri attesi. Inoltre, in questi sistemi nel definire la traiettoria che il sistema deve perseguire per conseguire la performance desiderata, la *path* è ottimale solo rispetto al modello utilizzato e rispetto alle assunzioni cui è soggetta la funzione obiettivo. Infatti, fondamentale in questo contesto è determinare i parametri della funzione obiettivo tenendo conto della struttura del modello e delle circostanze economiche che influenzano le variabili coinvolte nell'analisi⁶⁴. Quando il sistema è caratterizzato da comportamenti *forward-looking*, la consistenza della politica ottima viene meno e, se l'obiettivo, ad esempio, era la definizione di una politica di stabilizzazione il risultato potrebbe essere addirittura opposto, ossia condurre, se implementata, a situazioni d'instabilità del sistema economico.

Quindi, dopo un primo automatico adeguamento dall'ingegneria all'economia si è passati, da parte di quest'ultima, ad una modifica del modello di riferimento e ad un'implementazione che permettesse di tener conto della specificità del sistema economico. Tuttavia, le soluzioni prospettate hanno migliorato di poco la significatività del modello di base e, di conseguenza, la definizione delle politiche da implementare.

Il primo problema affrontato è stato il riconoscimento che il sistema economico è caratterizzato da incertezza. I continui cambiamenti cui è soggetto escludono che possa essere interpretato unicamente da un modello deterministico a

⁶² «It is the specification of the target – in essence the formulation of the control problem – that generates the problem of choice. Without this specification the control problem would be indeterminate, and the opportunity for choice would not occur», ERMINI L. (1988, p. 5).

⁶³ V. ERMINI L. (1988, p. 5).

⁶⁴ V. CHIARINI B. (1990).

parametri costanti. Quindi, da un modello deterministico di causa-effetto si è passati a considerare modelli stocastici in cui il comportamento della variabile di controllo viene definito secondo una distribuzione di probabilità ed in cui i parametri variano nel tempo. Per catturare elementi del sistema non considerati esplicitamente si introduce una componente di errore che ingloba tutto ciò che è incerto. L'utilizzo di modelli a parametri variabili ha, poi, permesso non solo di considerare una struttura economica variabile, ma anche di modellizzare il comportamento degli agenti. Tuttavia, sebbene, la modellistica econometrica sia stata modificata per tener conto di elementi sempre più realistici, c'è comunque da chiedersi fino a che punto questo adeguamento sia stato significativo in termini di *performance* del controllo.

L'uso dei modelli ad aspettative razionali e a parametri variabili, attraverso cui interpretare i mutamenti nel comportamento degli agenti e del sistema economico, senza dubbio ha significato un miglioramento della modellistica in termini di veridicità e di maggiore approssimazione alla realtà, ma quanto plausibile è di per sé, ad esempio, l'assunzione di aspettative razionali⁶⁵? Quanto efficace è stata la modellizzazione dell'incertezza⁶⁶ rispetto agli *shock* futuri che interesseranno l'economia e rispetto ai cambiamenti che subirà il sistema? La creazione di un modello non è cosa semplice. Molte sono le difficoltà inerenti la sua specificazione, ossia riguardo alle assunzioni che il ricercatore fa sul reale funzionamento del sistema, nonché la misurazione e la stima dei parametri.

Per quanto riguarda il primo punto, ovviamente ciò attiene all'impostazione del ricercatore e al desiderio-necessità di realizzare un modello quanto più realistico possibile. Per quanto riguarda il secondo punto è ovvio che il realismo del modello dipende da quanto bene i parametri siano stimati e dal ruolo che essi svolgono nel generare i dati. È difficile trovare un modo di modellare le serie storiche senza tener conto che il sistema generante può essere nel frattempo cambiato⁶⁷.

⁶⁵ Un presupposto base all'ipotesi di aspettative razionali è che la distribuzione di probabilità soggettiva coincida con quella oggettiva che ha generato i dati. «Questa definizione implica (ed è implicitata da) una definizione alternativa secondo la quale gli agenti economici non commettono errori sistematici *ex post* nel prevedere il futuro», VERCELLI A. (1987, p. 83).

⁶⁶ «[...] Bayesian or decision theoretic approach is unappealing, since this in general requires us to specify prior probabilities for a large set of alternative models, to then compute posterior probabilities from our sample, and to then give a combined forecast based on a loss function of the errors of forecast which is rather arbitrary», SARGAN J.D. (2003, p. 425).

⁶⁷ «This contrasts with other scientific areas, such as earth sciences or meteorology or astronomy/cosmology, where models are large with many variables, but the length of sample time periods used to estimate these models is relatively longer», SARGAN J.D. (2003, p. 424).

Inoltre, la modellistica della teoria dei controlli è fondata su un'assunzione di base che è la linearità del sistema economico. Sebbene vi sia una diffusa consapevolezza che il sistema economico sia non lineare, l'uso dell'ipotesi di linearità, più che per l'individuazione di una caratteristica specifica del modello economico, in economia è stata utilizzata per la sua semplicità⁶⁸.

Può essere il modello di riferimento lineare e possono essere le formulazioni matematiche della funzione obiettivo quadratiche o *piecewise*? L'ottimizzazione quadratica lineare che caratterizza la maggior parte delle applicazioni econometriche «*is undoubtedly attractive on analytical and computational grounds*»⁶⁹, ma essa «[...] *can be highly deficient in areas where economic behaviour is dominated by asymmetric costs of adjustment, irreversibilities, transactions costs or institutional and physical rigidities*»⁷⁰.

I modelli econometrici elaborati fino agli anni '80 sono approssimazioni della realtà, il cui valore descrittivo e predittivo è accettabile solo se si considera che i dati rimangono in qualche insieme limitato e che l'economia si trovi in una condizione di stabilità. Essi sono caratterizzati da una complessità che meglio si qualificherebbe adottando il termine di complicazione, esogena e non endogena, le cui cause non sono dipendenti dalla struttura del modello, ma derivano da fattori esterni ad esso. La complessità/complicazione di questi modelli è, dunque quantitativa e non qualitativa, ossia conseguenza di eccessiva informazione che non potendo essere trattata dallo scienziato⁷¹ viene inglobata nel termine di errore.

A questo punto c'è da chiedersi: nonostante i "ritocchi", cosa non ha funzionato nell'applicazione della teoria dei controlli all'economia, la tecnica o il modello di riferimento? Perché la teoria dei controlli non ha dato i risultati previsti? Perché una teoria fallisce?

L'esperienza empirica in ambito macroeconomico ha evidenziato i limiti alla nostra reale possibilità di fare previsioni in campo economico. Infatti, nonostante l'enorme quantità di ricerche, gli sviluppi di metodi econometrici sempre più

⁶⁸ Ad esempio PESARAN M.H. - POTTER S.M. (1992, p. S2) sottolineano che: «[...] autoregressive models can be estimated using standard regression packages and there is now a wide range of computer software packages available to estimate models with linear moving average components; the dynamics contained in estimated models can be completely characterized by their impulse response functions and directly related to linear models of the macroeconomy».

⁶⁹ v. PESARAN M.H. - POTTER S.M. (1992, p. S2).

⁷⁰ v. PESARAN M.H. - POTTER S.M. (1992, p. S2).

⁷¹ v. MATEOS DE CABO R. *et al.* (2002).

complicati e la disponibilità di dati la cui qualità è via via migliorata⁷², la *performance* delle previsioni degli aggregati macroeconomici è migliorata di poco.

Fino al 1980 i modelli econometrici hanno ricoperto un ruolo centrale nella previsione degli andamenti futuri delle variabili e nel loro controllo per il conseguimento di determinati obiettivi.

Astraendo i fenomeni d'interesse e traducendoli in forma matematica l'econometria ha analizzato le relazioni tra le variabili economiche. Notevoli sforzi e risorse sono stati impiegati per costruire modelli macroeconomici con numerose equazioni statisticamente calibrate, il cui obiettivo era quello di descrivere isolatamente un aspetto del comportamento del sistema economico. Dall'insieme di questi modelli e dei risultati da essi raggiunti si poteva avere la descrizione e l'andamento dell'intera economia⁷³.

Tutto questo, però, è stato fatto: utilizzando un insieme incompleto di fattori determinanti⁷⁴; costruendo modelli con variabili non osservate, la cui stima è ottenuta da dati la cui significatività è dubbia sia perché inficiati da errori di misurazione sia per l'aggregazione che di essi è fatta; ottenendo così correlazioni spurie utilizzando *proxy* di variabili; non distinguendo gli effetti di variabili multicollinari; assumendo forme funzionali lineari non conoscendo l'appropriata dimensione del regressore, inferendo non correttamente le cause dalle correlazioni; confondendo il significato statistico con quello economico dei risultati; fallendo nel porre in relazione la teoria economica all'econometria; utilizzando non adeguate ipotesi di esogeneità e di grandezza del campione.

Tra la fine degli anni '80 e gli inizi anni '90 l'uso di nuovi strumenti matematici ha evidenziato l'importanza dei fenomeni non-lineari e di fenomeni che le usuali tecniche econometriche etichettavano come casuali. L'assunzione di non-linearità e la conseguente introduzione della Teoria del caos hanno permesso di tener conto di comportamenti ciclici stazionari non derivanti da interferenze, ma proprie della dinamica del sistema, di comportamenti asimmetrici, di comportamenti volatili, di comportamenti discontinui caratterizzati da effetti soglia⁷⁵.

⁷² v. PHILLIPS P.C.B. (2003).

⁷³ v. HENDRY D.F. (1980).

⁷⁴ v. HENDRY D.F. (1980).

⁷⁵ v. MATEOS DE CABO R. *et AL.* (2002).

5. - Controllare l'imprevedibile e l'incontrollabile

«A chaotic motion is generally neither predictable nor controllable. It is unpredictable because a small disturbance will produce exponentially growing perturbation of the motion. It is uncontrollable because small disturbances lead only to other chaotic motions and not to any stable and predictable alternative»⁷⁶. Quest'affermazione individua chiaramente l'opinione sui sistemi caotici diffusa nella comunità scientifica fino al 1990.

Il lavoro di Ott, Grebogi, Yorke (1990), rappresenta la svolta nell'ambito della teoria dei sistemi caotici. Fu, infatti, descritta la possibilità di controllare il caos proprio sfruttando le caratteristiche che fino a quel momento erano state indicate come la causa principale della sua incontrollabilità: la sensibilità alle condizioni iniziali che invece rappresenta l'elemento chiave su cui si impernia la teoria del loro controllo.

Altre caratteristiche dei sistemi caotici sfruttate dalla teoria dei controlli sono la presenza di orbite periodiche instabili (UPO) e l'ergodicità della dinamica. La presenza di orbite instabili di diverso periodo consente che, nell'individuare, sulla base di un criterio di miglior *performance*⁷⁷, l'orbita da stabilizzare o l'orbita verso cui indirizzare la dinamica del sistema, sia possibile scegliere tra un'infinità di comportamenti, di orbite, naturalmente presenti nei sistemi caotici, ma soprattutto senza che ciò implichi un cambiamento della configurazione globale del sistema. Nel caso di un sistema non caotico, per poter raggiungere un desiderato obiettivo bisogna modificare proprio la configurazione globale del sistema. Inoltre è possibile scegliere una dinamica diversa semplicemente perché le condizioni di operatività del sistema sono cambiate⁷⁸.

La proprietà di ergodicità, con la quale è possibile individuare quante volte una traiettoria visita una regione dell'attrattore⁷⁹, garantisce che, attendendo per

⁷⁶ V. DYSON F. (1988, pp. 182-184).

⁷⁷ V. OTT E. *et AL.* (1990).

⁷⁸ «Few years ago, NASA scientists used only small amounts of residual of hydrazine fuel to send the spacecraft ISEE-3/ICE more than 50 million miles across the Solar System, thereby achieving the first scientific cometary encounter. This feat was made possible by the sensitivity of three-body problem of celestial mechanics to small perturbations, and would not have been possible in a nonchaotic system, in which a large effect typically requires large controls», SHINBROT H. *et AL.* (1993).

⁷⁹ Un sistema esiste in uno spazio delle fasi R^n -dimensionale. Dopo un certo periodo di tempo si stabilirà sempre in R^n , ma il suo comportamento sarà individuabile in un'area dello spazio più piccola. Quest'area rappresenta l'attrattore del sistema. Nei sistemi dinamici il concetto di attrattore rappresenta il contr'altare della nozione classica di equilibrio ma non è ciò che attrae sempre, ma ciò che impedisce dal volare lontano.

un certo periodo di tempo, la traiettoria caotica si avvicinerà ad una qualsiasi delle orbite che formano l'attrattore. Se una traiettoria si avvicina ad un'orbita, vi rimarrà vicina finché non verrà respinta a causa della sua instabilità scomparendo in *the chaotic sea*⁸⁰.

Il controllo del caos può essere diviso in due categorie:

- a) soppressione del caos;
- b) aumento del caos o *anticontrol of chaos*⁸¹.

In particolare quest'ultimo approccio, benché abbia attratto l'attenzione della letteratura scientifica solo negli ultimi anni, sembra essere la frontiera dei futuri sviluppi in questo settore. Esso sfrutta la capacità del sistema di auto-organizzarsi per ottenere ordine spazio temporale dal comportamento caotico delle singole unità del sistema.⁸² Qui non potrà essere trattato poiché, purtroppo, pochi sono i lavori sia teorici che empirici e non vi sono ancora applicazioni economiche.

L'approccio volto alla soppressione del caos (procedure di stabilizzazione), invece, benché di più difficile applicazione, è sicuramente quello più affermato⁸³.

Attraverso lo sfruttamento di queste caratteristiche, le tecniche di controllo del caos permettono di operare su una traiettoria caotica in prossimità dell'orbita che si vuole stabilizzare e, quindi, attraverso una piccola modificazione del parametro di controllo indurre il sistema alla dinamica desiderata⁸⁴.

La procedura attraverso cui ottenere la stabilizzazione dell'orbita periodica instabile consiste innanzitutto nell'esaminare la dinamica del sistema per individuare le orbite instabili di basso periodo⁸⁵ e gli eventuali punti fissi sull'attrattore. Scegliendo l'orbita periodica instabile che produce la miglior performance la si stabilizza attraverso l'uso di algoritmi che impongono piccole perturbazioni al parametro del sistema⁸⁶ oppure introducendo una forza additiva esterna all'originale sistema caotico⁸⁷.

⁸⁰ v. KANTZ H. - SCHREIBER T. (1997, p. 224).

⁸¹ v. ANDRIEVSKII B.R. - FRADKOV A.L. (2003).

⁸² v. MIRAMONTES O. *et AL.* (1993).

⁸³ v. SINGER J. *et AL.* (1992).

⁸⁴ v. BISHOP S.R. *et AL.* (1998).

⁸⁵ «According to recent result of Hunt and Ott (1996) the optimal value of an average of a given smooth function of the system state [...] is typically achieved by a low period unstable periodic orbit. Consequently, it makes sense to focus on such low-dimensional orbits», KOPEL M. (1997, p. 281).

⁸⁶ Si vedano su questo punto OTT E. *et AL.*, (1990); LIMA R. - PETTINI M. (1990); BRAIMAN Y. - GOLDBIRSCHE I. (1991).

⁸⁷ Si vedano su questo punto CHEN G. - DONG X. (1993); PYRAGAS K. (1992); JACKSON E.A. - GROSU I. (1995); BREEDEN J.L. (1994); TIAN Y. *et AL.* (2000).

Come per le tecniche di controllo delineate nel paragrafo 2, tra i metodi di stabilizzazione o di soppressione del caos distinguiamo i metodi *closed-loop* in cui la perturbazione applicata è determinata dallo stato del sistema e quelli *open-loop* in cui la perturbazione è indipendentemente dallo stato del sistema⁸⁸.

6. - *Open loop* o metodi di *non feed-back* e *closed loop* o metodi di *feedback*

I metodi *open loop* partono dalla constatazione che perturbazioni periodiche o stocastiche possono produrre rilevanti cambiamenti nella dinamica di un sistema caotico conducendo eventualmente alla stabilizzazione di alcuni comportamenti periodici.

Questa categoria di controlli include quelle strategie che considerano l'effetto di perturbazioni esterne (indipendenti dalla conoscenza della dinamica attuale del sistema) sull'evoluzione del sistema.

Un primo esempio⁸⁹ di controllo *open-loop* applicato ai sistemi caotici è da attribuire a Hübler, Lüscher (1989), i quali dalla conoscenza della dinamica di un modello suggerirono una modificazione di questa tale che la dinamica obiettivo diventasse una soluzione stabile. Per ottenere ciò è necessario avere un'equazione del modello per la dinamica e bisogna essere capaci di modificare sperimentalmente la *driving force* di queste equazioni con una modificazione abbastanza ampia. Quindi, attraverso una forza aperiodica specificamente designata, precedentemente calcolata utilizzando un modello globale del sistema, si indirizza il sistema all'andamento desiderato.

Questi approcci, pur sfruttando in modo ottimale le caratteristiche tipiche dei sistemi caotici, sono tuttavia limitati dal fatto che la loro azione non è *goal-oriented*, cioè lo stato finale periodico non può essere deciso dal controllore.

Appartengono, invece, alla categoria dei *closed loop* l'OGY, il cosiddetto OPF (*Occasional Proportional Feedback*) proposto da Hunt (1991) e Showalter *et al.* (1993), il metodo di Huebler, Plapp (1990) e Pyragas (1992), i quali applicano un *delay feed-back* ad una variabile del sistema.

Tutti questi metodi sono *independent model* nel senso che la conoscenza del sistema necessaria per selezionare la perturbazione è ottenuta osservando il sistema per un adeguato tempo di apprendimento⁹⁰.

⁸⁸ v. MIRUS K.A. - SPROTT J.C. (1999).

⁸⁹ v. HÜBINGER B. *et al.* (1994).

⁹⁰ v. BOCCALETTI S. *et al.* (2000).

La conoscenza dello stato del sistema non significa conoscenza della *law of motion* che lo regola, quanto piuttosto conoscenza della dinamica locale e della posizione delle orbite; conoscenze queste che possono essere acquisite direttamente dall'analisi empirica dei dati⁹¹.

7. - Il metodo Ott-Grebogi-Yorke

Il metodo Ott-Grebogi-Yorke (OGY) è contrassegnato da una serie di caratteristiche specifiche. Innanzitutto, la stabilizzazione dell'orbita periodica instabile (UPO) è ottenuta attraverso una piccola perturbazione del parametro poiché indirizza il sistema ad un andamento che è inerente la sua dinamica; in secondo luogo, la presenza di numerose e differenti UPO, cui corrispondono diversi andamenti del sistema, permette un'ampia scelta riguardo il comportamento che si vuole che il sistema assuma. Questi differenti comportamenti possono essere raggiunti attraverso piccole modifiche nel controllo di *feed-back* senza cambiare il regime globale del sistema⁹². L'idea base è quella di individuare alcune delle UPO di basso periodo o *steady state* instabili sull'attrattore e, attraverso l'algoritmo di controllo, muovere il sistema da un andamento caotico all'orbita scelta sulla base di un criterio di miglior performance del sistema⁹³. Data la natura ergodica del sistema si attende che la traiettoria caotica passi nelle prossimità dell'orbita su cui si desidera che essa si muova, dopodiché si applica una piccolissima variazione al parametro di controllo in modo tale da stabilizzare il sistema sull'orbita desiderata.

Infine, e questa rappresenta la più importante delle caratteristiche del metodo OGY, la sua applicazione non richiede un modello globale del sistema; infatti, ciò che è determinante è la conoscenza della dinamica del sistema nelle prossimità dell'orbita da stabilizzare. Una delle difficoltà dell'OGY riguarda la possibilità di acquisire informazioni dettagliate sulle orbite periodiche da stabilizzare e sui suoi "*manifold*"⁹⁴ stabili ed instabili. La dinamica locale che include l'UPO può essere ottenuta direttamente dall'analisi dei dati, utilizzando le tecniche dell'*embedding parameter*⁹⁵ e costruendo una *Poincaré section*⁹⁶ o il *close returns*⁹⁷. Una volta esa-

⁹¹ v. BOCCALETTI S. *et AL.* (2000).

⁹² v. HÜBINGER B. *et AL.* (1994).

⁹³ v. OTT E. *et AL.* (1990).

⁹⁴ v. TIAN Y. P. - CHEN G. (2001).

⁹⁵ Per un esempio si vedano SO P. - OTT E. (1995); OTT E. *et AL.* (1990); SO P. *et AL.* (1996).

⁹⁶ v. OTT E. *et AL.*, (1990).

⁹⁷ «If we are given an orbit of a dynamical system $v(n)$, we can locate the unstable periodic orbits ./.

minate posizione e stabilità attraverso il metodo dei *close returns* e della *Poincaré section*, si sceglie l'orbita che permette di ottenere la performance desiderata del sistema.

L'utilizzo di dati osservati per poter controllare un sistema caotico è particolarmente importante, in quanto per i sistemi caotici una stima dettagliata delle equazioni del processo spesso non è possibile. Questo metodo, quindi, può essere applicato anche quando l'equazione del sistema non è nota⁹⁸. Ciò non è invece possibile con i tradizionali metodi di controllo la cui applicabilità presuppone la conoscenza delle equazioni del sistema.

7.1 Un esempio di applicazione dell'OGY

L'applicazione del metodo OGY può così essere schematizzata:

1. innanzitutto, la dinamica del sistema deve essere nota e rappresentabile attraverso una funzione non lineare. Diversamente è possibile ottenerla sperimentalmente dai dati applicando le tecniche dell'*embedding parameter*. Nel nostro caso si utilizzerà una mappa quadratica del tipo $x_{n+1}=f(x_n, a)$ dove a è il parametro di controllo;
2. successivamente, si identifica l'orbita instabile su cui si vuole stabilizzare il sistema. Per cui se l'orbita prescelta, ad esempio, è di periodo T significa risolvere l'equazione $x_F=f^T(x_F, a)$ dove x_F rappresenta il punto fisso;
3. trovato il punto fisso si linearizza la mappa intorno ad esso. Questo punto è di particolare interesse in quanto la linearizzazione intorno al punto fisso non ci obbliga a conoscere con esattezza la mappa $x_{n+1}=f(x_n, a)$, ma è sufficiente conoscerla solo nell'intorno del punto fisso $x_F=f^T(x_F, a)$;
4. infine, si calcola e si applica a ciascuna iterazione la quantità per mantenere il sistema sull'orbita.

Un esempio può meglio chiarire quando detto. Consideriamo la mappa logistica⁹⁹:

within it by asking former recurrences of orbits. For this we ask when, if at all, an orbit returns to a close neighbourhood of its value at some time n . Close is defined using a small, positive number ε : for each data vector $y(n)$ in the original dynamics find the smallest time index $k > n$ such that $|y(n) - y(k)| < \varepsilon$, and using Euclidian distance in the $y(n)$ space is fine. This time $k - n$ is the recurrence time of the orbit», ABARBANEL H.D.I. (1996, p. 135).

⁹⁸ V. YAGASAKI K. - UOZUMI T. (1998).

⁹⁹ V. BOCCALETTI S. *et AL.* (2000).

$$(7.1.1) \quad x_{n+1} = ax_n(1 - x_n) \quad x \in [0, 1]$$

dove a che rappresenta il parametro di controllo è definito nell'intervallo $[0, 4]$.

La (7.1.1) è rappresentata geometricamente da una famiglia di parabole che, per $a > 0$ hanno la concavità rivolta verso il basso e intersecano l'asse delle ascisse nei punti $x=0$ e $x=1$, entrambi indipendenti dal parametro a ; hanno come asse di simmetria la retta $x=1/2$; il vertice, invece, dipende dal parametro a e si trova sull'asse di simmetria. Pertanto le parabole della famiglia (7.1.1), con l'aumentare del parametro a , presentano un vertice sempre più elevato. Non solo, ma per specifici valori a il sistema è caratterizzato da configurazioni -punti fissi stabili ed instabili-, che rappresentano gli attrattori del sistema, oppure da andamenti caotici, caratterizzati dal fatto che l'attrattore non è più un punto fisso ma un'intera regione. Infatti vediamo che per $0 < a < 1$ l'attrattore della mappa è il punto fisso stabile $x=0$, il sistema prima o poi convergerà a zero; per $1 < a < 3$ l'attrattore è il punto fisso stabile $x_F = 1 - 1/a$; un punto verso cui il sistema convergerà ma l'attrattore questa volta non è zero, ma è un numero che non dipende dalle condizioni iniziali ma dal parametro a ; per $3 < a < 1 + \sqrt{6}$ $x_F = 1 - 1/a$, invece, è instabile e l'attrattore è un'orbita stabile di periodo 2^{100} . Quando il valore di a aumenta ulteriormente otteniamo una sequenza di biforcazione con raddoppio di periodo. Abbiamo cioè punti di ciclo-2, ciclo-4, ciclo-8, ciclo 16 e così via. Quando il parametro assume valori superiori a 3,57 appare il caos. A partire da questo valore, chiamato anche numero di Feigenbaum, non c'è più convergenza e non ci sono più punti attrattori, ma l'attrattore è un'intera zona.

Consideriamo il caso in cui $a_0 = 3,8$ valore, questo, in corrispondenza del quale siamo in piena regione caotica. Supponiamo che il nostro obiettivo sia di stabilizzare la traiettoria caotica, ossia di spingere la traiettoria del sistema, originatasi da una condizione iniziale x_0 , sull'orbita stabile di periodo 2, ossia sull'attrattore del sistema in corrispondenza del punto fisso instabile $x_F = 1 - 1/a_0^{101}$.

Per ottenere ciò ad ogni iterazione dobbiamo far variare il parametro di controllo a_0 nell'intervallo $[a_0 - \delta, a_0 + \delta]$ dove $\delta < 1$, ossia la variazione dovrà essere piccolissima e positiva. È evidente che la particolarità e la significatività del con-

¹⁰⁰ v. GREBOGI C. - LAI Y.C. (1999).

¹⁰¹ Il valore di questo punto fisso sarà $x_F \approx 0,7368$ e quello di alcune delle corrispondenti componenti dell'orbita saranno $x_1 \approx 0,3737$ e $x_2 \approx 0,8894$ ricordando che $x_1 = f(x_2)$ e $x_2 = f(x_1)$.

trolo caotico è la piccola variazione del parametro di controllo per stabilizzare il sistema sull'orbita prescelta, ossia l'utilizzo di piccoli *input* per raggiungere ampi *output*. Infatti, se volessimo stabilizzare la traiettoria del sistema, nel caso del nostro esempio, dovremmo modificare il parametro a da 3,8 a 2,0, dovremmo cioè dirigere la traiettoria caotica in corrispondenza del punto fisso stabile $x_F=1-1/a$. Così facendo, però, innanzitutto modificherebbero la struttura del sistema eliminando il caos. Abbiamo visto, infatti, che per valori di $1 < a < 3$ il sistema non presenta comportamenti caotici, e quindi non più gestibile con le relative tecniche di controllo. In secondo luogo, sarebbe necessaria un'ampia variazione del parametro, intervento che niente di diverso avrebbe dalla metodologia del controllo tradizionale. L'aspetto peculiare del controllo dei sistemi caotici è, invece, quella di conseguire un'ampia performance con l'impiego minimo di risorse. L'idea, infatti, è quella di realizzare piccole variazioni di $a=a_0+\Delta a$ in ciascuna iterazione in modo tale che la differenza tra la traiettoria del sistema e l'orbita *target* sia estremamente piccola, prossima allo zero.

Ritornando al nostro esempio, fino ad ora abbiamo realizzato quanto specificato al punto 1), ossia scelta della funzione (7.1.1). Ora supponiamo che l'orbita che permetta al sistema di conseguire la migliore *performance* sia quella in corrispondenza del punto fisso instabile $x_F=1-1/a$, ossia l'orbita stabile di periodo. Individuata l'orbita, poiché la sua posizione dipende dal parametro di controllo, assumendo che la dinamica locale non vari molto in Δa , è possibile rappresentare la mappa logistica in prossimità dell'orbita *target* attraverso una sua linearizzazione¹⁰².

Indichiamo con x_n la traiettoria del sistema e con $x_{(i)}$, per $i=1, \dots, m$ il nostro obiettivo, cioè l'orbita di periodo m , dove $x_{(i+1)}=f(x_{(i)})$. Poiché il nostro obiettivo è che x_{n+1} si trovi in prossimità di $x_{(i+1)}$ per cui $x_{n+1}-x_{(i+1)}=0$, attraverso l'espansione in serie di Taylor di questa differenza, di cui trascuriamo i contributi del secondo ordine, la dinamica linearizzata in prossimità del componente $i+1$ è così ottenuta:

$$(7.1.2) \quad x_{n+1} - x_{i+1} = \frac{\partial f}{\partial x}(x_n - x_i) + \frac{\partial f}{\partial a} \Delta a_n = 0$$

dove le derivate parziali sono calcolate in corrispondenza di $x=x_i$ e $a=a_0$.

Risolviendo la (7.1.2) otteniamo

¹⁰² DUBÉ L.J. - DESPRÉS P. (2000).

$$(7.1.3) \quad a_0(1-2x_i)(x_n-x_i)+x_i(1-x_{i+1})\Delta a_n=0$$

Da cui consegue che:

$$(7.1.4) \quad \Delta a_n = a_0 \frac{[2x_i-1][x_n-x_i]}{x_i[1-x_i]}$$

e dove Δa_n rappresenta la modifica del parametro di controllo da applicare ad ogni iterazione affinché la traiettoria del sistema rimanga sull'orbita prescelta. Senza questo *fine tuning*, infatti, la traiettoria, a causa della sensibilità alle condizioni iniziali, divergerebbe dall'orbita esponenzialmente. Variazione, inoltre, applicabile solo quando la traiettoria x_n è in prossimità dell'orbita di periodo m , cioè quando $|x_n-x_i| < 1$. Il che implica che per ottenere questa modifica la variazione Δa sarà ogni volta molto piccola.

Se è possibile sfruttare i vantaggi di una piccola modifica del parametro di controllo, tuttavia il metodo OGY¹⁰³ ha comunque delle limitazioni. Prima fra tutte la necessità di conoscere l'orbita su cui stabilizzare la dinamica caotica, cui si aggiunge che tale orbita deve essere di basso periodo. Inoltre per applicare il controllo è necessario aspettare che la traiettoria del sistema si trovi nelle vicinanze dell'orbita periodica target. Per superare queste limitazione sono state elaborate varianti dell'originale metodo OGY¹⁰⁴.

È possibile, ad esempio, ridurre la dimensionalità dell'attrattore che, se monodimensionale,¹⁰⁵ può essere approssimato attraverso una singola funzione¹⁰⁶ della

¹⁰³ Per maggiori dettagli sul metodo OGY si veda CHEN G. (2001); CHANFREAU P. - LYYJYNNEN H. (1999); DITTO W.L. *et al.* (1995); DITTO W.L - SHOWALTER K. (1997); DUBÉ L.J. - DESPRÉS P. (2000); SHINBROT T. *et al.* (1993); OGORZALEK M. (1994); GREBOGI C. - LAI Y.C. (1997); BAYLY P.V. - VIRGIN L.N. (1994); BOCCALETTI S. *et al.* (2000).

¹⁰⁴ Si veda OTANI M. - JONES A. J. (1997); HÜBINGER B. *et al.* (1994) per l'applicazione dell'OGY con orbite di elevato periodo; DRESSLER U. - NITSCHKE G. (1992); SO P. - OTT E. (1995), per l'applicazione del controllo del caos attraverso *time-delay coordinates*, HÜBINGER B. *et al.* (1994); RITZ T. *et al.* (1997) per il controllo di orbite periodiche instabili caratterizzate da elevata instabilità.

¹⁰⁵ «In the surface of section, a continuous-time-periodic orbit appears as discrete time orbit cycling through a finite set points. [...] To embed a small neighbourhood of a point from x into X , we typically only require as many dimensions as there are coordinates of the point. Thus for our purpose, $M=D-1$ is generally sufficient. [...] Hence for the case considered ($D=3$), our surface of section is two dimensional», OTT E. *et al.* (1990, pp. 292-293).

¹⁰⁶ v. KANTZ H. - SCHREIBER T. (1997).

quale non è necessario individuare *manifold* stabili ed instabili. Un altro esempio è l'*Occasional Proportional Feedback* (OPF) proposto da Hunt (1991) e da Roy *et al.* (1992). Con questo metodo non è necessario individuare orbite periodiche e la perturbazione del parametro è abbastanza ampia. Il controllo è attivato ogni volta che la traiettoria entra nella prescritta regione. Non richiede conoscenza delle equazioni del sistema, tuttavia, essendo un metodo *trial and error* l'orbita risultante non può essere scelta in anticipo¹⁰⁷. A queste si aggiungono le varianti proposte da Otani e Jones (1997); Hübinger *et al.*, (1994) per l'applicazione dell'OGY con orbite di elevato periodo; Dressler e Nitsche (1992); So e Ott (1995) per l'applicazione del controllo del caos attraverso *time-delay coordinates*; Hübinger *et al.* (1994) e Ritz *et al.* (1997) per il controllo di orbite periodiche instabili caratterizzate da elevata instabilità.

8. - Il metodo di Pyragas o *delayed feedback control*

Le tecniche di controllo del caos finora descritte possono essere efficacemente applicate ai sistemi discreti. Il *Delay Feedback Control* (DFC) proposto da Pyragas (1992), rappresenta, invece, un esempio di controllo caotico applicato in modo particolare ai sistemi continui. Diversi autori hanno esaminato la possibilità di eliminare il caos utilizzando in vario modo variazioni continue periodiche o aperiodiche, ma nessuna di queste ha considerato che la perturbazione poteva essere ottenuta in forma di *feed-back*¹⁰⁸.

L'idea di base del DFC è, infatti, che le orbite periodiche instabili (UPO) possono essere stabilizzate applicando una variazione *time-dependent* nella forma di *feed-back* ai parametri di controllo del sistema. Nel *Delay Feedback Control* (DFC) l'idea base è quella di usare un *feed-back* ritardato. Più precisamente il *feed-back* dipende dalle differenze tra lo stato attuale e lo stato ritardato, dove il ritardo scelto è uguale al periodo dell'UPO da stabilizzare¹⁰⁹.

Consideriamo un'equazione differenziale n -dimensionale:

$$(8.1) \quad \dot{x}(t) = f(x(t), t) \quad (x \in R^n)$$

¹⁰⁷ v. KANTZ H. - SCHREIBER T. (1997).

¹⁰⁸ v. PYRAGAS K. (1992).

¹⁰⁹ v. NAKAJIMA H. (1997).

Per stabilizzare un'UPO in (8.1) col DFC si aggiunge il termine di *feedback* ritardato per cui abbiamo:

$$(8.2) \quad \dot{x}(t) = f(x(t), t) + K(t - T) - x(t)$$

dove K è una matrice $n \times n$ e T (*delay-time*) coincide con il periodo dell'UPO obiettivo. Ovviamente, nella (8.1), quando una traiettoria converge ad un'orbita periodica con periodo T , il termine di *feed-back* svanisce. Ciò implica che la traiettoria si sta muovendo sull'UPO obiettivo¹¹⁰.

Il vantaggio di questo metodo è che esso non richiede pregiudizialmente la totale informazione sull'UPO obiettivo, ma piuttosto usa un ritardo costante nel *feed-back controller* che dipende dal valore assunto dallo stesso¹¹¹. Inoltre, se con il metodo OGY la variazione del parametro di controllo può essere effettuata nel momento in cui il sistema è nelle prossimità del punto fisso, con il DFC non è necessario attendere che il sistema sia nelle vicinanze dell'orbita periodica prescelta da stabilizzare. Il vantaggio rispetto all'OGY è che la variazione può essere realizzata in qualsiasi momento. Il metodo non è, tuttavia, scevro da limitazioni. Per alcuni autori¹¹² la principale consiste nell'applicazione di questo metodo ai sistemi discreti. Il DFC non può stabilizzare l'orbita periodica instabile di un sistema discreto se la matrice Giacobiana del sistema linearizzato nell'intorno del punto fisso instabile presenta un numero dispari di autovalori reali maggiori di Nakajima (1997) dimostra, però, che la stessa limitazione esiste anche per i sistemi caotici continui.

9. - Il metodo del *Targeting*

La condizione essenziale per ottenere la stabilizzazione dell'orbita è che il sistema sia sufficientemente vicino al comportamento desiderato. La garanzia che ciò accada è data dal fatto che la dinamica sull'attrattore è ergodica. Tuttavia, se il sistema è lontano dallo stato desiderato prima che giunga in esso può essere necessario un lasso di tempo troppo lungo. Può accadere, cioè, che le prossimità di un dato punto dell'attrattore non siano visitate di frequente, per cui la dinamica

¹¹⁰ V. NAKAJIMA H. (1997).

¹¹¹ V. TIAN Y. - CHEN C. (2001).

¹¹² V. BIELAWSKI S. *et AL.* (1993); SCHUSTER H.G. *et AL.* (1997); USHIO T. (1996).

del sistema impiegherà molto tempo per avvicinarsi al punto desiderato. Ciò si tradurrà in un lungo tempo di attesa per applicare il controllo¹¹³.

Il metodo del *targeting*¹¹⁴ riduce notevolmente questo tempo di attesa. L'obiettivo è scegliere il *target*, cioè il punto di interesse sull'orbita che si vuole stabilizzare e spingere la traiettoria nelle prossimità di uno dei punti dell'orbita periodica stabile; dopodiché utilizzare l'algoritmo del controllo per mantenere l'orbita vicino al punto per il tempo desiderato.

Anche per il *targeting* non è strettamente necessario conoscere la *law of motion* potendo operare attraverso l'analisi dei dati¹¹⁵. Shinbrot *et al.* (1990) dimostrano, infatti, che il *targeting* può essere raggiunto anche quando il sistema è non perfettamente modellato. Ciò di cui si ha bisogno è un sufficiente ammontare di dati¹¹⁶ per stimare la legge di *motion* e conseguentemente le orbite obiettivo.

La differenza fondamentale tra la procedura di stabilizzazione e quella di *targeting* consiste nell'informazione necessaria per raggiungere il controllo¹¹⁷. Mentre il *targeting* cerca di raggiungere un controllo globale, la stabilizzazione fa uso di un controllo locale, cioè attiva la variazione del parametro quando la dinamica del sistema si avvicina all'orbita obiettivo. Quindi, mentre il controllo locale richiede solo un limitato ammontare di informazione sulla legge che regola il sistema, cioè richiede informazioni sulla posizione dello *steady state* (*saddle point*) e la derivata prima allo *steady state*¹¹⁸, costringendo, però, in alcuni casi a tempi di attesa piuttosto lunghi, il metodo di controllo globale riduce questi tempi ma implica maggiori informazioni sulla dinamica sistema.

Consideriamo la mappa logistica già utilizzata nell'esempio dell'OGY (7.1.1). Partiamo da una condizione iniziale¹¹⁹ di $x_0=0,8$ e $a_0=4$, il che significa che siamo in piena regione caotica. La mappa presenterà sia caos ergodico che topologico. Ricordiamo che quando $3 < a < 1 + \sqrt{6}$, è il caos del nostro esempio, la mappa logistica presenta un punto fisso $x_2=1-1/a$ instabile e un'orbita stabile di periodo 2 come attrattore¹²⁰.

¹¹³ « [...] In general, one expects that the average waiting time before a typical orbit approaches a given saddle periodic point is proportional to the dimension of the attractor», KOSTELIC E. J. - BARRETO E. (1997, p. 160).

¹¹⁴ v. SHINBROT T. *et al.* (1990).

¹¹⁵ v. KOPEL M. (1997).

¹¹⁶ v. KAAS L. (1998).

¹¹⁷ v. KOPEL M. (1997).

¹¹⁸ v. KAAS L. (1998).

¹¹⁹ v. KOPEL M. (1997).

¹²⁰ v. BOCCALETTI S. *et al.* (2000).

Nel nostro esempio il valore del punto fisso sarà: $x_f = 1 - 1/a = 0,75$. Supponiamo che questo sia il nostro obiettivo. Vogliamo, cioè, che il sistema raggiunga un intorno dell'orbita stabile di periodo 2, cioè il punto fisso $x_2 = 0,75$ attraverso una piccola variazione del parametro nell'intervallo $[3,9; 4]$. Partendo dal punto iniziale $x_0 = 0,8$ e variando a nell'intervallo $[3,9; 4]$. dopo la prima iterazione otterremo il seguente intervallo di condizioni iniziali:

$$x_1 = f_{[3,9; 4,0]}(0,8) \in (0,624; 0,64)$$

Consideriamo la nostra ipotesi di partenza con $a_0 = 4$ iterando avremo:

$$x_2 \in f_4(0,62; 0,64) = (0,921; 0,938)$$

$$x_3 \in f_4(0,921; 0,938) = (0,291; 0,232)$$

$$x_4 \in f_4(0,291; 0,232) = (0,825; 0,712)$$

In questa ultima iterazione abbiamo individuato un valore di $x_{n+1} \cong 0,75$, ossia 0,712. L'intervallo di x_1 è ampliato fino a contenere l'obiettivo dopo tre iterazioni. Quindi, individuato l'intervallo di interesse e poiché f_a è continua in a , otterremo il valore incognito del parametro che è $a^* \cong 3,93381$ come radice¹²¹ della seguente equazione¹²²:

$$(9.1) \quad f_4 \circ f_4 \circ f_4 \circ f_a(0,8) = 0,75$$

Con questo valore del parametro di controllo l'obiettivo di dirigere l'orbita verso il punto fisso instabile e quindi sull'orbita stabile di periodo 2 è conseguito

¹²¹ È chiaro che l'equazione ci darà 8 soluzioni che vengono scartate in quanto per l'ipotesi di partenza noi vogliamo ottenere una variazione piccola nell'intervallo $[3,9; 4,0]$. Quindi mentre le soluzioni $a = 4,6875; 5,8313; 6,14352$ non possono essere prese in considerazione in quanto l'intervallo di definizione del parametro è $[0; 4]$, le soluzioni $a = 0,106482; 0,418671; 1,5625$ implicano però una notevole variazione del parametro.

¹²² Trattasi di funzione di funzione delle singole iterazioni.

dopo 4 periodi intervenendo nel primo periodo. Questa modifica del parametro di controllo ha determinato una variazione consistente della dinamica del sistema: da una regione caotica ora il sistema si trova in uno *steady state*.

10. - Applicazioni del controllo del caos in economia

Le applicazioni della teoria del controllo del caos in economia non sono numerose. Dopo alcune riflessioni sull'argomento agli inizi degli anni '90, l'interesse degli economisti per questa problematica si è rafforzato solo di recente.

I primi esempi di utilizzo della teoria del controllo del caos da parte di un economista sono rinvenibili nei lavori di Gabisch (1984); Salzano (1993). Quest'ultimo in un modello di ciclo economico direttamente derivato da Samuelson (1939) e con le modifiche proposte da Gabisch (1984) evidenzia come controllare il caos in un'ipotesi di disavanzo di bilancio¹²³.

Altre applicazioni sono state proposte da fisici. Il modello proposto da Holyst (1996) in cui è applicato un OGY, è un'estensione non-lineare del modello di Richardson descritto da Behrens (1992). Le variabili che si comportano caoticamente sono le quote di mercato delle due imprese considerate nel modello. Il *decision-maker* di una delle imprese influenza l'economia controllando lo *stock* dei suoi investimenti sul mercato, tenendo conto delle vendite correnti di ambedue le imprese. Uno dei principali risultati è che una piccola variazione percentuale del corrispondente parametro di controllo, la funzione degli investimenti, permette di stabilizzare l'andamento del sistema. Il risultato di quest'intervento è l'aumento delle vendite per entrambe le imprese, sebbene il controllo sia stato implementato solo da una di esse. Sempre muovendo dallo stesso modello Holyst *et al.* (1997) hanno dimostrato che un controllo dell'andamento caotico delle quote di mercato, implementato da entrambe le imprese, aumenta la caoticità del sistema. In Holyst (2001) è analizzata l'efficacia del metodo DFC. Sempre nello stesso modello, il controllo è dimostrato potersi ottenere aggiungendo il ritardo alla variabile di stato (le vendite delle due imprese), attraverso dei cambiamenti dei parametri di controllo.

¹²³ «L'ipotesi di disavanzo di bilancio conferma tale risultato [aumento del caos economico]. Qui si evidenzia una riduzione del caos endogeno del sistema economico. Infatti una spesa pubblica costante in disavanzo aggiunge una componente costante alla domanda che riduce o stabilizza le oscillazioni», SALZANO M. (1993, p. 150).

Solo di recente il controllo del caos ha attirato l'attenzione degli economisti. Bala *et al.* (1998) partendo da un modello di *tâtonnement* economico con due agenti e due beni, in cui è stata dimostrata l'esistenza di comportamenti caotici¹²⁴, il controllo del caos è ottenuto attraverso l'impiego del metodo del *targeting*. Risultato del controllo è una situazione di *market clearing*.

Il metodo del *targeting* è anche applicato da Kopel (1997) in un modello di dinamica del mercato con agenti caratterizzati da razionalità limitata. Si dimostra che la comparsa di dinamiche complesse conduce ad *inferior results*. Un'attività di controllo del sistema, invece, si può tradurre in un miglioramento della *performance* del sistema, valutata in termini di profitti aggregati, ricavi di vendite aggregate e vendite medie. Kaas (1998) applica il controllo del caos sempre attraverso il *targeting* ad un modello macroeconomico. Dopo aver individuato che la miglior *performance* del sistema è ottenuta con una spesa pubblica piuttosto che con una politica di tassazione, l'autore ottiene la stabilizzazione dell'instabile equilibrio walrasiano.

11. - Conclusioni

Seguendo l'approccio diffuso in letteratura nella teoria dei controlli all'economia si possono individuare, di fatto, tre fasi¹²⁵. La prima ha riguardato essenzialmente la definizione dei modelli macroeconomici in un contesto deterministico per analizzare politiche di stabilizzazione; nella seconda, sulla base delle varie inconsistenze nella determinazione di politiche ottimali, l'attenzione si è rivolta a modelli di controllo che, nello stabilire il livello degli interventi da compiere, tenessero conto dell'incertezza. Qui, il contesto di analisi non è più quello deterministico, ma stocastico. Infine, nella terza fase si comincia ad interessarsi dei vari aspetti del controllo adattivo in cui la stocasticità viene rilevata nei suoi diversi aspetti (incertezza dei parametri, del sistema, errori di misurazione). A queste tre fasi abbiamo aggiunto quella che, ad opinione di chi scrive, si può prospettare come la quarta fase in cui si aprono nuove prospettive per il controllo dell'economia. Prospettive fornite dall'applicazione della teoria del caos all'economia.

Finora i *policy-maker*, intervenendo sul sistema economico, hanno ritenuto di poter operare su un ben definito sistema: concettualmente isolato e che nel corso

¹²⁴ v. SAARI D.G. (1985, 1991); DAY R.H. - PIANIGIANI G. (1991); BALA V. - MAJUMDAR M. (1992).

¹²⁵ v. SENGUPTA J.K. (1997).

del tempo non subisce cambiamenti radicali. Inoltre, il *policy-maker* assume che un numero sufficiente di dati permettono di prevedere, o quantomeno, di mostrare un *trend* dell'andamento futuro del sistema. Quindi, un intervento di politica economica produrrà un previsto e ben definito risultato e, nel caso in cui il sistema manifestasse un comportamento diverso da quello previsto dal modello, attraverso il controllo è possibile smorzare o eliminare l'oscillazione indesiderata. Il modello di riferimento è un modello lineare o linearizzabile.

Una delle principali critiche mosse a questi modelli è quella di essere fuorvianti per una reale comprensione del fenomeno economico e, soprattutto, di promuovere interventi di politica economica fundamentalmente inadeguati¹²⁶. L'ipotesi di linearità e di costanza dei parametri garantisce la gestibilità, in termini di calcolo, e la soluzione del sistema; tuttavia, la gestibilità e il raggiungimento di una soluzione unica e finita ha avuto un costo in termini d'interpretatività del sistema. Se da un lato, la dimostrazione della presenza di caos nelle serie storiche economiche e finanziarie (Faggini, 2011), è stata piuttosto debole, causata più che altro dalla mancanza di strumenti adeguati a confrontarsi, in particolar modo, con la loro brevità; dall'altro la letteratura teorica economica ampio spazio ha dato alla dimostrazione di come modelli tradizionali possono manifestare comportamenti *inusuali*, introducendo unicamente elementi di non-linearità (Faggini, 2009).

Nella realtà non è il comportamento caotico l'eccezione, bensì il comportamento lineare; non la molteplicità di equilibri quanto il singolo equilibrio. Nella matematica del caos non ci sono ipotesi né in contraddizione né speciali rispetto alla teoria economica¹²⁷. La dimostrazione di ciò è che i modelli economici in cui è stato sottolineato aversi caos sono modelli non solo ben noti agli economisti, ma costruiti su ipotesi *standard*. Diffusa è ormai la consapevolezza che il fenomeno economico è meglio approssimato da un modello non-lineare che manifesta comportamenti caotici, soluzioni multiple, piuttosto che da un modello lineare con soluzione unica e stabile e in cui dinamiche diverse dallo *steady state* devono essere giustificate con opportune specificazioni.

L'abbandono dell'interpretazione lineare dell'economia ha permesso l'indebolimento di molti dei presupposti su cui erano stati costruiti i modelli economici, e di conseguenza ha modificato le modalità con cui fare interventi di politica economica. Da questo punto di vista, un rilevante contributo sembra essere dato dalle tecniche di controllo del caos. Interventi di *policy* possono produrre *perfor-*

¹²⁶ v. BULLARD J.B. - BUTLER A. (1993); WOODFORD M. (1990).

¹²⁷ v. BULLARD J.B. - BUTLER A. (1993).

mance migliori sia rispetto ad un atteggiamento di *laissez faire* che dà luogo a dinamiche complesse Pareto inferiori¹²⁸, sia rispetto a politiche che inducano il sistema ad un inefficiente *steady state*.

Nelle applicazioni economiche sul controllo del caos, sia nelle procedure di stabilizzazione¹²⁹ che di *targeting*¹³⁰ o di uso combinato¹³¹, si è evidenziata la possibilità di modificare l'andamento del sistema nel senso di conseguimento della miglior *performance*, scegliendo tra la varietà disponibile in quanto il sistema è caotico. Come già evidenziato, in Holyst (1996) è stato dimostrato che la stabilizzazione dell'andamento caotico delle vendite di una delle imprese conduce ad un aumento delle vendite di entrambe le imprese presenti sul mercato. Ciò a sostegno di quanto sostenuto da Kopel (1997) che l'introduzione di una *chaotic policy* si ritiene possa condurre ad una situazione migliore per il sistema¹³², che non necessariamente deve essere interpretata come raggiungimento di un punto di equilibrio. Infatti, Huang (1995), mostra che in un modello *cobweb* instabile, ponendo un limite superiore non troppo ampio al tasso di crescita, un'impresa in concorrenza può ottenere, in una condizione di disequilibrio continuo, un profitto più ampio di quello che conseguirebbe in una condizione di equilibrio stabile.

L'obiettivo, dunque, deve essere individuare degli interventi di politica economica che permettono di conseguire una dinamica regolare ed efficiente, pur mantenendo la struttura qualificante del sistema. Ciò è possibile solo utilizzando il caos come teoria interpretativa dei fenomeni economici e le sue tecniche di controllo¹³³. Anche perché sistemi che manifestano comportamenti estremamente irregolari sono più facilmente controllabili di quelli in cui le irregolarità dipendono unicamente da *shock* stocastici, totalmente esterni al sistema. Inoltre, sfruttare la sensibilità alle condizioni iniziali per muoversi da un'orbita all'altra dell'attrattore, significa non solo poter scegliere tra diversi andamenti propri del sistema, ma anche che l'impiego dello strumento di controllo, in termini di risorse per conseguire un determinato obiettivo, sarà minore se paragonato a quello che si dovrebbe utilizzare, a parità di obiettivo, dovendo ricorrere a tecniche di controllo

¹²⁸ v. DAY R.H. (1999).

¹²⁹ v. HOLYST J.A. (1996).

¹³⁰ v. KOPEL M. (1997).

¹³¹ v. KAAS L. (1998).

¹³² « [...] stabilization of the system to the equilibrium (which is in their case a saddle point) leads to inferior results in terms of means sales than under the chaotic policy», KOPEL M. (1997, pp. 286).

¹³³ v. KASS L. (1998).

tradizionali. In sistemi non caotici l'effetto di un *input* sull'*output* è proporzionale a questo, mentre nei sistemi caotici la presenza della sensibilità alle condizioni iniziali rende questa relazione esponenziale: piccoli interventi grandi risultati. Inoltre, una limitazione di intervento del *decison-maker* dovuta a fattori diversi (vincolo di bilancio, accordi internazionali, concorrenza tra i settori e tra gli stati), in un contesto di analisi economica tradizionale, si tradurrebbe in una limitata possibilità di modificare il parametro di controllo del sistema. Se il sistema è caotico la limitazione di risorse da poter utilizzare non riduce la capacità di intervento del *policy-maker*, il quale è consapevole di poter comunque raggiungere l'obiettivo prefissatosi. Inoltre, solo nell'ambito dell'approccio lineare si deve condizionare la controllabilità del modello alla sua prevedibilità. Nel caso della caoticità, l'approccio tradizionale della teoria dell'informazione può essere considerato meno cogente, mentre, viceversa, il costo di acquisizione delle informazioni necessarie al controllo risulta più contenuto. Se uno dei motivi per cui i comportamenti caotici non erano condizioni desiderabili era la loro non prevedibilità di lungo periodo; ora attraverso i metodi di controllo, soprattutto quelli volti alla soppressione del caos, si amplia notevolmente anche la possibilità di una loro prevedibilità. Stabilizzare il sistema su un'orbita ben precisa significa poterne prevedere l'andamento, e quindi migliorare la capacità predittiva di questi modelli. I vantaggi, quindi, dell'approccio caotico si manifestano oltre che sul piano della correttezza metodologica anche sul piano dell'efficienza.

L'approccio caotico, dunque, non significa un totale abbandono o rimpiazzo dell'analisi tradizionale, ma semplicemente la comprensione¹³⁴, in virtù del connubio tra stocasticità e teoria strutturale, di alcuni fenomeni altrimenti non considerati. La teoria del caos può essere un'alternativa alla teoria tradizionale, rappresentando la riconciliazione tra la teoria e realtà economica e, in particolare, tra determinismo e casualità permettendo di capire "*the deterministic origin for stochastic processes*"¹³⁵.

Una precisazione, tuttavia, sembra doverosa. Nell'utilizzo delle tecniche di controllo del caos non bisogna perdere di vista le particolarità del sistema economico. La scelta di intervenire in uno specifico contesto, in virtù delle forti interrelazioni esistenti tra gli agenti, comporterà modifiche del contesto come effetto

¹³⁴ «The source of these new capabilities is that all of the information contained in the structure of the unknown dynamical system in state space can be shown to be embedded in the geometry of the chaotic attractor fractal, which asymptotically produces the stochastic behaviour of the solution paths in phase space», BARNETT W.A. - CHOI S.S. (1989, p. 143).

¹³⁵ v. BARNETT W.A. - CHOI S.S. (1989, p. 142).

diretto dell'intervento ma anche modifiche dovute alla reazione dell'intervento dei diversi agenti, come effetto indiretto. Non così nel campo della fisica, dove l'oggetto di studio non sono soggetti pensanti e soprattutto dove gli esperimenti possono essere riprodotti in laboratorio. Quindi, anche per la teoria del caos non possiamo operare estendendone l'applicazione all'economia senza tener conto della peculiarità di questa rispetto alla fisica dove, invece, la teoria del caos nasce, trovando ampio consenso e applicazione.

Indubbiamente, grosse difficoltà di preparazione dello *staff* degli economisti che devono adoperarli si frappongono alla costruzione di modelli caotici. Ciò è dovuto fondamentalmente al cambiamento di filosofia e d'impostazione che l'analista deve operare, più che a motivi tecnici. Considerati, però, i molti benefici che si possono ottenere, probabilmente varrebbe la pena un approfondimento di questa linea di ricerca, anziché continuare a profondere energie nella costruzione di modelli formalmente perfetti, in termini di ipotesi e di derivazioni, e, soprattutto, accademicamente riconosciuti, ma assolutamente inutili all'uomo della strada, il quale si chiede quotidianamente: «perché siamo stati in grado di andare sulla Luna ed esplorare Marte, ma non riusciamo a risolvere problemi come la disoccupazione?».

BIBLIOGRAFIA

- ABARBANEL H.D.I., *Analysis of Observed Chaotic Data*, Springer-Verlag, 1996.
- ANDRIEVSKII B.R. - FRADKOV A.L., «Control of Chaos: Methods and Applications. I. Methods», *Automation and Remote Control*, vol. 64, no. 5, 2003, pages 673-713.
- AOKI M., *Optimization of Stochastic Systems*, New York, Academic Press, 1967.
- .-, «Stochastic Control Theory in Economics: Applications and New Problems», *IFAC Symposium on Stochastic Control*, Budapest, 1974.
- .-, *Optimal Control and System Theory in Dynamics Economics*, North-Holland, NY, 1978.
- ARROW K.J., «Application of Control Theory to Economic Growth», in *Lectures in Applied Mathematics, Mathematics of the Decision Sciences*, Part 2, vol. 12, Providence, RI, American Mathematical Society, 1968.
- ATHANS M. - CHOW G.C., «Introduction to Stochastic Control and Economic Systems», *Annals of Economic and Social Measurement*, vol. 1, 1972, pages 375-384.
- ATHANS M. - KENDRICK D.A., «Control Theory and Economics: A Survey, Forecast, and Speculations», *IEEE Trans. Autom. Control*, vol.19, issue 5, 1974, pages 518-523.
- BALA V. - MAJUMDAR M., «Chaotic Tâtonnement», *Economic Theory*, vol. 2, 1992, pages 437-445.
- BALA V. - MAJUMDAR M. - MITRA T., «A Note on Controlling a Chaotic Tâtonnement», *Journal of Economic Behavior & Organization*, vol. 33, 1998, pages 411-420.
- BARNETT W.A. - CHOI S.S., «A Comparison between the Conventional Econometric Approach to Structural Inference and the Nonparametric Chaotic Attractor Approach», in BARNETT W.A. - GEWEKE J. - SHELL K. (eds.), *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity. Proceedings of the Fourth International Symposium in Economic Theory and Econometrics*, 1989, pages 141-212.
- BAYLY P.V. - VIRGIN L.N., «Practical Considerations in the Control of Chaos», *Physical Review E*, vol. 50, no. 1, 1994, pages 604-607.
- BEHRENS D., «Two- and Three-Dimensional Models on the Arms Race», *Diplomarbeit Institut für Ökonometrie, Operations Research und Systemtheorie, Technische Universität Wien*
- BIELAWSKI S. - DEROZIER D. - GLORIEUX P., «Experimental Characterization of Unstable Periodic Orbits by Controlling Chaos», *Physic Review A*, vol. 47, no. 4, 1993, pages 2493-2495.
- BISHOP S.R. - XU D. - LIAM C.Y. - CHAN E.S., «Applying Chaos Control in Periodic Windows», *Chaos, Solitons & Fractals*, vol. 8, no. 8, 1998, pages 1297-1305.
- BLACKBURN K., «Macroeconomic Policy Evaluation and Optimal Control Theory: A Critical Review of Some Recent Developments», *Journal of Economic Surveys*, vol.1, issue 2, 1987, pages 111-148.

- BOCCALETTI S. - GREBOGI C. - LAI Y.C. - MANCINI H. - MAZA D., «The Control of Chaos: Theory and Applications», *Physics Reports*, vol. 329, pages 103-197, 2000.
- BRAIMAN Y. - GOLDBIRSHCH I., «Taming Chaotic Dynamics with Weak Periodic Perturbations», *Physical Review Letters*, vol. 66, 1991, pages 2545-2548.
- BREEDEN L., «Open-Loop Control of Nonlinear Systems», *Physical Letters A*, vol. 190, 1994, pages 264-272.
- BRYSON A.E. JR. - YU-CHI H., *Applied Optimal Control*, Blaisdell, Waltham, Mass, 1969.
- BULLARD J. - BUTLER A., «Nonlinearity and Chaos in Economic Models: Implications for Policy Decisions», *Economic Journal*, vol. 103, no. 419, 1993.
- CANNON M.D. - CULLUM C.D. - POLAK E., *Theory of Optimal Control and Mathematical Programming*, McGraw-Hill, New York, San Francisco, London, 1970.
- CHANFREAU P. - LYYJYNNEN H., «Viewing the Efficiency of Chaos Control», *Journal of Nonlinear Mathematical Physics*, vol. 6, no. 3, 1999, pages 314-331.
- CHEN G., «Introduction to Chaos Control and Anti-Control», in LEUNG T. P. - QIN H.-S. (eds.), *Advanced Topics in Nonlinear Control Systems*, World Scientific Pub. Co., Singapore, 2001, pages 193-245.
- CHEN G. - DONG X., «From Chaos to Order-Perspectives and Methodologies in Controlling Chaotic Nonlinear Dynamic Systems», *International Journal of Bifurcation Chaos*, vol. 3, 1993, pages 1363-1409.
- CHIARINI B., *Teoria ed applicazioni del controllo ottimo*, Franco Angeli, Milano, 1990.
- CHOW G.C., «Optimal Stochastic Control of Linear Economic Systems», *Journal of Money, Credit, and Banking*, August, 1970, pages 291-302.
- .-, «How Much Could Be Gained by Optimal Stochastic Control Policies», *Annals of Economic and Social Measurement*, vol. 1, October, 1972, pages 391-406.
- .-, «Control Methods for Macroeconomic Policy Analysis», *The American Economic Review*, vol. 66, issue 2, 1976, pages 340-345.
- COLANDER D. - FÖLLMER H. - HAAS A. - GOLDBERG M. - JUSELIOUS K. - KIRMAN A. - LUX T. - SLOTH B., «The Financial Crisis and the Systemic Failure of Academic Economics», *Kiel Working Papers*, no. 1489, Kiel Institute for the World Economy, 2009.
- COOPER J.P. - FISCHER S., «A Method of Stochastic Control of Non-Linear Models and an Application», *Econometrica*, January, 1975, pages 147-162.
- CULLUM J., «Penalty Functions and Nonconvex Continuous Optimal Control Problems, Computing Methods in Optimization Problems», in ZADEH L.A. - NEUSTADT L.W. - VO BALAKRISHNAN A. (eds.), Academic Press, New York, 1969, pages 55-67.
- CURRIE D., «Macroeconomic Design and Control Theory - A Failed Partnership», *Economic Journal*, vol. 93, 1985, pages 285-306.
- DAY R.H., *Complex Economic Dynamics: vol. II. An introduction to Macroeconomic Dynamics*, The MIT Press, Cambridge Massachusetts, 1999.

- DAY R.H. - PIANIGIANI G., «Statistical Dynamics and Economics», *Journal of Economic Behaviour and Organisation*, vol. 16, 1991, pages 37-84.
- DE KORTE R.J. - SCHOUTEN J.C. - VAN DEN BLEEK C.M.V., «Experimental Control of a Chaotic Pendulum with Unknown Dynamics Using Delay Coordinates», *Physical Review E*, vol. 52, 1995, pages 3358-3365.
- DIEBOLD F.X., «The Past, Present, and Future of Macroeconomic Forecasting», *The Journal of Economic Perspectives*, vol. 12, 1998, pages 175-192.
- DITTO W.L. - SHOWALTER K., «Introduction: Control and Synchronization of Chaos», *Chaos*, vol. 7, no. 4, 1997, pages 509-511.
- DITTO W.L. - SPANO M.L. - LINDNER J.F., «Techniques for the Control of Chaos», *Physica D*, vol. 86, 1995, pages 198-211.
- DOBELL A.R., «Some Characteristic Features of Optimal Problems in Economic Theory», *IEEE Trans. Autom. Control*, vol. AC-14, issue 1, 1969, pages 39-46.
- DRESSLER U. - NITSCHKE G., «Controlling Chaos Using Time Delay Coordinates», *Physical Review Letters*, vol. 68, no. 1, 1992, pages 1-4.
- DUBÉ L.J. - DESPRÉS P., «The Control of Dynamical Systems Recovering Order from Chaos», in ITIKAWA Y. (ed.), *The Physics of Electronic and Atomic Collisions*, AIP Conference Proceedings, 2000, pages 551-570.
- DYSON F., *Infinite in All Directions*, New York, Harper and Row Publishers, 1988.
- ERMINI L., «The Limits of System Control Theory for Economic Policy-Making», *Working Papers in Economics*, no. 111, July-October, Department of Economics, The University of Australia, 1988.
- FAGGINI M., «Un approccio di teoria del caos all'analisi delle serie storiche economiche», *Rivista di Politica Economica*, luglio-agosto, 2005.
- .-, «Visual Recurrence Analysis: An Application to Economic Time Series», in SALZANO M. - COLANDER D. (eds.), *Complexity Hints for Economic Policy*, Springer Edition, 2007.
- .-, «Chaos and Chaotic Dynamics in Economics», *Nonlinear Dynamics, Psychology and Life Sciences*, vol. 13, 2009, pages 327-340.
- .-, «Chaotic Time Series Analysis in Economics. Balance and Perspectives», *Working Paper Series*, no. 25, Università di Torino, Department of Economics and Public Finance "G. Prato", 2011.
- FRIEDMAN B.M., «Optimal Economic Stabilization Policy: An Extended Framework», *Journal of Political Economy*, vol. 80, 1972, pages 1002-1022.
- GABISCH G., «Non-Linear Model of Business Cycle Theory», in HAMMER D. - PALLASCHNKE D. (eds.), *Selected Topics in Operational Research and Mathematical Economics, Lecture Notes in Economics and Mathematical Systems*, no. 226, Berlino, Springer Verlag, 1984, pages 205-222.
- GLEICK J., *Caos. La nascita di una nuova scienza*, Biblioteca Scientifica Sansoni, 1997.

- GREBOGI C. - LAI Y.C., «Controlling Chaotic Dynamical Systems», *Systems & Control Letters*, vol. 31, 1997, pages 307-312.
- . -, -. -, «Controlling Chaos», in SCHUSTER H.G. (ed.), *Handbook of Chaos Control*, Wiley, Vch, Verlag, 1999.
- HENDRY D.F., «Econometrics-Alchemy or Science?», *Economica*, vol. 47, issue 188, 1980, pages 387-406.
- HOLBROOK R.S., «Optimal Economic Policy and the Problem of Instrument Instability», *American Economic Review*, Mar., vol. 62, 1972, pages 57-65.
- HOLT C.C., «Linear Decision Rules for Economic Stabilization and Growth», *Quarterly Journal of Economics*, vol. 76, 1962, pages 20-45.
- HOLT C.C. - MODIGLIANI F. - MUTH J.F. - SIMON H.A., *Planning, Production, Inventories, and Work Force*, Englewood Cliffs, NJ, Prentice Hall, 1960.
- HOLYST J.A., «How To Control a Chaotic Economy», *Journal of Evolutionary Economics*, vol. 6, issue 1, 1996, pages 31-42.
- HOLYST J. - HAGEL T. - HAAG G., «Destructive Role of Competition and Noise for Control of Microeconomical Chaos», *Chaos, Solitons and Fractals*, no. 8, 9, 1997, pages 1489-1505.
- HOLYST J.A. - ZEBROWSKA M. - URBANOWICZ K., «Observations of Deterministic Chaos in Financial Time Series by Recurrence Plots, Can One Control Chaotic Economy?», *The European Physical Journal*, B 20, no. 4, 2001, pages 531-535.
- HÜBINGER B. - DOERNER R. - MARTIENSSSEN W., «Controlling Chaos Experimentally in Systems Exhibiting Large Effective Lyapunov Exponents», *Physical Review E*, vol. 50, no. 2, 1994, pages 932-948.
- HÜBLER A. - LÜSCHER E., «Resonant Stimulation and Control of Nonlinear Oscillators», *Naturwissenschaften*, vol. 76, 1989, pages 67-69.
- HUEBLER A.W. - PLAPP B.B., «Nonlinear Resonances and Suppression of Chaos in the Rf-Biased Josephson Junction», *Physical Review Letter*, vol. 65, 1990, pages 2302-2305.
- HUNT B.R., «Stabilizing High Period Orbits in a Chaotic System: The Diode Resonator», *Physical Review Letters*, vol. 67, 1991, pages 1953-1955.
- INTRILIGATOR M., «Applications of Optimal Control Theory in Economics», *Synthese*, vol. 31, issue 2, August, 1975, pages 271.
- JACKSON E.A. - GROSU I., «An Open-Plus-Closed-Loop (OPCL) Control of Complex Systems», *Physica D*, vol. 85, 1995, pages 1-9.
- KAAS L., «Stabilizing Chaos in a Dynamic Macroeconomic Model», *Journal of Economic Behavior and Organisation*, vol. 33, 1998, pages 313-332.
- KANTZ H. - SCHREIBER T., *Non Linear Time Series Analysis*, Cambridge University Press, 1997.
- KENDRICK D.A., «Applications of Control Theory to Macroeconomics», *Annals of Economics and Social Measurement*, vol. 5, issue 2, 1976, pages 171-190.

- KENDRICK D.A., «Control Theory with Applications to Economics», in ARROW K.J. - INTRILIGATOR M.D. (eds.), *Handbook of Mathematical Economics*, vol. I, North-Holland Publishing Company, 1981, pages 111-158.
- .-, *Stochastic Control for Economic Models*, 2nd Edition, The University of Texas, Austin Texas, US, 2002a.
- .-, *Stochastic Control for Economic Models: Past, Present and the Paths Ahead*, paper presented at the Ninth International Conference of the Society of Computational Economics at Aix-en-Provence in June, 2002b.
- KOPEL M., «Improving the Performance of an Economic System: Controlling chaos», *Journal of Evolutionary Economics*, vol. 7, 1997, pages 269-289.
- KOSTELIC E.J. - BARRETO E., «Targeting and Control of Chaos», in JUDD K. - MEES A. - TEO K.L. - VINCENT T.L. (eds.), *Control and Chaos*, Boston, Birkhäuser, 1997, pages 158-169.
- KU R. - ATHANS M., «On Adaptive of Linear Systems Using the Open Loop Feedback Optimal Control», *IEEE Transactions on Automatic Control*, vol. AC-18, no. 5, October, 1973.
- LIMA R. - PETTINI M., «Suppression of Chaos by Resonant Parametric Perturbations», *Physical Review A*, vol. 41, 1990, pages 726-733.
- LUX T. - WESTERHOFF F., «Economics crisis», *Nature Physics*, vol 5, issue 1, 2009, pages 2-3.
- MATEOS DE CABO R. - OLMEDO E. - SANCHO M. - VALDERAS J.M., «From Linearity to Complexity: Towards a New Economics», *Complexity and Philosophy Workshop*, ISCE Norwood, 2002.
- MEDIO A. - LINES M., *Nonlinear Dynamics: A Primer*, Cambridge University Press, 2001.
- MIRAMONTES O. - SOLÉ R.V. - GOODWIN B.C., *Antichaos in Ants: The Excitability Metaphor at Two Hierarchical Levels*, Proceedings of the Second European Conference on Artificial Life, Brussels, Belgium, 1993, pages 790-802.
- MIRUS K.A. - SPROTT J.C., «Controlling Chaos in Low- and High-Dimensional Systems with Periodic Parametric Perturbations», *Physical Review E*, vol. 59, no. 5, 1999, pages 5313-5324.
- NAKAJIMA H., «On Analytical Properties of Delayed Feedback Control of Chaos», *Physics Letters A*, vol. 232, 1997, pages 207-210.
- NORMAN A.L., «On the Relationship between Linear Feedback Control and First Period Certainty Equivalence», *International Economic Review*, vol. 15, issue 1, 1974, pages 209-215.
- OGORZALEK M., «Chaos Control: How to Avoid Chaos or Take Advantage of It», *Journal of the Franklin Institute*, vol. 331B, no. 6, 1994, pages 681-704.
- OTANI M. - JONES A.J., «Guiding Chaotic Orbits», *Research Report*, no. 130, Imperial College of Science Technology and Medicine, London, 1997.
- OTT E. - GREBOGI C. - YORKE J.A., «Controlling Chaos», *Physical Review Letters*, vol. 64, no. 11, 1990, pages 1196-1199.

- PESARAN M.H. - POTTER S., «Introduction to Special Issue», *Journal of Applied Econometrics*, vol. 5, 1992, pages S1-S7.
- PHILLIPS A.W., «Stabilisation Policy in a Closed Economy», *Economic Journal*, vol. 64, 1954, pages 290-323.
- .-, «Stabilisation Policy and the Time-Form of Lagged Responses», *Economic Journal*, vol. 67, 1957, pages 265-277.
- PHILLIPS P.C.B., «Laws and Limits of Econometrics», *The Economic Journal*, vol. 113, issue 486, March, 2003, pages C26-C52.
- PINDYCK R.S., «Optimal Stabilization Policies Via Deterministic Control», *Annals of Economic Systems and Social Measurement*, vol. 5, 1972, pages 239-248.
- .-, «Optimal Policies for Economic Stabilization», *Econometrica*, vol. 41, no. 3, May, 1973, pages 529-560.
- POLAK E., *Computational Methods in Optimization*, Academic, New York, 1971.
- .-, «An Historical Survey of Computational Methods in Optimal Control», *SIAM Review*, vol. 15, issue 2, Part 2, Anniversary Supplement (April), 1973, pages 553-584.
- PRASADA RAO G., «Elements of Control Systems, International Centre for Water and Energy Systems», PO Box 2623, Abu Dhabi, UAE, luglio, <http://www.eolss.com/eolss/47a.htm>, 2003.
- PYRAGAS K., «Continuous Control of Chaos by Self-Controlling Feedback», *Physical Letters A*, no. 170, 1992, pages 421-428.
- RITZ T. - SCHWEINSBERG A.S.Z. - DRESSLER U. - DOERNER R. - HÜBINGER B. - MARTIENSSEN W., «Chaos Control with Adjustable Control Times», *Chaos, Solitons & Fractals*, vol. 8, no. 9, 1997, pages 1559-1576.
- ROY R. - MURPHY T.W. - MAIER T.D. - GILLS Z. - HUNT E.R., «Dynamical Control of a Chaotic Laser: Experimental Stabilization of a Globally Coupled System», *Physical Review Letters*, vol. 68, 1992, pages 1259-1262.
- SAARI D.G., «Erratic Behavior in Economic Models», *Journal of Economic Behavior and Organization*, vol. 16, 1991, pages 3-35.
- .-, «Iterative Price Mechanisms», *Econometrica*, vol. 53, 1985, pages 1117-1132.
- SALZANO M., *Le variabili fiscali nel modello neokeynesiano*, Liguori Editore, 1993.
- SAMUELSON P., «Interaction between the Multiplier Analysis and the Principle of Acceleration», *Review of Economic Statistics*, vol. 21, 1939, pages 75-78.
- SARGAN J.D., «Current Problems in Econometrics: A Personal View», *Econometric Theory*, vol. 19, issue 3, June, 2003, pages 423-428.
- SCHNEIDER F. - KIRCHGÄSSNER G., «Financial and World Economic Crisis: What did Economists Contribute?», *Public Choice*, no. 140, 2009, pages 319-327.
- SCHUSTER H.G. - STEMMLER M.B., «Control of Chaos by Oscillating Feedback», *Physical Review E*, vol. 56, no. 6, 1997, pages 6410-6417.

- SENGUPTA J.K., «Optimal Stabilization Policy with a Quadratic Criterion Function», *Review of Economic Studies*, vol. 37, 1970, pages 127-146.
- .-, *Control Theory Methods in Economics*, Kluwer Academic Publishers, 1997.
- SHINBROT T. - GREBOGI C. - OTT E. - YORKE J.A., «Using Small Perturbations to Control Chaos», *Nature*, vol. 363, 1993, pages 411-417.
- .-, «Using Chaos to Direct Trajectories to Targets», *Physical Review Letters*, vol. 65, no. 26, 1990, pages 3215-3218.
- SHOWALTER K. - PETROV V. - GASPAR V. - MASERE J., «Controlling Chaos in the Belousov-Zhabotinsky Reaction», *Nature*, vol. 361, 1993, pages 240-243.
- SIMON H.A., «Dynamic Programming under Uncertainty with a Quadratic Criterion Function», *Econometrica*, vol. 24, January, 1956, pages 74-81.
- SINGER J. - WANG Y.Z. - BAU H.H., «Controlling a Chaotic System», *Physical Review Letters*, vol. 66, no. 9, 1992.
- SO P. - OTT E., «Controlling Chaos Using Time Delay Coordinates Via Stabilization of Periodic Orbits», *Physical Review E*, vol. 51, no. 4, 1995, pages 2955-2962.
- SO P. - OTT E. - SCHIFF S.J. - KAPLAN D.T. - SAUER T. - GREBOGI C., «Detecting Unstable Periodic Orbits in Chaotic Experimental Data», *Physical Review Letters*, vol. 76, no. 25, pages 4705-4708, 1996.
- SUBBARAM NAIDUS D., «Featured Review of Books on Optimal Control», *International Journal Robust Nonlinear Control*, vol. 10, 2000, pages 1345-1358.
- THEIL H., «A Note on Certainty Equivalence in Dynamic Planning», *Econometrica*, vol. 25, 1957, pages 346-349.
- .-, *Optimal Decision Rules for Government and Industry*, Amsterdam, North Holland, 1964.
- TIAN Y.C. - TADE M.O. - TANG J.Y., «Non-linear Open-Plus-Closed-Loop (NOPCL) Control of Dynamic Systems», *Chaos, Solitons & Fractals*, vol. 11, 2000, pages 1029-1035.
- TIAN Y.P. - CHEN G., «A Separation Principle for Dynamical Delayed Output Feedback Control of Chaos», *Physics Letters A*, vol. 284, 2001, pages 31-42.
- TINBERGEN J., *Economic Policy: Theory and Design*, Amsterdam, The Netherlands, North Holland Publishing Co, 1956.
- .-, *On the Theory of Economic Policy*, Amsterdam North-Holland Publishing Company, 1952.
- TSE E. - ATHANS M., «Adaptive Stochastic Control for a Class of Linear Systems», *IEEE Transaction on Automatic Control*, vol. AC-17, no. 1 February, pages 450-490, 1972.
- .-, «Adaptive Stochastic Control for Linear Systems-I: Solution Methods; II Asymptotic Properties and Simulation Results», *Proc. IEEE Conf. on Decision and Control*, Austin Texas, 1970.
- TURNOVSKY S.J., «Optimal Stabilization Policies for Deterministic and Stochastic Linear Systems», *Review of Economic Studies*, vol. 40, 1973, pages 79-96.

- TUSTIN A., *The Mechanism of Economic Systems*, Heinemann, London and Harvard University Press, Cambridge, Mass, 1953.
- USHIO T., «Limitation of Delayed Feedback Control in Nonlinear Discrete-Time Systems», *IEEE Transactions on Circuits and Systems I*, vol. 43, 1996, pages 815-816.
- VERCELLI A., *Keynes dopo Lucas. I fondamenti della macroeconomia*, La Nuova Italia Scientifica, 1987.
- WALRAS L., *Éléments d'économie politique pure ou théorie de la richesse sociale*, Imprimerie L. Corbaz & Cie., Lausanne, 1874.
- WONHAM W.M., «Optimal Stationary Control of a Linear System with State Dependent Noise», *SIAM Journal Control*, vol. 5, 1967, pages 486-500.
- WOODFORD M., «Equilibrium Models of Endogenous Fluctuations: An Introduction», *NBER Working Paper*, no. 3360, 1990.
- YAGASAKI K. - UOZUMI T., «Controlling Chaos Using Nonlinear Approximations and Delay Coordinate Embedding», *Physics Letters A*, vol. 247, 1998, pages 129-139.
- ZANGWILL W.I., «Nonlinear Programming via Penalty Functions», *Management Sciences*, vol. 13, 1967, pages 344-358.
- .-, *Nonlinear Programming: A Unified Approach*, Prentice-Hall International Series in Management, Englewood Cliffs, NJ, 1969.

Sulla *political economy* della spesa pubblica nell'Italia liberale

Emma Galli*

La Sapienza Università di Roma

Roberto Ricciuti**

Università di Verona e CESifo

This paper empirically analyzes the demand for public goods in Italy over the period 1861-1913 in a political economy perspective and concentrates on fiscal illusion. The analysis suggests that the relationship between public expenditure and deficit is typically a long-run one. There is also a negative short-run relationship between public expenditures and the price of private goods showing complementarity between private and public goods; instead, in the long-run substitutability prevails. The impact of both income distribution and the number of voters on the public expenditures is variable over time.

[JEL Classification: H3; H61; N4].

Keywords: *political economy*; Italia liberale; spesa pubblica; illusione fiscale.

* <emma.galli@uniroma1.it>, Dipartimento di Analisi Economiche e Sociali, La Sapienza Università di Roma;

** <roberto.ricciuti@univr.it>, Dipartimento di Scienze Economiche, Università degli Studi di Verona, CESifo, Munich.

Desideriamo ringraziare Giovanni Federico, Antonio Pedone, Fabio Padovano e due anonimi referee per i loro commenti a precedenti versioni di questo lavoro. Vale la formula liberatoria.

«Contenere la spesa pubblica in così ristretti confini, che potessero bastare ad essa le risorse derivanti da un mite e riguardoso ordinamento tributario, sarebbe stato senza dubbio desiderabile, ma non era possibile nei primi tempi della vita nazionale [...] Da tre principali fonti prendeva alimento il crescere delle pubbliche spese, o traevano origine le difficoltà a limitare talune di esse: le necessità militari, le opere pubbliche, l'ordinamento amministrativo del paese»,

PLEBANO A. (1900, p. 2).

1. - Introduzione

Diversi studi di storia economica hanno analizzato la politica macroeconomica in Italia con particolare attenzione alla politica fiscale e all'evoluzione del debito pubblico (Toniolo e Ganugi, 1992; Marongiu, 1996*a,b*; Zamagni, 1998; Artoni e Biancini, 2003; Ricciuti, 2008). Con riferimento al periodo 1861-1914 la maggior parte dei lavori sulla finanza pubblica sono stati in un certo senso funzionali all'analisi dello sviluppo economico, concentrandosi sul ruolo svolto dallo Stato nel processo di crescita dell'economia italiana o sui rapporti Stato-industria. Le politiche pubbliche sono state esaminate a livello microeconomico, con riferimento agli effetti che la costruzione della rete ferroviaria ed il loro rendimento o l'applicazione delle tariffe protezionistiche hanno avuto sullo sviluppo di specifici settori industriali (Fenoaltea, 2006). Questi studi, tuttavia, non considerano il ruolo che le istituzioni e i processi politici hanno avuto sui comportamenti della classe politica del tempo nella formazione delle decisioni di bilancio e di finanziamento del settore pubblico. Fanno eccezione, in tal senso, i lavori di Ferrero e Santagata (1981) che analizzano il ciclo economico-elettorale nel Regno d'Italia nel periodo 1867-1913; di Fratianni e Spinelli (1982) che interpretano la dinamica della spesa pubblica nel periodo 1861-1979 attraverso tre modelli alternativi, e precisamente di domanda dei beni pubblici, di gruppi di interesse e redistributivo; di Brosio e Marchese (1988) che elaborano e verificano empiricamente modelli di tipo redistributivo, adattandoli alle caratteristiche politico-istituzionali dei sistemi di governo che si sono succeduti in Italia a partire dal 1861; di Fiorino e Ricciuti (2011) che analizzano l'aumento della spesa pubblica nell'Italia liberale mediante un modello di gruppi di interesse.

Il nostro lavoro si inserisce in questo filone e si propone di verificare l'ipotesi che i governi dell'Italia liberale abbiano fatto frequente ricorso ai disavanzi come

modalità di finanziamento della spesa pubblica nonostante in quel periodo la dottrina e la pratica sostenessero il principio del pareggio del bilancio.

In particolare rispetto al contributo di Brosio e Marchese (1988) la nostra analisi innova in tre direzioni: in primo luogo, include l'illusione fiscale tra le possibili determinanti della spesa; in secondo luogo utilizza una metodologia che permette di considerare gli effetti di retroazione che la spesa pubblica può avere sulle variabili che la dovrebbero spiegare. Infine, utilizza i nuovi dati di contabilità nazionale prodotti da Fenoaltea (2005) che hanno un profilo diverso rispetto a quelli utilizzati in precedenza e modificano pertanto l'andamento di alcune delle variabili rilevanti.

Il lavoro è organizzato come segue: nel paragrafo 2 si discute la dinamica della spesa pubblica nell'Italia liberale includendo, accanto alle determinanti tradizionali, l'ipotesi dell'illusione fiscale nella formulazione di Buchanan e Wagner (1977), ovvero l'ipotesi che i contribuenti preferiscono il finanziamento della spesa pubblica con il debito piuttosto che con la tassazione corrente. Nel paragrafo 3 si specifica il modello empirico e nel paragrafo 4 si discutono i risultati dell'analisi. Seguono alcune considerazioni conclusive.

2. - Spesa pubblica e illusione fiscale nell'Italia liberale

La letteratura economica identifica una serie di fattori socioeconomici (il reddito, la popolazione, i prezzi dei beni pubblici e privati, il livello di tollerabilità dell'imposizione fiscale) e politico-istituzionali (sistemi politici, leggi elettorali, estensione del suffragio) da cui dipende la domanda di spesa pubblica (Brosio e Marchese, 1986). Accanto a questi fattori può assumere un ruolo rilevante l'ipotesi che la classe politica utilizzi il *deficit spending* come canale di finanziamento della spesa a causa di un fenomeno di illusione fiscale¹.

¹ La teoria dell'illusione fiscale è stata formulata originariamente da PUVIANI A. (1897) e successivamente ripresa da BUCHANAN J.M. (1960). Secondo Puviani il confronto tra i costi dell'attività pubblica (imposte) ed i suoi benefici (servizi pubblici) non è sempre percepito dai cittadini in modo adeguato. Per una rivisitazione della teoria nelle sue varie dimensioni si rimanda a GOETZ C. e DA EMPOLI D. (1976) e OATES W.E. (1988). L'illusione fiscale è in contrasto con il teorema dell'equivalenza ricardiana tra imposte e debito pubblico proposto da BARRO R. (1974), per quanto Ricardo stesso mostrasse un certo scetticismo in proposito; infatti scriveva: «In point of economy, there is no real difference in either of three modes: for twenty millions in one payment, one million per annum for ever, or 1,200,000 for 45 years, are precisely the same value; but people who pay taxes never so estimate them, and therefore do not manage their private affairs accordingly. We are too apt to think that war is burdensome only in proportion to what we are at the moment called to pay for it in taxes, reflecting on the probable duration of such taxes», (RICARDO D., 1951-1820, pp. 186-187).

Una specificazione dell'approccio di illusione fiscale è rappresentato dalla scelta della combinazione debito/imposte e dai suoi effetti sulla spesa pubblica, sviluppata da Buchanan e Wagner (1977). La possibilità che il comportamento individuale non sia sempre caratterizzato da informazione e razionalità perfetta è alla base di tale fenomeno. Generalmente l'illusione fiscale si verifica in presenza delle seguenti condizioni:

- 1) i votanti-contribuenti non sono consapevoli dei futuri obblighi fiscali causati dai *deficit* correnti;
- 2) scontano gli impegni futuri ad un tasso di interesse più alto del tasso di interesse sui titoli del debito pubblico;
- 3) hanno vite finite e valutano l'onere futuro durante la vita più dell'onere imposto alle generazioni future. Secondo Buchanan e Wagner (1977) la sostituzione del finanziamento della spesa pubblica attraverso imposte correnti con il ricorso al debito pubblico riduce i prezzi dei beni e servizi così come sono percepiti dai votanti-contribuenti che non anticipano completamente gli oneri fiscali futuri impliciti nel ricorso al debito. Una volta che queste, o alcune di queste condizioni, siano soddisfatte, si determinerà un'elasticità negativa della domanda per i beni e servizi pubblici rispetto al prezzo percepito, per cui la spesa pubblica tenderà ad aumentare e il ricorso al *deficit* a celarne o differirne il costo. In altri termini, il disavanzo riduce il prezzo effettivo pagato dai contribuenti per i servizi forniti dal settore pubblico e quindi ne aumenta la domanda. Ne consegue che la classe politica tende ad evitare il ricorso a un aumento della tassazione per finanziare eventuali incrementi di spesa.

Sottoporremo a verifica empirica questa proposizione con riferimento al comportamento della classe politica del Regno d'Italia nel primo cinquantennio postunitario caratterizzato da un governo monarchico rappresentativo di tipo parlamentare, in cui il Parlamento e il governo erano sostanzialmente preposti alle decisioni di politica fiscale. In particolare, lo Statuto Albertino del 1848 prescriveva che nessun tributo potesse essere imposto o riscosso senza l'approvazione delle Camere e la sanzione regia (art. 30), sanzione che, peraltro, si rivelò il più delle volte di carattere formale. Inoltre, l'art. 10 imponeva la presentazione in via preliminare alla Camera elettiva di ogni legge di imposizione dei tributi e di approvazione dei bilanci e dei conti dello Stato. Senza considerare che la prassi emersa nell'attività parlamentare aveva visto la prevalenza pressoché assoluta della Camera dei Deputati e del governo nelle decisioni in materia fiscale in tutto il periodo di interesse².

² Per una storia politica-parlamentare italiana nel primo cinquantennio postunitario si vedano MARANINI G. (1967) e GHISALBERTI G. (1986).

In un sistema in cui il suffragio ristretto e la base uninominale del collegio rendono più diretto il collegamento tra eletti ed elettori, il grado di controllo degli elettori sull'operato del governo è elevato. Si consideri in proposito che nel Regno d'Italia il numero dei cittadini aventi il diritto di voto passò dal 2% della popolazione residente nel 1861 all'intera popolazione adulta maschile (sopra i trenta anni o che avesse assolto il servizio militare) nel 1912. La percentuale effettiva dei votanti si dimostrò generalmente bassa, soprattutto a causa della mancata partecipazione degli strati meno abbienti tra gli aventi diritto.

L'estensione del corpo elettorale avvenne in due passaggi principali. La percentuale degli aventi diritto al voto passò dal 2,2% nel 1880 al 6,9% nelle elezioni del 1892 e successivamente dall'8,3% del 1909 al 23,2% corrispondente al suffragio universale nel 1913. I requisiti di capacità e di censo richiesti per gli elettori e il collegamento diretto tra questi e l'eletto, facilitato sia dalle limitate dimensioni della circoscrizione elettorale che dalla base uninominale del collegio, comportavano la selezione di una classe politica alquanto omogenea, costituita dai cosiddetti notabili, cittadini che emergevano all'interno di ogni singolo collegio ponendosi in una posizione di preminenza nell'ambiente locale. Inoltre il sistema elettorale del doppio turno con il ballottaggio tra i due candidati che, nell'ipotesi di mancato conseguimento del *quorum* richiesto per l'elezione a primo scrutinio, avevano raccolto il maggior numero di suffragi, portava gli elettori alla convergenza sulle posizioni più moderate, sfumando gli estremismi degli schieramenti in competizione (Ballini e Ridolfi, 2002). Questo sistema rimase in vigore inalterato fino alla riforma elettorale del 1882 sui requisiti degli elettori e sullo scrutinio di lista realizzata dalla Sinistra (leggi 22 gennaio 1882 e 7 maggio 1882). Il corpo elettorale ne risultò fortemente allargato, giungendo a comprendere tutti coloro che avessero raggiunto il ventunesimo anno di età e sapessero leggere e scrivere, o che pagassero una certa somma di imposta diretta. Naturalmente l'acquisizione del diritto di voto non comportò un'immediata capacità di sfruttarlo.

Differenti comportamenti fiscali in periodi caratterizzati da suffragi diversi possono naturalmente derivare da interessi diversi: il nostro elettore decisivo appartiene ad un gruppo elitario che, almeno fino al 1912, detiene il potere politico. I rappresentanti eletti sono in buona parte proprietari o rappresentanti di proprietari, interessati al bilancio pubblico per migliorare o consolidare le loro situazioni patrimoniali e proteggere la proprietà anche dal governo. La composizione della spesa pubblica nell'Italia del tempo risulta pertanto poco articolata: servizi istituzionali, come amministrazione, ordine pubblico e difesa, e infrastrutture, in particolare capitale sociale di sostegno allo sviluppo delle attività imprenditoriali sono le voci di spesa principali, oltre agli interessi sul debito.

Coloro che sono privi del diritto di voto possono influire sulle scelte solo indirettamente, attraverso la pressione sociale sulla classe dirigente (il malcontento diffuso, le rivolte popolari, frequenti soprattutto nel sud, il fenomeno del brigantaggio) oppure ricercandone il favore (la concezione di un liberalismo oligarchico nelle origini ma democratico negli obiettivi era tipica del pensiero moderato dell'epoca). In un sistema oligarchico l'*élite* politica ed economica è contraria a qualunque forma di tassazione a fini redistributivi, diversamente dalle classi meno abbienti. Queste, pertanto, possono ricorrere alla strada della rivoluzione per ottenere il diritto di voto e conseguentemente qualche forma di redistribuzione. Per ridurre il rischio di rivoluzione (che distrugge parte della ricchezza nazionale, e quindi dell'*élite* stessa) i ricchi concedono gradualmente l'estensione del suffragio, a partire dalle classi medie che hanno interessi redistributivi meno forti rispetto a quelli delle classi povere (Acemoglu e Robinson, 2006).

In seguito all'allargamento del suffragio alle classi meno abbienti l'elettore mediano, che è quello decisivo nella determinazione delle politiche, ha un reddito inferiore rispetto a quello appartenente all'*élite* e quindi un maggiore interesse verso una politica di redistribuzione del reddito dall'oligarchia alla classe povera (Meltzer e Richard, 1981).

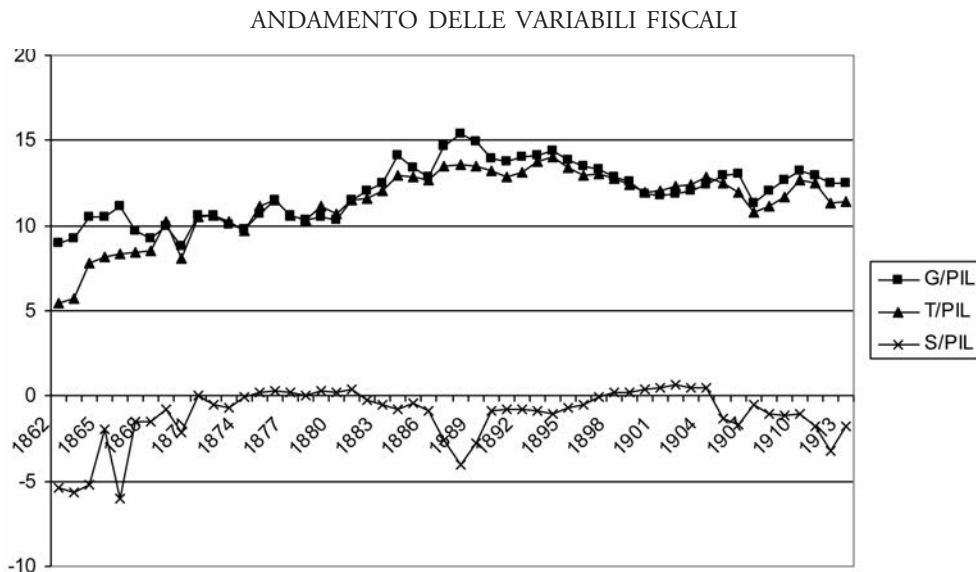
Si consideri inoltre che il pagamento degli interessi sul debito ereditato dagli Stati preunitari e la spesa militare (che rappresentano i tre-quarti della spesa pubblica totale in tempo di pace) richiedevano, in un'ottica di finanza ortodossa in cui il principio del bilancio in pareggio era tendenzialmente condiviso sia dalla Destra che dalla Sinistra, un ampio utilizzo dello strumento fiscale e, ciononostante, rendevano a volte inevitabile il ricorso al *deficit spending*. Il resto della spesa era rappresentato dall'Amministrazione generale, di modesta entità e dalla spesa produttiva (prevalentemente per l'istruzione e le infrastrutture) che oscillava nel cinquantennio tra il 20% e il 15% della spesa totale. Complessivamente la spesa pubblica ammontava all'8,98% del PIL nel 1862 e al 12,7% nel 1913. Fino al 1880 ha oscillato intorno al 10% del PIL, con un leggero incremento dovuto alle spese militari nel 1870, per poi crescere in maniera contenuta e fluttuare tra il 12% e il 13% circa.

La pressione fiscale era passata dal 5,41 nel 1862 al 10,62% dieci anni dopo, per poi assestarsi su valori compresi tra il 12% e il 13% circa tra il 1882 e i primi anni del 1900 fino a diminuire leggermente nel 1913 all'11,43% (Ragioneria Generale dello Stato, 1969; Fenoaltea, 2005)³. In conclusione, dal 1862 al 1913 il

³ Per un'analisi comparata dei diversi *dataset* relativi alle voci di bilancio si veda FEDERICO G. (2010).

valore medio del saldo di bilancio è stato di -1,12%, ha superato il 5% in alcuni anni durante il primo decennio unitario ed ha raggiunto livelli analoghi tra la fine degli anni 1880 e i primi anni 1890 ed alla vigilia della prima guerra mondiale. Il Grafico 1 mostra gli andamenti della spesa pubblica (G/PIL), delle entrate (T/PIL) e del *surplus* di bilancio (S/PIL).

GRAF. 1



3. - Il modello empirico: variabili, dati e metodologia

Al fine di analizzare le determinanti della spesa pubblica in Italia nel periodo 1861-1913 utilizziamo una specificazione della funzione di domanda di beni e servizi pubblici dell'elettore mediano generalmente stimata in letteratura (Mueller, 1989), con l'aggiunta dell'illusione fiscale⁴:

⁴ Nella stima della funzione di domanda può emergere un problema di identificazione nel caso si utilizzino soltanto le variabili prezzi e quantità e quindi l'informazione disponibile non sia sufficiente a distinguere la stima della funzione di domanda da quella di offerta. Generalmente per risolvere questo problema è necessario avere ulteriori informazioni. Ad esempio, supponiamo che nel tempo la curva di domanda si sposti in seguito a variazioni del reddito o delle preferenze dei consumatori, mentre la curva di offerta rimanga relativamente stabile. In questo caso le informazioni sui prezzi e le quantità servono ad identificare la curva di offerta. Viceversa, nel caso sia la curva di offerta a mostrare maggiore variabilità rispetto a quella di domanda, sarà quest'ultima ad essere identificata. Nel caso di interesse, da un lato l'inserimento di informazioni quantitative relative ad una pluralità di fattori che operano dal lato della domanda e, dall'altro, la relativa stabilità di quest'ultima in un periodo caratterizzato da cambiamenti lenti e graduali, ci confermano che si tratta di una stima della funzione di domanda.

$$(1) \log Q_t = a_1 \log P_t + a_2 \log YPC_t + a_3 \log D_t + a_4 \log V_t + a_5 \log POP_t + a_6 \log F_t + u_t$$

dove Q rappresenta la spesa pubblica totale⁵, P è il prezzo dei beni privati⁶, YPC è il reddito *pro-capite*, D è l'asimmetria della distribuzione dei redditi, V è il numero degli aventi diritto al voto, POP è la popolazione che fruisce dei beni pubblici, F è la misura indiretta della preferenza per il disavanzo e il suo coefficiente l'elasticità rispetto al prezzo percepito della domanda di beni e servizi pubblici ed u è il termine di errore. Il segno atteso per il coefficiente a_1 può essere positivo o negativo a seconda della relazione di complementarità o sostituzione tra beni privati e beni pubblici, mentre all'aumentare del prezzo del bene pubblico la sua domanda diminuisce. All'aumentare del reddito medio, se i beni pubblici sono normali, allora la loro domanda aumenta ($a_2 > 0$). Il modello di Meltzer e Richard (1983) mette in relazione la quota della spesa sul reddito con l'estensione del suffragio e il rapporto tra reddito medio e reddito mediano, che è un indice sintetico della distribuzione dei redditi. Il segno atteso del coefficiente a_5 è pertanto positivo in quanto, a causa di un effetto reddito, un'estensione del suffragio abbassa il reddito dell'elettore mediano⁷. In linea teorica, è anche ipotizzabile un effetto di congestione-prezzo che influisce sulle conseguenze che l'estensione del suffragio ha sul prezzo che l'elettore decisivo deve pagare. Tuttavia se il bene offerto è pubblico, come in questo caso, la spesa pubblica dovrebbe aumentare quando un incremento nel numero di coloro che sono avvantaggiati dalla spesa pubblica provoca una riduzione, *ceteris paribus*, del prezzo che ciascuno deve pagare. Questa considerazione rafforza il segno atteso positivo di a_4 . Inoltre l'estensione del diritto di voto, *ceteris paribus*, riduce il trasferimento di reddito che ciascun elettore riceve mettendo in atto un effetto di sfruttamento. Anche in questo caso la scarsa incidenza delle spese per trasferimenti sul totale della spesa pubblica ci fa ritenere questo effetto empiricamente ininfluenza. Anche il segno atteso di a_5 è positivo

⁵ Il volume fisico delle quantità consumate non è osservabile né lo sono i prezzi cui queste quantità sono acquistate. È possibile osservare solo le voci di spesa del bilancio pubblico.

⁶ In letteratura generalmente si inserisce tra i regressori il prezzo dei beni pubblici. Dal momento che l'affidabilità dei dati forniti dall'ISTAT (1957) sui prezzi dei beni pubblici è stata messa in discussione da lavori successivi (Giovanni Federico, comunicazione personale) ci limitiamo a considerare nella nostra analisi empirica solo il prezzo dei beni privati.

⁷ Si consideri, tuttavia, che se l'ampliamento del corpo elettorale avviene in un periodo di sviluppo del reddito, l'elettore mediano di oggi può essere più ricco in assoluto di quanto fosse ieri l'elettore mediano. Non è sicuramente questo il caso dell'Italia, in cui i tassi di sviluppo del reddito nel primo cinquantennio non sono stati molto elevati.

perché all'aumentare della popolazione aumenta la congestione nel consumo di alcuni beni forniti dallo stato e quindi ne deve aumentare l'offerta.

Utilizzare il reddito medio presenta il limite di sottostimare il reddito dell'elettore decisivo, in quanto la distribuzione del reddito è fortemente asimmetrica e la maggioranza degli individui ha un reddito inferiore a quello medio⁸. Se è ipotizzabile che un incremento del reddito medio (YPC) e della disuguaglianza della distribuzione (D) producano un incremento del reddito dell'elettore mediano (a_2 e $a_3 > 0$), un'estensione del diritto di voto dovrebbe abbassarlo e avere un impatto positivo sulla spesa pubblica ($a_2 > 0$). In una democrazia a suffragio ristretto generalmente i più ricchi coincidono con gli aventi diritto al voto; l'estensione del suffragio comporta l'accesso alle urne di gruppi di popolazione più poveri facendo diventare decisivo un elettore più povero rispetto al precedente che chiede maggiore spesa pubblica redistributiva.

L'elemento centrale della nostra relazione empirica è la considerazione del prezzo *percepito* dei servizi pubblici come determinante della quantità domandata e la conseguente decisione di finanziamento della spesa mediante il ricorso al disavanzo. In altri termini, ci proponiamo di verificare se gli effetti dell'illusione fiscale sul modo in cui il contribuente percepisce la propria posizione economica in termini di costi e benefici abbiano avuto un ruolo nella determinazione della frazione di spesa pubblica che dovrà essere finanziata dalla tassazione corrente ($F = T/G$).

La letteratura empirica ha stimato la funzione di domanda sia nei livelli che nelle variazioni, al fine di ottenere stime di lungo e di breve periodo della variabilità della spesa pubblica rispetto alle variabili indipendenti individuate. Se la stima dell'equazione in livelli può presentare problemi di regressione spuria data la non stazionarietà delle serie, la stima dell'equazione in differenze prime omette le informazioni di lungo periodo. Allo scopo di correggere questi problemi, imponiamo una struttura dinamica sulla relazione tra *deficit* e spesa pubblica del tipo "meccanismo di correzione dell'errore" che combina entrambi i modelli in livelli e in differenze prime, attraverso l'uso delle variabili differenziate e dei ritardi nei livelli opportunamente scelti. Secondo il modello di correzione dell'errore, l'equazione (I) diventa:

⁸ Per risolvere questo problema BROSIO G. e MARCHESE C. (1986, p. 131) ricorrono ad una serie di indicatori che collegano gli andamenti del reddito mediano: il reddito medio della popolazione (YPC), l'asimmetria della distribuzione dei redditi (D) e la quota della popolazione avente diritto al voto (V).

$$(2) \quad \Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t$$

dove y_t è un vettore delle K serie storiche discusse precedentemente, $\Pi = -(I_K - A_1 - \dots - A_p)$, $\Gamma_i = -(A_{i+1} + \dots + A_p)$, A_i essendo matrici di coefficienti ($K \times K$) e u_t è un termine di errori non osservabili. Il primo termine dell'equazione (2) contiene i parametri di breve periodo ed è stazionario, mentre il secondo termine contiene i parametri di lungo periodo ed è integrato di ordine 1. La spesa Q è determinata da F , la quota di tassazione corrente. È possibile anche un fenomeno di causazione inversa: l'aumento del gettito fiscale, rendendo nuove risorse disponibili, spinge i politici ad aumentare la spesa. La struttura autoregressiva del modello stimato permette di considerare questo aspetto⁹.

La spesa pubblica (Q) è misurata in termini reali. Il livello dei prezzi dei beni (P) è l'indice dei prezzi al consumo (Fenoaltea, 2002). Il reddito *pro-capite* (YPC) è misurato in termini reali con il prodotto interno lordo al costo dei fattori (Fenoaltea, 2005). In mancanza di un buon indice della distribuzione del reddito (D), si è utilizzata una misura di disuguaglianza alla Kuznets, che fa riferimento alle differenze nei prodotti *pro-capite* tra settori (Fenoaltea, 2005) ed è calcolata come rapporto tra deviazione standard e la media del reddito *pro-capite* nei settori dell'agricoltura e dell'industria¹⁰. La variabile V è rappresentata alternativamente dal numero effettivo di votanti o dal numero degli aventi diritto al voto (ISTAT, 1958; Ghisalberti, 1986). La variabile POP è data dalla popolazione totale pre-

⁹ Stimando il modello con il metodo dei minimi quadrati senza struttura autoregressiva, si ottengono *p-value* del test di Hausman che vanno da 0,038 a 0,744 e che indicano come l'ipotesi nulla di esogenità possa essere respinta.

¹⁰ La costruzione di un rapporto salari/PIL come misura della disuguaglianza dei redditi avrebbe comportato la formulazione di una serie di ipotesi discutibili. FENOALTEA S. (2002) fornisce i salari orari dei lavoratori *unskilled* dell'agricoltura e dell'industria. Per ottenere il monte salari sarebbe necessario conoscere la durata della giornata media di lavoro in questi settori e la forza lavoro anno per anno. ERCOLANI F. (1969) fornisce, per il periodo di nostro interesse, una stima per gli anni 1881, 1901 e 1911; pertanto bisognerebbe interpolare questi dati rispetto ai diversi cicli economici (supponendo una qualche elasticità tra produzione ed occupazione). Resterebbero fuori da questa stima tuttavia i dati riguardanti i lavoratori pubblici (per i quali Ercolani fornisce un numero indice). Si consideri inoltre che BOURGUIGNON F. e MORRISSON C. (2001) forniscono una stima della distribuzione del reddito dal 1820 al 1992 per un ampio numero di paesi del mondo. In particolare per l'Italia – con riferimento al periodo di nostro interesse – la quota dei salari è stimata negli anni 1870 e 1910, è costante e simile a quella della Francia. Questi ultimi dati ci sono tuttavia sembrati inadatti ad una stima temporale. Alla luce di queste considerazioni abbiamo preferito utilizzare l'indicatore alla Kuznets.

sente sul territorio nazionale (ISTAT, 1958). La variabile F è misurata dal rapporto tra le entrate tributarie e la spesa pubblica. Tutti i valori reali sono calcolati in base 1911 e sono espressi in logaritmi.

4. - Risultati

Analizziamo dapprima la struttura stocastica delle serie per verificarne la stazionarietà o non stazionarietà. A tal fine, si è utilizzata la statistica Augmented Dickey-Fuller. Tutte le sette serie continue si sono rivelate non stazionarie nei livelli (Tavola 1), mentre la variabile dicotomica per le guerre naturalmente è stazionaria. Dato che le serie sono tutte non stazionarie siamo interessati a vedere se sono cointegrate utilizzando il *trace test* di Johansen (1991). I risultati presentati nella Tavola 2 sono ottenuti inserendo un numero di ritardi uguale a 4, come indicato dall'Akaike Information Criterion. La Tavola 2 fornisce nella prima colonna il valore del test (rapporto di verosimiglianza, RV), nella seconda colonna il relativo *p-value*, e nelle successive tre i valori critici del test al 10%, 5% e 1%. Non si può respingere, alla soglia del 5%, l'ipotesi nulla che vi siano 6 vettori di cointegrazione tra le variabili considerate.

TAV. 1

TEST ADF DI RADICI UNITARIE

	Statistica	Trend	Lag
Q	-0,833	SI	1
YPC	-0,103	SI	2
D	-2,594	SI	4
V	-1,978	SI	1
P	-1,749	NO	2
POP	-2,282	NO	7
F	-1,978	SI	1

I valori critici per il test ADF con *trend* sono -3,96, -3,41 e -3,13 rispettivamente ai livelli dell'1%, 5% e 10%. I corrispondenti valori critici per il test senza *trend* sono -3,43, -2,86 e -2,57. Il numero di *lag* è determinato secondo l'Akaike Information Criterion.

Per stimare il modello di correzione dell'errore abbiamo fissato a 3 il numero di ritardi delle variabili espresse nelle differenze prime in base all'Akaike Information Criterion e aggiunto l'intercetta ed il *trend*. Per la stima abbiamo utilizzato la procedura di Johansen (1991). Nei grafici seguenti rappresentiamo le funzioni di risposta ad impulso della spesa pubblica rispetto alle variabili descritte in pre-

cedenza. Gli intervalli di confidenza del 95% sono ottenuti mediante il metodo di Hall (1992) con una procedura di *bootstrapping* con 1.000 ripetizioni. La Tavola 3 riporta le stime del VECM.

TAV. 2

TEST DI COINTEGRAZIONE DI JOHANSEN

	<i>RV</i>	<i>P</i>	90%	95%	99%
0	416,96	0,000	181,08	187,25	199,18
1	304,00	0,000	144,81	150,35	161,11
2	220,65	0,000	112,54	117,45	127,04
3	145,47	0,000	84,27	88,55	96,97
4	92,28	0,000	60,00	63,66	70,91
5	52,63	0,003	39,73	42,77	48,87
6	28,07	0,024	23,32	25,73	30,67

Nel Grafico 2 viene rappresentata la risposta della spesa pubblica ad uno *shock* sul prezzo dei beni privati. Nel breve periodo l'impatto è negativo, indicando una complementarità tra beni privati e pubblici; nel lungo periodo, invece, esiste una relazione di sostituzione tra beni privati e beni pubblici. L'attendibilità di questo risultato risiede nella limitatezza delle risorse propria di un'economia, come quella italiana del tempo, in cui lo sviluppo economico è appena agli inizi e lo sforzo fiscale richiesto ai contribuenti alquanto intenso: il costo opportunità della spesa pubblica è rappresentato dalla spesa privata.

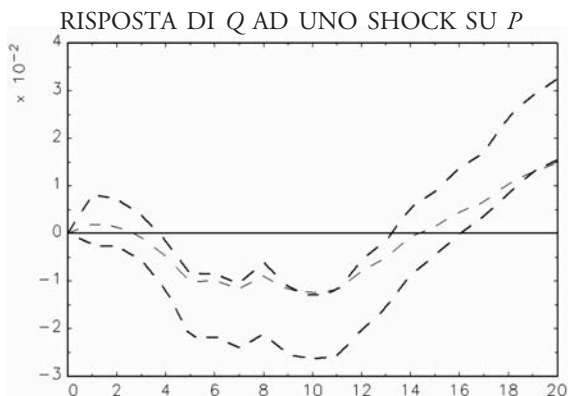
A seguito di uno *shock* sul reddito (Grafico 3), si evidenzia dapprima un aumento della spesa pubblica, e poi una sua riduzione. L'effetto nel breve-medio periodo è tipico dei paesi a basso reddito, dove questi beni sono considerati superiori (legge di Wagner). Nel lungo periodo invece si osserva uno spostamento della domanda verso i beni privati, ulteriore evidenza di quella sostituibilità tra spesa pubblica e privata a cui si è fatto riferimento in precedenza.

TAV. 3

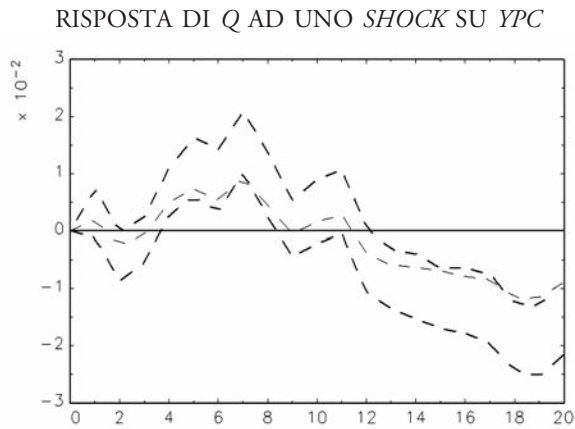
LA STIMA DEL VECM

ΔQ_{t-1}	-0,325** (0,148)	ΔP_t	-0,475* (0,243)
ΔQ_{t-2}	-0,114* (0,053)	ΔP_{t-1}	-0,520* (0,303)
ΔQ_{t-3}	0,185 (0,183)	ΔP_{t-2}	-0,333* (0,287)
ΔYPC_t	-0,290*** (0,025)	ΔP_{t-3}	0,272 (0,234)
ΔYPC_{t-1}	0,215*** (0,035)	ΔPOP_t	-0,563** (0,274)
ΔYPC_{t-2}	-0,538*** (0,165)	ΔPOP_{t-1}	0,564* (0,322)
ΔYPC_{t-3}	0,012 (0,442)	ΔPOP_{t-2}	0,542* (0,369)
ΔD_t	-0,328** (0,147)	ΔPOP_{t-3}	-0,993** (0,394)
ΔD_{t-1}	0,161* (0,059)	ΔF_t	-0,243*** (0,034)
ΔD_{t-2}	0,220 (0,593)	ΔF_{t-1}	-0,433 (0,505)
ΔD_{t-3}	0,964 (0,615)	ΔF_{t-2}	0,756* (0,362)
ΔV_t	0,288** (0,096)	ΔF_{t-3}	0,552* (0,377)
ΔV_{t-1}	-1,864*** (0,502)	Intercetta	-5,756*** (1,066)
ΔV_{t-2}	-0,127 (0,142)	Trend	-0,395* (0,293)
ΔV_{t-3}	0,516** (0,144)		
R^2 corr.	0,983		
N	49		

GRAF. 2



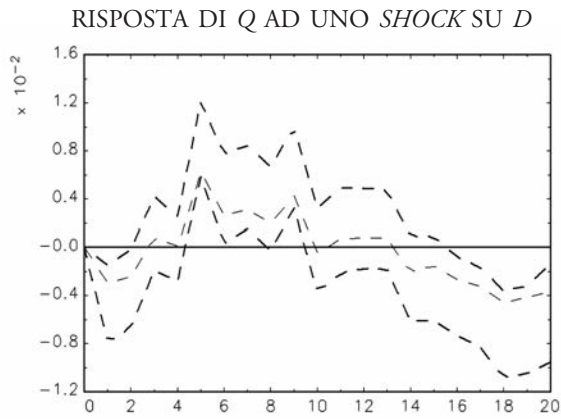
GRAF. 3



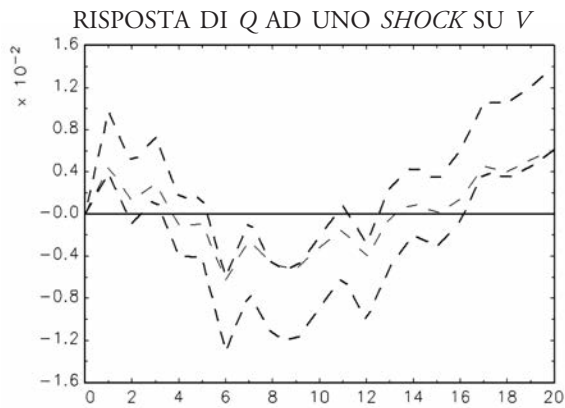
Il Grafico 4 analizza l'effetto di uno *shock* sulla disuguaglianza di reddito. Qui si possono individuare tre diverse fasi. Nel brevissimo periodo si assiste ad una riduzione di Q , seguita da un suo aumento sopra il livello iniziale nel medio periodo (ma di limitata significatività) ed infine una nuova riduzione. L'aumento del numero dei votanti (Grafico 5) ha un andamento speculare: ad un aumento della spesa pubblica segue una riduzione (questa volta significativa) ed infine un aumento. Nel periodo in esame sembrano avere un effetto molto lento i cambiamenti delle regole del gioco. È in una dimensione temporale di lungo periodo che le decisioni di politica fiscale rispecchiano i mutamenti nelle regole elettorali. Le estensioni successive del suffragio all'epoca inserivano nel gioco elettorale persone dotate via via di minore potere e capacità politica. Non è pensabile che esse fossero in grado di incidere rapidamente sulle decisioni di bilancio se si considera che il 2% della popolazione che nel 1861 aveva diritto di voto costituiva un *élite* ristretta ma influente (professionisti, alti burocrati e ufficiali, nobili, proprietari terrieri), in grado di condizionare il funzionamento del sistema politico anche con mezzi diversi dalla semplice partecipazione elettorale. Nel Grafico 6 uno *shock* sulla popolazione risulta non avere effetti sulla spesa pubblica.

Infine, uno *shock* su F ha un effetto positivo sulla spesa nel brevissimo e nel lungo periodo, mentre ha un effetto negativo nel medio (Grafico 7). Questo risultato suggerisce che la relazione tra aumenti della spesa pubblica e *deficit* sia una relazione di lungo periodo.

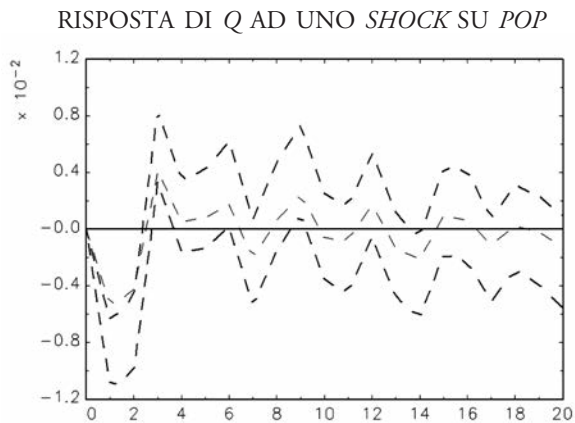
GRAF. 4



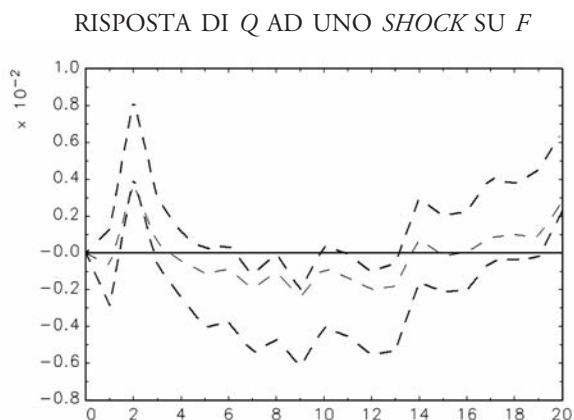
GRAF. 5



GRAF. 6



GRAF. 7



5. - Conclusioni

In questo lavoro abbiamo stimato un modello di domanda di spesa pubblica in Italia nel periodo 1861-1913, che inserisce tra le determinanti l'illusione fiscale al fine di esaminare il comportamento della classe politica del tempo nelle decisioni di finanza pubblica. Dall'analisi empirica emerge che la relazione tra crescita della spesa pubblica e disavanzi di bilancio è stata fondamentalmente una relazione di lungo periodo. Le altre variabili sono generalmente in linea con le previsioni della teoria economica. Si rivela di un certo interesse la relazione negativa di breve periodo tra spesa pubblica e prezzi dei beni privati, indicando una complementarità tra beni privati e pubblici; nel lungo periodo, invece, emerge una relazione di sostituzione tra beni privati e beni pubblici in un contesto di risorse limitate quale quello dell'economia italiana del tempo in cui lo sviluppo economico è appena agli inizi e lo sforzo fiscale richiesto ai contribuenti alquanto intenso. L'effetto, positivo nel breve periodo, del reddito sulla spesa pubblica che caratterizza paesi nelle fasi iniziali dello sviluppo economico, diventa negativo nel lungo periodo dove si osserva uno spostamento della domanda verso i beni privati in un'ottica di sostituibilità tra spesa pubblica e privata. L'impatto della disuguaglianza del reddito sulla spesa pubblica è oscillante, come pure quello del numero dei votanti, suggerendo che nel periodo in esame i cambiamenti delle regole del gioco hanno avuto un effetto molto lento. D'altronde le estensioni successive del suffragio all'epoca inserivano nel gioco elettorale persone con potere e capacità politica via via minore che non erano in grado di utilizzare pienamente e rapidamente il diritto di voto per influire sulle decisioni di bilancio.

BIBLIOGRAFIA

- ACEMOGLU D. - ROBINSON J.A., *Economic Origins of Dictatorship and Democracy*, Cambridge University Press, 2006.
- ARTONI R. - BIANCHINI S., «Il debito pubblico dall'unità ad oggi», in CIOCCA P. - TONIOLO G. (a cura di), *Storia Economica d'Italia*, Roma-Bari, Laterza, 2003.
- BALLINI P. - RIDOLFI M. (a cura di), *Storia delle campagne elettorali in Italia*, Milano, Bruno Mondadori, 2002.
- BARRO R.J., «Are Government Bonds Net Wealth?», *Journal of Political Economy*, vol. 82, 1974, pages 1095-1117.
- BROSIO G. - MARCHESE C., *Il potere di spendere*, Bologna, Il Mulino, 1986.
- .-, «The Growth of Government under Different Redistributive Rules: A Long Term Study of the Italian Case», *Public Finance Quarterly*, vol. 16, 1988, pages 439-463.
- BUCHANAN J.M., *Fiscal Theory and Political Economy*, Chapel Hill, University of North Carolina Press, 1960.
- BUCHANAN J.M. - WAGNER R.E., *Democracy in Deficit*, New York, Academic Press, 1977.
- BOURGUIGNON F. - MORRISON C., «Data Sources for Inequality among World Citizens, 1820-1992», 2001, www.delta.ens.fr/XIX.
- ERCOLANI F., «Documentazione statistica di base», in FUÀ G. (a cura di), *Lo sviluppo economico in Italia*, Milano, Franco Angeli Editore, 1969.
- FEDERICO G., «Always on the Brink: Piedmont and Italy», in CARDOSO J.L. - LAINS P. (eds.), *Paying for the Liberal State*, Cambridge University Press, 2010.
- FENOALTEA S., «Production and Consumption in Post-Unification Italy: New Evidence, New Conjectures», *Rivista di Storia Economica*, no. 18, 2002, pages 251-298.
- .-, «The Growth of the Italian Economy, 1861-1913: Preliminary Second-Generation Estimates», *European Review of Economic History*, 9, 2005, pages 273-312.
- .-, *L'Economia Italiana dall'Unità alla Grande Guerra*, Roma-Bari, Laterza, 2006.
- FERRERO M. - SANTAGATA W., «Il ciclo del "galantuomo": manovra economica e scadenze elettorali nel Regno d'Italia, 1867-1913», *Micros*, n. 2, 1981, pp. 25-37.
- FIORINO N. - RICCIUTI R., «State Protection, Interest Groups and Government Spending in Italy, 1876-1913», *Homo Oeconomicus*, no. 27, 2011, pages 443-461.
- FRATIANNI M. - SPINELLI F., «The Growth of Government in Italy: Evidence from 1861 to 1979», *Public Choice*, no. 39, 1982, pages 221-243.
- GHISALBERTI C. *Storia costituzionale d'Italia 1848/1948*, Bari, Laterza, 1986.
- GOETZ C. - DA EMPOLI D., «La teoria dell'illusione finanziaria di Amilcare Puviani», *Studi di Finanza Pubblica*, Milano, Giuffrè, 1976.
- HALL P., *The Bootstrap and Edgeworth Expansion*, New York, Springer-Verlag, 1992.

- ISTAT, «Le rilevazioni statistiche in Italia dal 1861 al 1956», *Annali di Statistica*, serie VIII, voll. 5-8, Roma, 1957.
- .-, *Sommario di statistiche storiche italiane*, Roma, 1958.
- JOHANSEN S., «Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models», *Econometrica*, vol. 59, 1991, pages 1551-1581.
- MARANINI G., *Storia del potere in Italia 1848-1967*, Firenze, Le Monnier, 1967.
- MARONGIU G., *La politica fiscale della destra storica (1861-1876)*, Torino, Einaudi, 1996a.
- .-, *La politica fiscale della sinistra storica (1876-1896)*, Torino, Einaudi, 1996b.
- MELTZER A.H. - RICHARD S.F., «A Rational Theory of the Size of Government», *Journal of Political Economy*, vol. 89, 1981, pages 914-927.
- .-, -.-, «Tests of a Rational Theory of the Size of Government», *Public Choice*, vol. 47, 1983, pages 231-265.
- MUELLER D.C., *Public Choice II*, Cambridge University Press, 1989.
- OATES W.E., «On the Nature and Measurement of Fiscal Illusion: A Survey», *Taxation and Fiscal federalism: Essays in honour of Russell Mathews*, Anu Press, Sydney, 1988.
- PLEBANO A., *Storia della finanza italiana nei primi quarant'anni dell'unificazione*, voll. I-II-III, Torino, Roux e Viarengo Editori, 1900.
- PUVIANI A., *Teoria della illusione nelle entrate pubbliche*, Perugia, 1897.
- RAGIONERIA GENERALE DELLO STATO, *Il bilancio dello Stato italiano dal 1862 al 1967*, Roma, Poligrafico dello Stato, 1969.
- RICARDO D., «Funding System», 1820, in SRAFFA P. (ed.), *The Works and Correspondence of David Ricardo*, Cambridge University Press, 1951.
- RICCIUTI R., «The Quest for a Fiscal Rule: Italy, 1861-1998», *Cliometrica*, vol. 2, 2008, pages 259-274.
- TONIOLO G. - GANUGI P., «Il debito pubblico italiano in prospettiva secolare, 1876-1947», in *Il disavanzo pubblico come problema strutturale*, Ente Einaudi (a cura di), Bologna, Il Mulino, 1992.
- ZAMAGNI V., *Dalla periferia al centro*, Bologna, Il Mulino, 1990.
- .-, «Il debito pubblico italiano 1861-1946: ricostruzione della serie storica», *Rivista di Storia Economica*, vol. 14, 1998, pp. 207-242.

Le scelte di localizzazione delle opere pubbliche: il fenomeno *Nimby*

Roberta Occhilupo Giuliana Palumbo Paolo Sestito*
Banca d'Italia

The paper shows that siting procedures taking local preferences into account should be preferred when the heterogeneity in preferences across communities is large relative to that in constructing and operating costs. Preferences are better elicited through auction-like mechanisms rather than multilateral negotiations if: credible information about the risks are available; conflicting preferences at the local level can be preliminarily aggregated; compensations are mainly monetary. The conflict between highly centralized siting procedures and highly decentralized administrative institutions is identified as a critical point.

[JEL Classification: D73; D82; H7].

Keywords: scelta localizzativa; preferenze locali; costi di costruzione e gestione; meccanismi d'asta; negoziazione.

* <robertamargherita.occhilupo@bancaditalia.it>; <giuliana.palumbo@bancaditalia.it>; <paolo.sestito@bancaditalia.it>; Area Ricerca Economica e Relazioni Internazionali. Gli Autori desiderano ringraziare l'Agenzia di ricerche informazione e società (ARIS) per avere fornito i dati ed Emmanuele Somma per il contributo dato alla loro elaborazione. Si ringraziano inoltre i rappresentanti delle imprese contattate per i molti interessanti spunti di analisi forniti. Un particolare ringraziamento va a Gian Luigi Albano, Fabrizio Balassone, Silvia Boccato, Fulvio Rossi e Luisa Torchia per le utili discussioni, a Christian Iaione per i preziosi suggerimenti sulla parte istituzionale, ai *referee* della *Rivista di Politica Economica* per i loro commenti su una versione precedente del lavoro. Gli Autori restano i soli responsabili di eventuali errori. Le opinioni espresse sono personali e non coinvolgono l'istituzione di appartenenza. L'analisi normativa non tiene conto delle novità legislative intervenute nel corso del 2012.

Il miglior modo per accelerare la realizzazione di un impianto eolico negli Stati Uniti è aggirare la tempesta scatenata dalle opposizioni locali. La regola generale dell'industria per evitare del tutto il sorgere del fenomeno Nimby è proporre la localizzazione di un progetto ad almeno 30 miglia dal più vicino Starbucks.

(Business Week, 2009).

1. - Introduzione

L'opposizione agli investimenti infrastrutturali è fenomeno ormai endemico, in Italia come altrove. A essere coinvolte sono sia le infrastrutture in senso proprio, sia molti impianti produttivi rispetto ai quali l'accresciuta sensibilità ambientale, e i connessi standard fissati dal legislatore sono vissuti da molte imprese occidentali come una sorta di ostacolo competitivo rispetto ai paesi emergenti. Queste difficoltà si manifestano non solo per le più tradizionali industrie inquinanti, ma anche per quelle che dovrebbero essere le industrie "verdi" e sostenibili del futuro: come riportato da Business Week nel 2009, negli Stati Uniti l'installazione di pale eoliche viene ritenuta dagli esperti del campo possibile solo se a una distanza di almeno 30 miglia dal più vicino *Starbucks*.

La trattazione del fenomeno è peraltro complicata da vari aspetti, di tipo sia statistico sia concettuale. Sotto il primo profilo, la misurazione delle contestazioni agli investimenti infrastrutturali e dei loro effetti è scarna e poco sistematica: poco si sa su quali tipologie di investimenti attirino maggiori contestazioni e su quali aspetti, di contesto socio-economico e istituzionale, si correlino maggiormente con la presenza di contestazioni; ancor meno si sa sugli effetti delle contestazioni e sulle modalità con cui le istanze sottostanti le contestazioni possano essere tenute in conto senza costituire un blocco per il paese. Sotto il secondo profilo, poco si sa sulle diverse tipologie di contestazione, il *Nimby* (*Not In My BackYard*) – definito come contestazione alla localizzazione di una certa opera e non necessariamente dell'opera in quanto tale – spesso sconfinando in una sindrome *Banana*, ovvero *Build Absolutely Nothing Anywhere Near Anything (or Anyone)*.

In quanto segue si presenta una trattazione sistematica del *Nimby* per l'Italia, coniugando una schematizzazione concettuale del fenomeno, dei fattori sottostanti e delle modalità alternative di risoluzione del contrasto di interessi che ne è alla base, con una esposizione dell'*iter* amministrativo seguito per l'individuazione della scelta di localizzazione delle infrastrutture.

La schematizzazione concettuale servirà innanzitutto a distinguere il *Nimby* – inteso come opposizione a una certa localizzazione da parte delle comunità interessate – dal-

l'opposizione valoriale a questa o quella tipologia di investimenti. In pratica le due cose possono spesso essere frammischiate perché il fatto che una certa fetta di popolazione possa *a priori* osteggiare un certo tipo di impianti tenderà ad amplificare e magari a “giustificare” l'opposizione di una certa comunità motivata dai suoi interessi particolari; similmente, il fatto che una certa opera sia obiettivamente poco motivata (o motivata sulla base di argomenti mal espressi e poco verificabili) in termini di costi e benefici per l'intera collettività nazionale, rende più probabile l'insorgere di una opposizione legata agli interessi particolari della specifica comunità ospitante. Una definizione più circoscritta del *Nimby* consente però di analizzarne meglio le ragioni economiche – il contrasto tra le utilità generali e complessive di una data opera e le disutilità specifiche della singola comunità ospitante – i fattori di contesto che possono favorire l'insorgere delle opposizioni e le modalità con cui uno “scambio” e un bilanciamento tra interessi “generali” ed interessi “particolari” può essere realizzato.

Sulla base di questa schematizzazione sarà poi possibile esaminare in che misura le regole amministrative esistenti in Italia siano o meno coerenti con l'obiettivo di ottenere il bilanciamento di interessi sopra detto.

2. - Alcune evidenze empiriche per l'Italia

Le uniche informazioni disponibili sul fenomeno *Nimby* in Italia derivano dal censimento delle notizie a stampa condotto dall'Osservatorio permanente *Nimby Forum* dell'*Agenzia di Ricerche informazione e società (ARIS)*. Del fenomeno *Nimby* viene peraltro data una definizione particolare: il fatto che, con riferimento a un impianto o a un'infrastruttura, progettata o anche già in parziale esercizio, la stampa *riporti* notizia di una contestazione. Le statistiche sui casi rilevati mostrano che il numero di impianti contestati è cresciuto nel tempo, soprattutto con riguardo alle nuove contestazioni che sono passate da 85 del 2006 a 158 del 2010 (Tav. 1). Come atteso, i settori maggiormente interessati sono quelli dell'energia (soprattutto degli impianti di produzione) e dei rifiuti (soprattutto termovalorizzatori) (Tav. 2)¹. Riguardo ai soggetti che partecipano alla protesta², quelli mag-

¹ Disponiamo di tali informazioni solo fino al 2009. Il totale degli impianti nella Tavola 2 non coincide con quello della Tavola 1 perché per 44 impianti non disponiamo delle informazioni relative al settore di appartenenza.

² Per ciascun impianto contestato, la banca dati rileva l'orientamento – favorevole, neutrale, contrario – all'impianto di vari *stakeholders*: amministratori pubblici locali, amministratori pubblici nazionali, cittadini, associazioni ambientaliste, gestore/costruttore, associazioni di categoria, sindacati, enti pubblici, altro.

giormente attivi e più spesso con posizioni contrarie sono gli amministratori pubblici locali, seguiti dai cittadini (Graf. 1). Nel confronto con le altre categorie, i due soggetti insieme rappresentano in media circa il 63 per cento del totale delle posizioni espresse (favorevoli, neutrali e contrarie) (41 per cento amministratori pubblici locali e 22 per cento cittadini).

TAV. 1

IMPIANTI CONTESTATI NEL PERIODO 2005-2010						
	2005	2006	2007	2008	2009	2010
impianti contestati	237	173	193	267	278	320
di cui nuovi*	–	85	106	136	156	158

* non presenti nell'anno/negli anni precedenti.

Nel complesso le informazioni disponibili forniscono un quadro piuttosto parziale del fenomeno *Nimby* nel nostro paese. La mancanza di rilevazioni sistematiche sugli impianti *a rischio* di contestazione – che non vengono però contestati – non consente di quantificare l'effettiva entità del fenomeno né di stimare con precisione quali siano le condizioni di contesto e le tipologie di impianto più soggette a *Nimby*. La mancanza di informazioni sugli esiti delle contestazioni non consente di dire nulla sugli effetti del *Nimby* sugli investimenti infrastrutturali (in termini di costi e tempi di realizzazione) né di confrontare diversi assetti di sua risoluzione.

Alcune indicazioni sulla natura del fenomeno che utilizzano misure di intensità della contestazione³ e di permanenza nel tempo nello stato di contestazione, costruite impiegando i dati del *Nimby Forum* sono state fornite da Occhilupo *et al.* (2011).

³ Il numero di articoli con orientamento contrario all'impianto su stampa locale e il numero di posizioni contrarie espresse da amministratori locali e cittadini.

TAV. 2

IMPIANTI CONTESTATI NEL PERIODO 2005-2009:
DISTRIBUZIONE PER MACRO-SETTORE

	2005	2006	2007	2008	2009
Energia	51	55	75	117	159
<i>di cui: imp. produzione</i>	45	38	62	97	132
<i>rigassificatori</i>	2	8	11	13	10
<i>altro</i>	4	9	2	7	17
Infrastrutture	16	21	24	22	21
<i>di cui: autostradale</i>	0	7	2	2	2
<i>ferroviaria</i>	6	5	6	7	4
<i>stradale</i>	7	4	8	8	11
<i>generica</i>	3	5	8	5	4
Rifiuti	129	94	89	122	93
<i>di cui: discariche</i>	35	16	21	31	16
<i>termovalorizzatori</i>	57	54	45	55	42
<i>altro</i>	37	24	23	36	35
Altro	4	1	4	2	5
Totale	200	171	192	263	278

Correlando queste misure a varie caratteristiche dell'impianto e dei territori interessati gli autori trovano che, coerentemente a quanto emerge dall'analisi descrittiva, l'intensità della protesta si è accresciuta nel tempo (anche se il riferimento è solo a un quinquennio) ed è maggiore nei casi che presentano rischi elevati e concentrati (termovalorizzatori, rigassificatori, discariche, impianti di produzione dell'energia). La complessità dell'opera – in termini di pluralità di siti, frammentazione istituzionale e livelli di governo coinvolti – e le difficoltà di dialogo tra questi ultimi sono fattori che tendono a esacerbare l'intensità della protesta⁴.

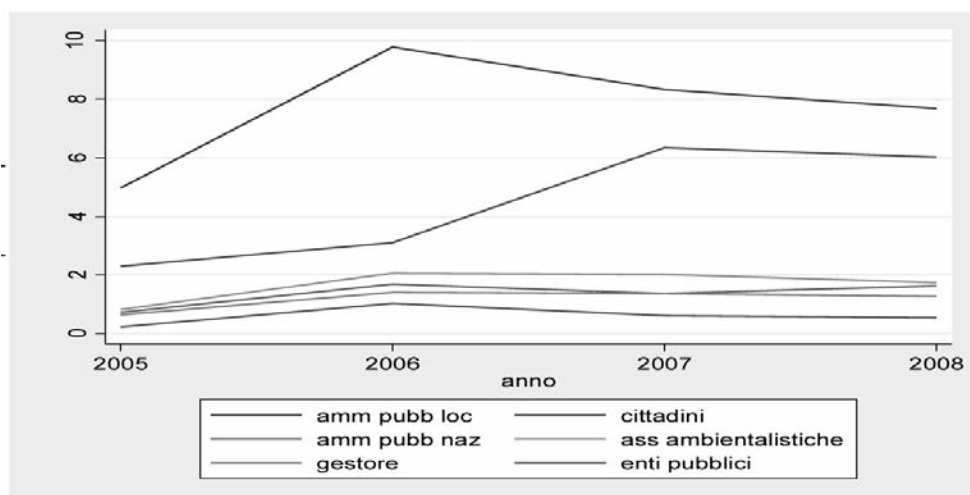
Pur con le cautele dovute alla natura delle informazioni adoperate, questi primi risultati empirici sembrano dunque evidenziare, da un lato, la potenziale rilevanza del contesto istituzionale sul fenomeno *Nimby*, dall'altro, la presenza di ragioni economiche sottostanti, legate alla natura di "*public bad*" degli impianti contestati. Le opposizioni sono maggiori per gli impianti che presentano un più forte con-

⁴ In particolare l'intensità della protesta è minore laddove sindaco e governo regionale hanno lo stesso orientamento politico, rispetto al caso in cui il sindaco appartiene ad un partito all'opposizione nella regione o vi sia una lista civica o un commissario di governo, ovvero laddove l'impianto risiede in una sola regione, rispetto al caso in cui sono coinvolte più regioni aventi governi con eguale o diverso orientamento politico. È anche da segnalare come le variabili di congruenza politica sono in generale più rilevanti nel caso delle contestazioni degli amministratori locali, laddove invece i comportamenti della cittadinanza ne sono scarsamente influenzati. Ciò sembrerebbe essere coerente con la natura a volte strumentale delle contestazioni da parte degli amministratori locali.

trasto tra la natura “diffusa” dei benefici prodotti, che vengono goduti dall’intera collettività, e quella “concentrata” dei costi, che ricadono interamente sulla comunità ospitante. Nei paragrafi che seguono cercheremo di approfondire entrambi questi aspetti, sia analizzando le modalità con cui bilanciamento tra interessi generali ed interessi particolari può essere realizzato, sia con riferimento al se e al come tale bilanciamento è stato raggiunto nel contesto italiano.

GRAF. A

POSIZIONI CONTRARIE ESPRESSE PER CATEGORIE DI SOGGETTI IN RAPPORTO AGLI IMPIANTI CONTESTATI



3. - Cosa è il *Nimby*

L’acronimo *Nimby* (“*Not In My BackYard*”) indica il fenomeno di opposizione alla localizzazione di nuovi impianti o infrastrutture da parte delle comunità direttamente interessate⁵. Esso va quindi distinto dalla opposizione *tout court* all’effettuazione di un determinato investimento ritenuto inutile o troppo costoso; almeno in linea di principio, infatti, il *Nimby* riguarda situazioni in cui vi è consenso circa la desiderabilità sociale dell’infrastruttura⁶ e l’opposizione nasce dal

⁵ D’ora in poi useremo i due termini in maniera intercambiabile, salvo laddove si vorranno evidenziare differenze tra impianti che servono a produrre beni e servizi vendibili sul mercato ed infrastrutture pubbliche il cui utilizzo sia prevalentemente a titolo gratuito.

⁶ I benefici netti complessivi, al netto anche delle disutilità sopportate dalla comunità ospitante, sono strettamente positivi per almeno una delle localizzazioni/varianti tecnicamente possibili.

fatto che, a fronte di benefici diffusi all'intera collettività, i costi sono concentrati sulla comunità ospitante.

I costi connessi con una determinata localizzazione sono generalmente di due tipi: costi di costruzione e gestione dell'infrastruttura e disutilità che gravano sulla comunità ospitante. I primi riguardano aspetti di natura tecnica e possono variare tra le diverse localizzazioni, di norma per motivi strettamente tecnologici (la connotazione orografica ed altimetrica di un sito, la necessità o meno di porre in essere accorgimenti di sicurezza per via di rischi sismici, ecc.)⁷. Tali costi sono generalmente noti al costruttore/gestore⁸. Le seconde possono essere prettamente economiche (venir meno di usi alternativi del territorio, perdita di valore degli immobili ecc.), ma più spesso riguardano rischi per l'ambiente e la salute, danni al paesaggio. Il fatto che questa seconda tipologia sia particolarmente rilevante rende in genere difficile valutare gli effetti negativi che ciascuna comunità sopporta dall'ospitare l'infrastruttura. Soprattutto, le differenze tra le diverse comunità possono riflettere diversità nella struttura di preferenze (ad es. il fatto che comunità più ricche plausibilmente valutano maggiormente taluni beni ambientali, normalmente ritenuti beni di lusso, o che comunità situate in contesti economici diversi possono attribuire un peso diverso alle ricadute sull'occupazione derivanti dalla presenza dell'infrastruttura) o circostanze particolari della localizzazione dell'infrastruttura in un certo territorio (si pensi all'unicità di ciascun paesaggio e alla soggettività della valutazione dello stesso). La struttura di preferenze generalmente rappresenta un'informazione privata di ciascuna comunità.

Questa caratteristica rende complesso il compito di un decisore pubblico che si ponga l'obiettivo di massimizzare i benefici complessivi del progetto al netto di tutti i costi coinvolti, quelli di costruzione specifici a ogni sito, ma anche le disutilità che la comunità ospitante sopporta. Laddove la localizzazione sia predeterminata, la difficoltà consiste nel definire un'eventuale "compensazione" dei costi che la comunità

⁷ A variare tra localizzazioni potrebbe essere anche la produttività futura dell'impianto. Una localizzazione potrebbe essere preferibile da un punto di vista tecnologico perché l'infrastruttura in quel sito potrebbe essere più efficace o meno costosa da gestire e mantenere.

⁸ Anche per quanto riguarda i termini costruttore/gestore useremo entrambi in maniera intercambiabile. Nel seguito useremo anche in maniera intercambiabile i termini costi di costruzione e costi di costruzione e gestione, salvo laddove debba essere esplicita una distinzione tra problematiche connesse con la costruzione e con la gestione dell'infrastruttura (tale distinzione potrebbe ad esempio essere rilevante laddove le condizioni connesse con la successiva gestione dell'impianto non vengano pienamente internalizzate dal costruttore o proponente dell'impianto).

ospitante deve sopportare e che sarebbe opportuno venissero “internalizzati” da chi il progetto propugna e porta avanti. Nel caso in cui si possa scegliere tra più possibili localizzazioni, diventa difficile individuare quella con costi minori.

4. - Come si può gestire il *Nimby*

Non rientra tra gli obiettivi di questo lavoro analizzare le caratteristiche e le proprietà dei meccanismi ottimi di localizzazione delle infrastrutture soggette a *Nimby*. Ai nostri fini preme soltanto concentrarci su due aspetti del problema, evidenziando in proposito l'importanza delle variabili già prima introdotte: differenze tra siti potenziali nei costi di costruzione dell'infrastruttura ed eterogeneità nella struttura di preferenze tra comunità potenzialmente interessate alla localizzazione.

Il primo aspetto su cui ci concentriamo concerne la scelta del momento e delle modalità con cui viene attivato il confronto con le comunità interessate (e conseguentemente dei soggetti che sono coinvolti). Al riguardo è possibile distinguere due “approcci” cui di seguito faremo riferimento come approccio *ex post* e approccio *ex ante*. Il primo segue una logica di tipo *top-down*: la sequenza degli eventi vede il gestore (o il decisore pubblico) adottare la scelta localizzativa sulla base di valutazioni tecniche e formulare il progetto dettagliato dell'opera senza informare o coinvolgere preventivamente la popolazione o gli enti locali interessati. Il confronto avviene successivamente e solo con *la* comunità prescelta; riguarderà forme di compensazioni e varianti finalizzate a ridurre o annullare le disutilità dell'opera, massimizzandone l'accettazione. L'approccio *ex ante* si basa su una logica di tipo *bottom-up*: tutte *le* comunità interessate (a seconda dei contesti istituzionali, solo gli enti locali ovvero tutti gli *stakeholders*: enti locali, cittadini, associazioni, ecc.) vengono coinvolte in una fase preliminare in cui sono discusse le scelte principali dell'intervento infrastrutturale (localizzazione, eventuali varianti e misure compensative in senso proprio).

Il secondo aspetto su cui vorremmo focalizzare l'attenzione riguarda la scelta, nel caso in cui l'approccio utilizzato sia di tipo *ex ante*, di un meccanismo più o meno competitivo per l'individuazione della localizzazione e delle misure di compensazione.

4.1 Ex post vs ex ante

La desiderabilità di un approccio di tipo *ex post* o *ex ante* è influenzata dalla variabilità relativa delle due tipologie di costi – costi di costruzione dell'infrastruttura e disutilità che gravano sulla comunità ospitante – tra le diverse possibili lo-

calizzazioni. Laddove i primi presentano un'elevata eterogeneità, relativamente all'eterogeneità nelle disutilità percepite dalle diverse comunità, la localizzazione ottimale sarà *ceteris paribus* più facilmente perseguita affidando la scelta al gestore dell'opera. In questo caso, infatti, per plausibili differenze nelle preferenze tra le diverse comunità, la scelta della localizzazione ottima sarà "dominata" dalle considerazioni sui costi di costruzione e gestione dell'impianto, su cui è il gestore il soggetto che possiede le maggiori informazioni.

L'individuazione preventiva della localizzazione ottima non esaurisce però il problema. Le criticità della successiva contrattazione, mirante a far accettare l'opera da parte della comunità prescelta in cambio di una compensazione monetaria esplicita o di qualche variante che innalzi i costi per il gestore ma attenui le disutilità per la comunità ospitante, vengono anzi semmai esacerbate. In particolare, il fatto che la contrattazione avvenga dopo che la scelta localizzativa è stata effettuata pone il gestore in una condizione di "lock in": avendo formulato un progetto e sostenuto i relativi costi egli sopporterebbe una perdita elevata qualora decidesse di abbandonarlo, di modificarne la localizzazione o le caratteristiche essenziali. Questa condizione conferisce un elevato potere contrattuale alla comunità prescelta che, nel confronto bilaterale *ex post*, avrà la possibilità di chiedere compensazioni che riflettono non solo le disutilità da esse effettivamente sopportate ma anche il costo-opportunità sostenuto dal gestore qualora decidesse di abbandonare il progetto. L'esito potrebbe essere il riconoscimento di compensazioni troppo elevate o la scelta del gestore di non intraprendere l'investimento: le infrastrutture finirebbero col costare "troppo" e col farsi meno spesso del necessario.

In tale ambito assume particolare rilevanza il contesto istituzionale (e politico) all'interno del quale la contrattazione ha luogo. La capacità della comunità locale di estrarre rendita e ottenere compensazioni elevate è, infatti, maggiore laddove il suo consenso sia un elemento necessario. In un ipotetico regime in cui la scelta localizzativa (del gestore o di un'autorità pubblica sovraordinata) non sia contestabile, è evidente che il rischio di *lock in* non esiste. Il contesto istituzionale che rileva non è però solo quello connesso con la formale attribuzione del potere di scelta localizzativa. In un sistema fortemente decentrato, in cui ampi poteri e prerogative in tema di governo del territorio sono riconosciute agli enti locali, la singola comunità locale può comunque avere ampi margini per ostacolare il processo di effettiva costruzione dell'infrastruttura. Altri aspetti del contesto istituzionale che contano sono inoltre quelli attinenti la previsione di tempi e modalità pre-stabiliti per l'effettuazione del confronto *ex post*. Fra gli aspetti del contesto (politico o istituzionale) rilevanti vi è infine anche la più complessiva interazione tra

livelli di governo, perché i rischi di *lock in* sono ridotti laddove il livello statale comunque svolge una funzione di indirizzo e di incentivo all'internalizzazione delle esternalità nei confronti del livello locale⁹.

4.2 Ex ante: *aste e approccio contrattuale*

Nel caso opposto, in cui l'eterogeneità delle preferenze tra le comunità potenzialmente ospitanti risulti elevata rispetto a quella dei costi di costruzione dell'impianto, la localizzazione ottima viene a dipendere soprattutto dalle preferenze delle diverse comunità. Essenziale diviene il definire una procedura che meglio conduca al loro corretto disvelamento. Nelle sue linee generali, il problema economico di gestione del *Nimby* assume caratteri simili a quelli che caratterizzano la fornitura di beni pubblici: la presenza di asimmetria informativa tra il gestore (decisore pubblico) e le comunità potenzialmente deputate a ospitare l'infrastruttura circa le preferenze di queste ultime; la difficoltà, in assenza di meccanismi opportuni, di indurre una corretta rivelazione. Ciascuna comunità ha infatti un interesse a "esagerare" la portata delle proprie disutilità, perché così facendo può sperare di evitare di essere la localizzazione prescelta, sopportandone i costi connessi, o comunque di ottenere compensazioni più elevate.

La letteratura economica normalmente affronta questo problema, in termini tanto descrittivi quanto prescrittivi, cercando di individuare meccanismi in grado di indurre le comunità interessate a rivelare correttamente le loro preferenze. In generale i meccanismi individuati condividono alcuni profili del meccanismo Vickrey-Clarke-Groves (VCG) di rivelazione delle preferenze per i beni pubblici¹⁰. Alcuni dei principali meccanismi proposti sono Kunreuther e Kleindorfer (1986), O' Sullivan (1993), Perez-Castrillo e Wettstein (2002)¹¹, tutti relativi al caso relativamente più semplice in cui si possano definire solo compensazioni monetarie in attenuazione delle disutilità attribuite da ciascuna delle comunità potenzialmente ospitanti l'infrastruttura. Kunreuther e Kleindorfer (1986) propongono un'asta in busta chiusa

⁹ Sotto questo punto di vista, il venir meno dei grandi partiti nazionali su base ideologica e le modifiche al regime elettorale nelle amministrazioni locali (anche se in questo caso va detto che nei Comuni è stato anche introdotto un vincolo al numero di mandati dei sindaci, che dovrebbe portare ad una asimmetria tra primo e secondo mandato) potrebbero aver aggravato le difficoltà rispetto a 20 anni or sono. Per quanto argomenti simili sono stati a volte sostenuti nel dibattito in Italia non siamo però in grado di fornire alcuna evidenza empirica precisa in proposito.

¹⁰ CLARKE E. (1971); GROVES T. (1973); VICKREY W. (1961).

¹¹ Si vedano anche MINEHART D. e NEEMAN Z. (2002); LAURENT-LUCCHETTI J. e LEROUX J. (2009).

in cui ciascuna comunità dichiara quanto vorrebbe ricevere per ospitare l'infrastruttura. La comunità che fa l'offerta più bassa vince e riceve la propria offerta sotto forma di compensazione. Le altre comunità pagano un'imposta pari alla propria offerta diviso il numero delle comunità meno uno. Se le comunità adottano una strategia *maximin* (la comunità *i* sceglie l'azione migliore assumendo che la comunità *j* scelga l'azione che più danneggia *i*) l'equilibrio è efficiente. O' Sullivan (1993) propone un'asta in busta chiusa in cui la comunità che sottomette l'offerta più bassa ospita l'infrastruttura e riceve l'offerta più alta sotto forma di compensazione. L'equilibrio di Nash del gioco assicura una soluzione efficiente. Infine, Perez-Castrillo e Wettstein (2002) propongono un'asta a offerte multiple: ciascuna comunità sottopone un vettore di offerte – una per ciascuna delle comunità potenzialmente ospitanti incluso la propria – che rappresentano la disponibilità a pagare affinché l'infrastruttura sia localizzata in quella comunità. È imposto il vincolo che la somma delle offerte sottomesse sia pari a zero. In aggiunta ciascuna comunità annuncia la sua localizzazione preferita. La comunità che riceve la più elevata offerta complessiva è selezionata e ciascuna comunità paga il relativo trasferimento. In caso di parità, è selezionata una comunità a caso tra quelle che sono state scelte come preferite da almeno uno dei partecipanti. Un equilibrio di Nash del gioco esiste sempre e implementa l'allocazione efficiente. Il meccanismo assicura pareggio del bilancio.

Nonostante, in generale, questi meccanismi assicurino in teoria decisioni efficienti – la comunità selezionata per ospitare l'infrastruttura è quella con le disutilità più basse – nella pratica essi sono scarsamente utilizzati, mentre tendono a prevalere approcci contrattuali. Di seguito cercheremo di individuare alcuni fattori che possono razionalizzare l'impiego limitato di meccanismi d'asta nel contesto qui considerato.

Costi privati e comuni. – Un'ipotesi comune a tutti i meccanismi d'asta descritti è che le comunità partecipanti all'asta siano preventivamente in grado di attribuire un costo alle disutilità derivanti dalla localizzazione dell'infrastruttura nel proprio territorio. In pratica una simile valutazione dipende, oltre che da una componente “privata”, nota alla comunità, che riflette la sua struttura di preferenze¹², anche

¹² La valutazione sulla base di informazioni “private” può anche essere difficile perché talune cose possono essere ritenute valori “incommensurabili” (un paesaggio che non ha prezzo!). Nella logica delle scelte comunque da fare, ciò può però significare attribuire un valore molto elevato a certe disutilità (nell'esempio prima fatto il perdere un certo paesaggio) o anche nel seguire una logica lessicografica nelle proprie valutazioni (ad esempio escludere comunque certi impianti che superino una certa soglia di rischio sì da ridefinire le caratteristiche “ammissibili” di certe infrastrutture), senza che ciò modifica la natura logica del problema qui esaminato.

da una serie di caratteristiche tecniche proprie dell'impianto. Questa determinante del costo è "comune" a tutte le localizzazioni; su di essa ciascuna comunità possiede però una sua "credenza privata", che è funzione di ciò che essa sa circa le caratteristiche effettive dell'infrastruttura. In presenza di questa componente comune l'asta cessa di essere una soluzione efficiente: aggiudicataria dell'asta potrebbe essere non la comunità coi costi di localizzazione più bassi ma quella più "ottimista". L'inefficienza è tanto più elevata quanto maggiore è l'incertezza relativa alla componente comune; quando questa diventa così elevata da rendere ininfluenti le preferenze private circa le disutilità arrecate dalla infrastruttura, l'allocatione diviene del tutto casuale. In altre parole, l'asta non riesce più a svolgere la sua funzione di meccanismo di estrazione delle preferenze private¹³. Più in generale, anche laddove le comunità non differiscano nelle loro credenze private circa le caratteristiche obiettive dell'infrastruttura, è plausibile ritenere che tutte saranno indotte a innalzare le compensazioni monetarie richieste – innalzando il costo che il costruttore deve sopportare per realizzare l'infrastruttura, sì da rendere la realizzazione di questa meno probabile - se avverse all'ambiguità insita nel non conoscere con certezza le caratteristiche oggettive dell'infrastruttura e la natura dei rischi a cui si può andare incontro¹⁴.

Tali inefficienze possono essere mitigate dal gestore, che meglio conosce i rischi insiti nell'infrastruttura e ha un interesse a fornire informazioni che consentano alle comunità di precisare la loro valutazione dei costi della localizzazione¹⁵. L'effettiva capacità di trasmettere tali informazioni in modo credibile trova tuttavia un ostacolo nella loro natura non facilmente verificabile e nell'interesse del gestore a diffondere informazioni "ottimistiche". In tale contesto la comunicazione può avvenire solo attraverso un processo graduale di acquisizione di fiducia reciproca

¹³ Cfr. GOEREE J. K. e OFFERMAN T. (2003) per un'analisi del funzionamento dei meccanismi d'asta in presenza di "*private and common values*".

¹⁴ L'avversione all'ambiguità è cosa diversa dall'avversione al rischio (cfr. sul punto in questione GHIRARDATO P. e MARINACCI M., 2002). Se un certo impianto comporta certi rischi, ma le diverse comunità hanno una diversa avversione al rischio, è ottimale che la comunità meno avversa al rischio sia quella (*ceteris paribus*) prescelta per la localizzazione. L'avversione all'ambiguità attiene al fatto che il rischio dell'impianto non è, almeno non agli occhi delle comunità interessate, calcolabile, il che induce a cautelarsi avverso lo stesso, richiedendo compensazioni più elevate.

¹⁵ L'argomento è speculare a quello appena esposto: l'acquisizione di informazioni più precise circa i rischi insiti nell'infrastruttura determina una riduzione nelle richieste di compensazioni.

e uno sforzo di assimilazione e chiarificazione che è più proprio di una relazione di tipo negoziale¹⁶.

Compensazioni non monetarie. – La letteratura sui meccanismi d’asta prima sinteticamente descritta assume che le preferenze delle diverse comunità siano sintetizzabili in un equivalente monetario. In altre parole, si ipotizza che vi sia un’unica possibile variante del progetto d’infrastruttura e che rispetto a questa le diverse comunità siano chiamate a dichiarare l’ammontare di compenso monetario che desiderano ricevere per essere disposte a ospitare l’infrastruttura (o che sono disposte a pagare pur di non ospitarla). In pratica le compensazioni assumono più spesso la forma di servizi finalizzati e opere suppletive, ovvero di modifiche alle caratteristiche del progetto definite al fine di attenuarne la disutilità percepita. La diffusione di queste forme di compensazione può essere spiegata con la loro maggiore efficienza: si tratta spesso di interventi il cui valore per la comunità è superiore al costo per il gestore di realizzarle. Laddove un’asta è per sua natura più adatta a gestire compensazioni monetarie in senso proprio, l’identificazione delle misure di attenuazione/adeguamento delle caratteristiche di un’infrastruttura o dei benefici suppletivi da fornire in natura a una certa comunità è più consona a un meccanismo contrattuale e di confronto tra comunità e soggetto proponente l’infrastruttura¹⁷.

Chi partecipa all’asta? – Finora si è parlato di costi di localizzazione percepiti da una comunità nel suo insieme, laddove evidentemente valutazioni e decisioni devono esser fatte aggregando le opinioni di più soggetti. Ad esempio, all’interno della stessa comunità possono coesistere interessi e preferenze divergenti e tra loro in conflitto; le disutilità generate dall’impianto possono ricadere su più comunità pur essendo l’impianto materialmente localizzato all’interno di una sola di esse¹⁸;

¹⁶ Cfr. DEWATRIPONT M. e TIROLE J. (2005) per un’analisi economica della comunicazione in un contesto caratterizzato da conflitto di interessi e natura “soft” dell’informazione. Il punto evidenziato dagli autori è che la comunicazione è soggetta a “moral hazard in teams” e quindi può aversi trasmissione credibile di informazioni soltanto quando simultaneamente il “sender” e il “receiver” compiono uno sforzo, rispettivamente, di chiarificazione e assimilazione.

¹⁷ Un argomento simile è stato utilizzato nel contesto del *procurement* (cfr. BAJARI P., MCMILLAN R. e TADELIS S., 2008) per dimostrare la preferibilità di procedure negoziali di selezione del contraente in presenza di appalti complessi in cui la stazione appaltante ha difficoltà nel definire nel contratto iniziale tutte le caratteristiche rilevanti del lavoro da svolgere.

¹⁸ Ad esempio il Rapporto dell’Osservatorio *NimbyForum* del 2009 documenta la presenza di siti a beneficio della popolazione di un comune e situati al confine di un altro, che ne percepisce quasi solo i costi di localizzazione, il che ne fomenta l’opposizione nonostante l’assenza di poteri formali di intervento in capo allo stesso.

se la scelta della localizzazione coinvolge comuni appartenenti a regioni diverse, il consenso su una determinata localizzazione può richiedere il parere positivo di livelli diversi di governo. Il funzionamento di un meccanismo d'asta presuppone che questi diversi interessi vengano preventivamente aggregati in modo che per ciascuna possibile localizzazione possa essere individuato un unico interlocutore le cui preferenze siano espressione e mettano insieme le preferenze di tutti gli *stakeholders* interessati. La difficoltà di compiere tale processo, e di compierlo prima che l'asta vera e propria abbia luogo, può spiegare la maggiore efficacia di un meccanismo contrattuale, che non richiede che tutti gli interessi siano preventivamente aggregati. La loro ricomposizione è affidata a un processo di negoziazione multilaterale, che può essere più o meno regolamentato e istituzionalizzato, nell'ambito del quale il gestore o il decisore pubblico assume la funzione di mediatore e il compito di individuare la soluzione in grado di ottenere il consenso di tutte le parti coinvolte¹⁹.

I limiti dei meccanismi d'asta qui evidenziati non vanno interpretati nel senso di escludere che vi possano essere situazioni in cui essi possono svolgere una funzione utile, ma piuttosto di indicare alcune "condizioni" che devono essere soddisfatte perché l'asta sia un meccanismo efficiente. In primo luogo, al fine di evitare che l'incertezza sui rischi insiti nell'infrastruttura ne comprometta l'efficacia, il ricorso all'asta richiede la presenza di un soggetto che sia in grado di generare informazioni credibili, o di certificare quelle fornite dal gestore, in modo da consentire che tali informazioni possano definire, senza ambiguità, i valori posti a base d'asta. Questo sposta il problema economico su un piano diverso che concerne la scelta dei criteri sulla base dei quali tale soggetto è selezionato e i meccanismi attraverso i quali se ne costruisce una reputazione di competenza e credibilità. In secondo luogo, il ricorso all'asta è maggiormente proficuo nei casi in cui, per caratteristiche delle opere o funzionamento dei meccanismi istituzionali, l'aggregazione preventiva delle preferenze sia più facilmente realizzabile (l'impianto non genera esternalità negative sui comuni limitrofi, l'opera è di natura puntuale e non riguarda un intero tracciato con scelte localizzative da realizzare in maniera sequenziale, vi è un basso rischio che possano emergere minoranze vocifere) ovvero laddove non vi siano significativi benefici nel coinvolgimento più diretto delle comunità al fine di definire possibili varianti dell'opera apprezzate

¹⁹ Più in generale, la scelta delle comunità ammesse a partecipare all'asta è un aspetto rilevante e di non facile determinazione in considerazione delle distorsioni nei comportamenti che l'inclusione o l'esclusione di una di esse può determinare.

dalla comunità e di costo limitato per il costruttore. Sintetizzando i vari elementi discussi, possiamo quindi dire che il modello negoziale *ex ante* è sostanzialmente da preferire all'asta laddove le caratteristiche dell'opera siano relativamente *fuzzy* e non chiaramente, e credibilmente, definite *a priori* e sintetizzabili in una semplice metrica monetaria.

D'altra parte, proprio perché connessi a situazioni più *fuzzy*, i processi di tipo negoziale si caratterizzano per avere costi di transazione più elevati, legati alla necessità di discutere con più soggetti (più comunità, più *stakeholders*) e di tenere conto delle sollecitazioni e delle proposte di modifica del progetto che provengono dai vari interlocutori. Per questa ragione, anche laddove superiori a quelli ottenibili con un meccanismo d'asta, i benefici di un approccio negoziale possono risultare fortemente sminuiti in presenza di un contesto istituzionale caratterizzato da elevata frammentazione degli interessi e dei centri decisionali, dal sovrapporsi di competenze sullo stesso ambito geografico di più livelli di governo (comune, provincia, regione etc.) e da una non chiara attribuzione delle competenze. L'eccessivo allungamento dei tempi che può derivarne può avere ricadute negative sulla stessa probabilità di successo della concertazione. In particolare può accadere che gli enti locali, magari a seguito di mutamenti nel quadro politico generale, rinneghino le intese di massima nei successivi passi del processo. Il problema del "*commitment*" politico diviene maggiormente severo in presenza di forte instabilità politica, elevata conflittualità e frequenti cambi di maggioranze.

Le criticità ora ricordate in merito ai diversi meccanismi di gestione del *Nimby* comportano che, in modo non direttamente specifico a nessun di essi, condizioni politiche e di contesto divengono rilevanti nello spiegare l'insorgere delle proteste e l'intensità e criticità delle stesse. In particolare, sul fenomeno *Nimby* possono incidere i costi di agenzia legati al gioco politico tra rappresentanti e cittadini e la possibilità che le motivazioni dei primi possano divergere da quelle dei cittadini. Può impattare il ciclo politico-elettorale, per cui rappresentanti prossimi ad una scadenza elettorale possono essere portati ad anteporre gli interessi della loro *constituency* rispetto a quelli generali o a distorcere i loro comportamenti con l'obiettivo di evidenziare la propria *toughness* a fini elettorali; possono impattare gli incentivi di carriera degli amministratori locali che possono trovare conveniente uniformarsi alle posizioni adottate dal partito di appartenenza a livello centrale – generalmente più attente a perseguire gli interessi nazionali – al fine di costruirsi una reputazione di fedeltà.

5. - Il modello procedimentale italiano

In Italia la costruzione (e l'ampliamento) di un impianto presuppongono un *iter* autorizzativo particolarmente complesso che coinvolge in varia misura una pluralità di amministrazioni pubbliche²⁰. Esso è scandito in due fasi distinte: la prima termina con l'adozione della scelta sulla localizzazione dell'opera; la seconda con il rilascio dei provvedimenti indispensabili a consentire la costruzione e il funzionamento dell'impianto.

Tradizionalmente, la scelta localizzativa vera e propria è adottata a seguito di una negoziazione o un confronto che coinvolge esclusivamente i soggetti chiamati a rilasciare una serie di atti di assenso, ovvero: *a*) tutti i livelli di governo afferenti sul territorio direttamente interessato; *b*) le amministrazioni settoriali preposte alla cura di interessi sensibili (es. tutela della salute, dell'ambiente, dei beni culturali e dell'incolumità pubblica); *c*) i gestori di opere interferenti.

In linea generale, non tutte le diverse istanze hanno eguale voce in capitolo. In alcune tipologie di progetti (come le infrastrutture di interesse concorrente statale e regionale e quelle energetiche) il consenso della regione/provincia autonoma è indispensabile per il raggiungimento di un accordo. Nelle altre il loro dissenso può essere formalmente superato rinviando la determinazione conclusiva del procedimento localizzativo al Consiglio dei Ministri. In generale, le amministrazioni settoriali preposte alla cura di interessi sensibili non sono titolari singolarmente di alcun potere di veto; anche in questo caso eventuali conflitti inter-istituzionali sono superabili ricorrendo al Consiglio dei Ministri.

Non sono invece previsti, non almeno in maniera formalizzata e tanto meno obbligatoria, meccanismi istituzionalizzati di informazione, coinvolgimento e ascolto dei cittadini in quanto tali né durante né a valle della fase localizzativa (ad es. nella eventuale definizione di varianti al progetto finalizzate ad attuare le possibili disutilità dell'impianto, in quello che abbiamo definito come meccanismo *ex ante*, cfr. par. precedente). I cittadini possono opporsi alla scelta localizzativa solo inviando – nell'ambito del sub-procedimento di Valutazione di Impatto Ambientale (VIA) – osservazioni scritte, che però non vincolano né influenzano la

²⁰ La legislazione italiana prevede diversi *iter* procedimentali: ordinari per opere di interesse nazionale; ordinari per opere di interesse statale e infrastrutture strategiche; speciali per infrastrutture energetiche e di telecomunicazione; emergenziali o per commissari per opere "selezionate". In quanto segue ci si concentrerà sulle logiche generali dei diversi procedimenti e sulla evoluzione più recente dell'*iter* decisorio. Su caratteristiche e criticità dei singoli procedimenti localizzativi si rinvia a IAIONE C. (2010); ISAE (2008); ITALIADECIDE (2009); LABOCETTA A. M. (2008); MERLONI F. (2007).

decisione finale; tra l'altro, la ristrettezza dei tempi entro i quali le osservazioni possono essere inviate e il non agile accesso alla procedura riducono ulteriormente la capacità dei cittadini di manifestare la propria opposizione per via scritta²¹. A differenza di quanto ad esempio accade in Francia e nel Regno Unito, l'Italia non prevede forme obbligatorie di consultazione e dialogo diretto con le comunità locali sotto forma di dibattiti pubblici, *public hearings* e *inquires*. L'inchiesta pubblica, che trova disciplina nel Testo unico dell'ambiente, non è utilizzata²².

A valle della scelta localizzativa, ai cittadini, così come agli stessi enti locali, non è certo precluso il ricorso all'autorità giudiziaria, l'insorgere di forme di contenzioso essendo anzi favorito dalla complessità delle procedure. È inoltre ovvio che, pur se formalmente e direttamente non consultati, i cittadini possono trovare ascolto, collettivamente, nelle scelte politiche degli enti locali.

Tornando alle interazioni tra i diversi soggetti pubblici, già in sede di scelta localizzativa va rilevato come le procedure che consentirebbero di superare i dissensi espressi da talune amministrazioni non siano poi di fatto adoperate. Il Con-

²¹ Cfr. CORSETTI A., FERRARA R., FRACCHIA F. e OLIVETTI RASON N. (2008).

²² L'inchiesta pubblica è un peculiare modulo procedimentale utilizzato in settori in cui l'attività istruttoria appare particolarmente complessa a causa sia della molteplicità degli *stakeholders* coinvolti sia degli interessi sensibili toccati. Se utilizzata durante la fase di localizzazione di opere pubbliche essa consente di coinvolgere attivamente i vari attori interessati (inclusi i cittadini) e di analizzarne a valutarne contestualmente le posizioni espresse. In Italia l'inchiesta pubblica può essere convocata dall'autorità tenuta a emettere il decreto di compatibilità ambientale nell'ambito della VIA. Su richiesta dell'autorità o su iniziativa del soggetto proponente, può essere sostituita dallo svolgimento di un contraddittorio tra quest'ultimo e i soggetti pubblici e privati che hanno presentato pareri e osservazioni. Cfr. TONETTI A. (2010); CASINI L. (2007, 2010). L'inchiesta pubblica prevista in Italia si caratterizza per una disciplina alquanto scarna che la allontana dagli strumenti partecipativi previsti in Francia e nel Regno Unito. A differenza di quanto accade in sede nazionale, la regione Toscana, tradizionalmente sensibile al tema della partecipazione popolare ai procedimenti di interesse collettivo e ispirandosi ampiamente al modello partecipativo francese, con la l. r. n. 69 del 2007 ha introdotto il dibattito pubblico regionale come forma di confronto e dialogo in merito alla localizzazione e alla realizzazione di opere pubbliche con forte impatto ambientale, territoriale e sociale per la comunità regionale. Il dibattito si articola in una procedura particolarmente scandita e condotta da un'autorità amministrativa indipendente all'uopo costituita (l'Autorità regionale per la garanzia e la promozione della partecipazione). Il suo svolgimento avviene nella fase preliminare alla scelta localizzativa e consente un ampio coinvolgimento di vari attori istituzionali e sociali oltre che dei singoli cittadini. L'Autorità valuta e ammette le proposte di dibattito pubblico (che possono essere avanzate dal soggetto proponente, dal soggetto realizzatore, dagli enti locali coinvolti o da un determinato numero di cittadini) così come può promuovere progetti partecipativi diversi. Al termine del dibattito, il soggetto titolare dell'intervento può discostarsi dalle conclusioni a cui il dibattito è giunto solo argomentando tale scelta.

siglio dei Ministri generalmente non adotta un provvedimento di contenuto contrario alla posizione espressa da suoi singoli membri o dalla regione coinvolta. Le ragioni sono essenzialmente politiche: la prassi è quella di cercare comunque un'intesa, magari allargando la trattativa anche ad altri possibili tavoli di scambio; la politicizzazione della decisione d'altro canto fa sì che i contrasti che inevitabilmente sorgerebbero tra istituzioni diverse acquistino un *flavour* tutto politico e non più tecnico-istituzionale. E' possibile che vi influisca la natura comunque eterogenea delle coalizioni che hanno governato l'Italia e la debolezza tecnica delle amministrazioni coinvolte, i cui eventuali contrasti divengono per l'appunto una questione politica e non meramente tecnica.

Soprattutto, ulteriori ostacoli procedurali sorgono nella fase realizzativa. Assunta la scelta localizzativa, l'avvio effettivo del progetto infrastrutturale è infatti soggetto al rilascio di una serie di altri atti amministrativi di competenza degli enti locali minori e delle amministrazioni preposte alla cura di interessi sensibili. Tali atti hanno da accertare la compatibilità del progetto a piani paesaggistici, strumenti urbanistici e altri vincoli concernenti principalmente la tutela dell'ambiente, della salute e del paesaggio²³. In questa fase vengono altresì negoziati interventi di compensazione e mitigazione tra soggetto proponente e (soprattutto) enti locali minori per massimizzare l'accettazione dell'opera. Anche stante il fatto che la scelta localizzativa è ormai intervenuta (nella logica di quello che nel paragrafo precedente abbiamo definito come approccio *ex post*), questo può essere ulteriore causa di *impasse*.

Il potere che in questa fase (soprattutto) gli enti locali minori – con meno voce in capitolo nella fase precedente di scelta localizzativa vera e propria – sono in grado di sfruttare deriva dalla possibilità che essi di fatto hanno di rallentare la realizzazione e il funzionamento dell'opera ritardando il rilascio di una serie di atti amministrativi come il permesso di costruzione, le concessioni edilizie, le autorizzazioni per opere connesse e di altri provvedimenti previsti in materia urbanistica e paesaggistica²⁴. Non è, poi, escluso che gli stessi enti provvedano al ritiro

²³ Molti di questi adempimenti rientrano nella competenza legislativa regionale e nella potestà amministrativa comunale.

²⁴ Trattandosi di atti autorizzativi che involgono interessi pubblici ritenuti meritevoli di una tutela rafforzata o che devono essere rilasciati in presenza di vincoli ambientali, paesaggistici e culturali o, ancora, che rientrano nell'ambito dell'esercizio dell'attività discrezionale della p.a., tali provvedimenti non rientrano nella disciplina della Segnalazione certificata di inizio attività (SCIA) recentemente introdotta allo scopo di facilitare, velocizzare e rendere meno costoso l'avvio e l'esercizio dell'attività di impresa.

di atti già rilasciati attraverso il ricorso all'esercizio dei poteri di autotutela²⁵.

La complessità delle procedure e la stratificazione normativa da un lato, gli ampi margini di discrezionalità riconosciuti alle amministrazioni competenti al rilascio degli atti amministrativi dall'altro, favoriscono inoltre l'insorgere di contenzioso amministrativo da parte delle imprese, degli enti locali minori e dei cittadini (nonché delle loro associazioni), contenzioso che rallenta ulteriormente la realizzazione dell'opera e contribuisce a innalzarne i costi. Inoltre, soprattutto negli ultimi anni, la complessità e la disomogeneità del quadro normativo, derivanti dall'intreccio di competenze statali e regionali sul punto e da numerose incertezze applicative connesse con il nuovo Titolo V della Costituzione, hanno anche favorito il ricorso alla Corte costituzionale.

Alcune indicazioni sull'importanza del contenzioso come fattore di stallo degli investimenti infrastrutturali sono desumibili dall'ultimo Rapporto pubblicato dal Dipartimento per l'Energia del Ministero dello Sviluppo economico sull'andamento delle autorizzazioni concernenti la realizzazione o il potenziamento di centrali termoelettriche di potenza superiore a 300 MW termici²⁶. Il Rapporto segnala che dal 2002 a oggi sono stati autorizzati 45 progetti mentre 46 sono in corso di valutazione. Complessivamente, i casi che hanno registrato l'insorgere di contenzioso sono 15 per gli impianti realizzati e 11 nel caso di quelli ancora in corso di valutazione. La differenza è peraltro plausibilmente legata al fatto che nel caso di progetti ancora in corso di valutazione le occasioni di contenzioso, e gli incentivi a ricorrere avverso decisioni già prese, sono per definizione inferiori. Gli investimenti che più spesso fanno insorgere il contenzioso sono quelli relativi a impianti nuovi (14 casi su 29, nel caso degli impianti nuovi autorizzati, contro 1 su 16 nel caso di modifiche autorizzate a impianti già esistenti). Anche se è impossibile stabilire un rapporto di causalità, l'insorgere di contenzioso è chiaramente correlato con il fatto che l'investimento autorizzato sia poi stato effettivamente realizzato²⁷: per gli impianti nuovi, si hanno 6 casi di contenzioso tra i 21 investimenti realizzati e ben 4 casi di contenzioso negli 8 non ancora realizzati; per le modifiche autorizzate a impianti esistenti, nessuno degli investimenti realizzati ha avuto contenzioso, mentre in un caso su 3 tra gli investimenti non ancora realizzati si ha contenzioso.

²⁵ Ciò avviene nell'ipotesi in cui reputino che siano sopraggiunti motivi di interesse pubblico, in caso di mutamento della situazione di fatto e di nuova valutazione dell'interesse pubblico.

²⁶ Si tratta del procedimento speciale ordinario per le infrastrutture energetiche.

²⁷ Il contenzioso potrebbe causare la mancata realizzazione bloccando i lavori; potrebbe discendere dal fatto che il mancato completamento fa insorgere, soprattutto in questo caso ad iniziativa dell'impresa interessata, forme di contenzioso a fini di rivalsa.

TAV. 3

PROGETTI RELATIVI A CENTRALI TERMOELETTRICHE DI POTENZA
OLTRE I 300 MW TERMICI PRESENTATI A PARTIRE DAL 2002

Progetti autorizzati	
Impianti nuovi	29
<i>di cui realizzati</i>	21
<i>di cui con contenzioso</i>	6
(TAR: 8, CdS: 1, straordinario al Presidente della Repubblica: 1)	
<i>di cui non realizzati</i>	8
<i>di cui con contenzioso</i>	4
(TAR: 13, CdS: 8, Corte d'appello: 1)	
Modifica impianti esistenti	16
<i>di cui realizzati</i>	13
<i>di cui con contenzioso</i>	0
<i>di cui non realizzati</i>	3
<i>di cui con contenzioso</i>	1
(TAR: 13, CdS: 8, Corte d'appello: 1)	
Progetti in corso di valutazione	
Impianti nuovi	34
<i>di cui con contenzioso</i>	8
(TAR: 11, straordinario al Presidente della Repubblica: 1)	
Modifica impianti esistenti	12
<i>di cui con contenzioso</i>	3
(TAR: 5, Consiglio di Stato: 1)	

Fonte: MINISTERO DELLO SVILUPPO ECONOMICO, Dipartimento per l'Energia.

Nel quadro ora ricordato degli assetti istituzionali italiani, per molti versi tradizionale, l'ultimo decennio ha visto due novità. Da un lato, il rafforzamento dei poteri locali, e in particolare di quelli regionali (il cui consenso è necessario infatti per talune tipologie di opere; cfr. *supra*), anche a seguito del nuovo Titolo V della Costituzione, introdotto nel 2001, che ha rafforzato la valenza degli stessi, tanto in sede di scelta localizzativa in senso stretto quanto, più in generale, in materia di governo del territorio. Dall'altro, la tendenza del legislatore nazionale a risolvere l'aporia tra assetti istituzionali generali sempre più decentrati e primato statale in sede di scelta localizzativa rafforzando quest'ultimo anche e soprattutto con strumenti "straordinari". Il ruolo degli enti locali minori è stato reso comprimibile con l'adozione di procedimenti particolari per la localizzazione delle infrastrutture e degli insediamenti strategici e di interesse nazionale²⁸. Per superare le situazioni

²⁸ Nello stesso procedimento ordinario per le opere di interesse statale, non affidate a procedure speciali di tipo commissariale, la regola maggioritaria ha sostituito quella dell'unanimità a partire dal 2009, per effetto della modifica apportata dall'art. 20, comma 10 *bis*, del d.l. n. 185 del 2008, convertito in l. n. 2 del 2009. All'atto pratico, l'area di rilevanza di questa innovazione è peraltro ridotta dal fatto che si è accresciuto il ricorso a procedimenti più semplificati come quello ordinario per le infrastrutture e gli insediamenti strategici.

di stallo si è fatto ricorso in misura crescente a procedure straordinarie commissariali. Il ricorso a tali procedure, inizialmente circoscritto a ipotesi effettivamente eccezionali (come le catastrofi naturali), si è poi progressivamente esteso a interventi connessi a opere ritenute essenziali per scongiurare crisi ambientali o infrastrutturali gravi, fino a diventare l'ordinario regime localizzativo per le c.d. opere prioritarie. Di pari passo, i poteri riconosciuti ai commissari sono aumentati con il progressivo riconoscimento di poteri straordinari ampiamente derogatori e sostitutivi delle competenze ordinariamente assegnate alle amministrazioni statali e locali, potenzialmente inclusive anche di quelle competenze attinenti la successiva fase di implementazione delle opere²⁹.

In assenza di informazioni sistematiche non siamo in grado di fornire una stima degli effetti di questo maggior ricorso alle procedure commissariali, una valutazione che dovrebbe tra l'altro tener conto della eterogeneità di situazioni specifiche, di poteri accordati e di modalità effettive di esercizio degli stessi in cui lo strumento commissariale si è concretato. In generale, lo strumento commissariale si è fatto a volte apprezzare per una maggiore rapidità; sul piano procedurale, si è in genere ottenuta una migliore scansione dei tempi procedimentali, con una unificazione delle competenze in capo a un unico organo. Tuttavia, l'ampio ricorso a tali procedure – che ha posto e pone numerosi dubbi sul piano della legittimità costituzionale – rischia sempre di sfociare in episodi di radicalizzazione dei conflitti con la popolazione e gli enti locali, completamente estromessi dal procedimento localizzativo a monte, nella successiva fase di implementazione dell'opera e ha fatto spesso registrare una riduzione della trasparenza dell'attività amministrativa.

L'indirizzo seguito dal legislatore nazionale è comunque in apparente contrasto con talune tendenze regolatorie internazionali e con una prassi anche nazionale di intensificazione, da parte di molte imprese impegnate in Italia nel comparto delle opere pubbliche, del confronto con gli attori locali.

Secondo quanto è emerso nel corso di colloqui intercorsi con alcuni tra i maggiori operatori nazionali impegnati nel settore infrastrutturale, la maggior parte delle imprese che realizzano impianti con forte impatto sulla tutela della salute, del territorio e del paesaggio cerca di “superare” il modello formale attraverso il

²⁹ Cfr. FRANCHINI C. (2009); AMOROSINO S. (2009); IAIONE C. (2010). I concreti confini di tali poteri derogatori variano a seconda della specifica fattispecie, essendo di volta in volta fissati dal singolo decreto di nomina del commissario. Potenzialmente però è evidente che si superano, almeno sulla carta, le prime descritte *chances* degli enti minori di porre “bastoni tra le ruote” in un momento successivo alla scelta localizzativa propriamente detta.

ricorso a consultazioni/concertazioni che si svolgono con gli enti locali e, in alcuni casi, con la popolazione interessata nella fase precedente all'avvio del procedimento localizzativo. L'utilizzo di tali pratiche è funzionale a massimizzare il consenso alla costruzione dell'opera a monte; secondo le imprese ciò permetterebbe da un lato, di scongiurare gli ostacoli amministrativi che gli enti locali minori possono porre durante la fase realizzativa, dall'altro, di prevenire le proteste della comunità interessata dall'intervento infrastrutturale. Generalmente si instaura un rapporto privilegiato con la regione, che svolge un ruolo di coordinamento, frapponendosi tra le imprese e gli enti locali minori³⁰. L'esigenza di instaurare forme di dialogo diretto con le comunità locali è avvertita da alcune imprese che operano in settori percepiti come particolarmente rischiosi per la salute e/o che hanno un forte impatto sul territorio e sul paesaggio (si tratta principalmente delle imprese del settore rifiuti e di quello elettrico). In tali incontri le imprese informano i cittadini su rischi e vantaggi connessi alla realizzazione dell'opera (limitando una loro strumentalizzazione da parte di comitati e associazioni che tendono ad amplificare i rischi), di recepire i motivi alla base di eventuali loro opposizioni (a tal fine sono spesso effettuati sondaggi, mappature degli *stakeholders*, analisi economiche e sociali) e di accrescere la propria reputazione³¹.

Sul piano più strettamente normativo, gli indirizzi diffusi cui si faceva riferimento riguardano tanto gli auspici comunitari³², quanto il crescente ricorso a, e la procedimentalizzazione di, forme di ascolto e confronto con la cittadinanza. Dagli anni settanta è cresciuta la diffusione di pratiche partecipative sotto forma di *public hearings* preliminari e funzionali all'adozione delle decisioni di localiz-

³⁰ Sempre durante la fase preventiva, al fine di ottenere più facilmente le autorizzazioni necessarie le imprese propongono la realizzazione contestuale di opere di mitigazione ambientale e/o accessorie che recano vantaggi diretti alla comunità locale e in generale al contesto territoriale (principalmente, interventi di razionalizzazione della rete esistente, interventi che rendono compatibile l'opera con il contesto paesaggistico e territoriale e riducono gli elementi impattanti).

³¹ La maggior parte delle imprese ha evidenziato che l'utilizzo di forme di dialogo con gli enti locali ha preso avvio a seguito dell'introduzione del decentramento amministrativo e legislativo, costituzionalizzato nel 2001 con la modifica del Titolo V. Precedentemente, le imprese utilizzavano un tipico approccio *ex post*. La scelta localizzativa era adottata sulla base della programmazione statale e la progettazione era effettuata tenendo conto di valutazioni più che altro di natura tecnica. Gli enti locali e le comunità interessate erano del tutto esclusi dal procedimento di localizzazione; per limitare loro opposizioni erano utilizzate compensazioni sotto forma di prospettive di crescita occupazionale ed economica. Per opporsi alle scelte adottate a livello centrale non restava che il ricorso all'autorità giudiziaria.

³² Si vedano in particolare la Convenzione di Aarhus e la legislazione comunitaria in materia ambientale.

zazione e progettazione di opere pubbliche. In Francia e Regno Unito, in particolare, l'inchiesta pubblica, che rappresenta il principale strumento istruttorio per la localizzazione e realizzazione di infrastrutture, trova una disciplina compiuta e organica³³.

6. - Conclusioni

Dalla trattazione del problema sinora svolta, è evidente come il fenomeno *Nimby* non sia una semplice aberrazione patologica. Pur se a volte fomentata dai più diversi fattori (la polemica e la competizione politica, la complessità dell'*iter* istituzionale e procedurale di decisione, i vantaggi non sempre ben dimostrati dell'opera in quanto tale, ecc...), l'emergere di un'opposizione alla localizzazione di una infrastruttura nasconde un problema oggettivo, legato al fatto che, a fronte di benefici diffusi, costi e disutilità si concentrano sulla comunità ospitante.

Il quesito che occorre porsi è perciò quello circa la adeguatezza delle modalità con cui le istanze sottostanti il *Nimby* vengono gestite al fine di pervenire a decisioni ottimali di investimento infrastrutturale e di sua localizzazione. Gestire bene il *Nimby* dovrebbe significare definire le regole d'un processo decisionale mirante a scegliere la localizzazione che garantisce che i benefici complessivi del progetto al netto di tutti i costi coinvolti – quelli di costruzione specifici a ogni sito, ma anche le disutilità che la comunità ospitante sopporta – siano massimi.

Il lavoro evidenzia come l'individuazione del meccanismo ottimale di gestione del *Nimby*, e in particolare dei tempi e delle modalità di coinvolgimento delle comunità interessate, è influenzata dalle caratteristiche concrete dell'opera e dal contesto istituzionale. Ove i costi di costruzione e gestione dell'opera differiscano poco tra i diversi siti relativamente alle differenze nelle disutilità percepite dalle diverse comunità, è opportuno che queste siano ascoltate e indotte a un veritiero disvelamento delle loro preferenze. A tale scopo, i meccanismi d'asta, che pure la teoria individuerrebbe come quelli più efficienti, non sempre rappresentano la soluzione ottimale come suggerito anche dal fatto che in pratica essi sono poco adoperati. A limitare l'efficienza dei meccanismi d'asta, e far prevalere approcci che prevedono una negoziazione multilaterale e sequenziale con tutte le parti interessate, contribuiscono diversi fattori. La difficoltà per le comunità di acquisire informazioni credibili sulle caratteristiche tecniche dell'opera, così da individuarne correttamente i rischi connessi ed evitare che le richieste di compensazioni siano

³³ Cfr. CASINI L. (2007); CASSESE S. (2007); ITALIADECIDE (2009).

gonfiate dal timore circa la presenza di rischi ad esse non noti; la difficoltà di incorporare all'interno di una procedura d'asta richieste di varianti al progetto o di opere suppletive che attenuino le disutilità per la comunità ospitante a costi più contenuti rispetto a compensazioni puramente monetarie; l'incerta definizione dei confini delle comunità e dei rappresentanti delle stesse che dovrebbero esprimersi sul problema.

Queste considerazioni suggeriscono che il crescente interesse nei confronti dei meccanismi di asta, tanto a livello internazionale (il caso forse più rilevante è quello spagnolo) quanto in Italia³⁴, per produrre risultati positivi dovrà preventivamente mirare alla definizione di una serie di "condizioni". Innanzitutto occorre che l'asta sia attentamente disegnata, anche per tener conto di aspetti quali il costo netto complessivo (per l'intera collettività nazionale) delle compensazioni che si è disposti a mettere a disposizione. In secondo luogo, occorre prevedere la presenza di meccanismi istituzionali autorevoli e terzi atti a fornire garanzie circa i rischi effettivi connessi con l'opera e i comportamenti del gestore successivi alla sua realizzazione. Infine, il ricorso a meccanismi d'asta andrebbe limitato agli interventi che, per caratteristiche proprie dell'opera da realizzare (opere puntuali piuttosto che a rete, impianti che non generano esternalità sulle comunità vicine), consentano di definire con precisione i soggetti le cui preferenze si vuole indurre a rivelare. Accrescere il ricorso alle aste è perciò possibile e auspicabile, ma non è la soluzione universalmente da ricercare, né rappresenta una facile scorciatoia.

In generale, si dovranno perciò usare più strumenti, ciascuno dei quali richiede il rispetto di certe condizioni di contesto, e i diversi strumenti possono dover essere tra loro combinati. Nella maggior parte delle situazioni concrete, considerazioni tecniche (che impattano sui costi di costruzione e successiva gestione dell'opera) e preferenze delle comunità assumono entrambe rilievo, suggerendo che la soluzione ottima possa dover prevedere elementi dell'uno e dell'altro: la preventiva individuazione da parte del costruttore (o del decisore pubblico) dei siti che per caratteristiche tecniche meglio si prestano a ospitare l'infrastruttura e il successivo ricorso a meccanismi (d'asta o negoziali) di ascolto delle comunità prescelte al fine di individuare quella con le minori disutilità.

Quanto sinora detto ci consente altresì di evidenziare alcune criticità del sistema italiano di gestione del fenomeno *Nimby*. In primo luogo, la situazione italiana sembra caratterizzarsi per una sostanziale aporia intrinseca. Il tradizionale

³⁴ Prima che il referendum sancisse il definitivo abbandono del progetto di ricorso al nucleare, l'asta era stata proposta per la scelta dei siti ove localizzare le scorie nucleari.

approccio centralistico – che presuppone il coinvolgimento del territorio solo dopo l'adozione della scelta localizzativa e che per certi aspetti il legislatore ha di recente ripreso e rafforzato immaginando di affidare alla figura dei commissari il superamento delle situazioni di stallo diffuse sul territorio – si innesta su un apparato burocratico-amministrativo che, a seguito della modifica del Titolo V della Costituzione, è diventato profondamente decentrato e che in teoria imporrebbe la partecipazione di tutti i livelli di governo alle decisioni relative a programmazione, localizzazione e realizzazione delle infrastrutture. Il risultato è che i poteri locali, formalmente esautorati dai processi di scelta localizzativa, hanno poi, anche per via dei poteri amministrativi che per altri versi a loro fanno capo, modo e maniera di ostacolare la realizzazione delle opere ponendo ostacoli nei passaggi burocratici e procedimentali successivi. È a tale questione che ha cercato di dare soluzione, a un estremo, l'informale metodo di confronto *ex ante* sviluppato da molti operatori, all'altro il ricorso allo strumento commissariale. Come detto, non siamo in grado di esporre una valutazione compiuta dei dettagli delle prassi di un tipo e dell'altro concretamente sviluppatasi, anche con profonde eterogeneità interne, negli ultimi anni. Senz'altro, una maggiore stabilità normativa, e una maggiore coerenza dei vari provvedimenti speciali introdotti con le regole generali del sistema, renderebbe il quadro procedimentale più efficace.

Vicenda che in maniera paradigmatica mostra alcuni dei principali profili critici evidenziati nel lavoro è il noto caso dell'opposizione alla realizzazione della linea ferroviaria Torino-Lione messa in atto dagli enti locali minori, dai comitati e dalle comunità della Val di Susa. A differenza di quanto è accaduto in Francia in cui si è fatto ricorso a procedure volte a favorire la massima partecipazione di tutti gli attori interessati sin dalle fasi iniziali del processo decisionale, in Italia si è adottato un tradizionale approccio di tipo *top down*: l'approvazione del progetto è avvenuta senza aver preventivamente coinvolto gli enti locali minori e le comunità interessate e senza aver tenuto adeguatamente conto delle problematiche connesse al territorio. Ciò ha determinato il sorgere di una radicale opposizione dei comuni, non intenzionati a cedere prerogative e poteri di governo del territorio a loro legittimamente riconosciuti, e dei comitati e della popolazione locale, fermamente convinti della inopportunità dell'opera. Per sbloccare la situazione di stallo che ne è derivata nel 2005 è stato istituito un Osservatorio come sede di confronto di tutte le amministrazioni pubbliche coinvolte nella realizzazione dell'infrastruttura. A seguito di numerosi incontri in cui sono state analizzate le criticità ambientali, sociali ed economiche del progetto sono state individuate alcune soluzioni che il decisore politico ha successivamente adottato con l'approvazione

nel 2010 di un nuovo progetto preliminare³⁵. Questa soluzione non ha, però, consentito di riavviare i lavori. Le criticità caratterizzanti il caso TAV sono infatti anche segnaletiche della mancanza di una chiara e condivisa analisi dei costi e dei benefici connessi con la realizzazione dell'opera e della scarsa capacità di progettazione a monte. L'opera in quanto tale, anche indipendentemente dai suoi possibili costi di localizzazione, non è infatti chiaramente profittevole e sembra supportata più da motivazioni politiche che economiche. I vantaggi che scaturirebbero dalla sua realizzazione, difficilmente dimostrabili, hanno rappresentato una base piuttosto vaga di definizione di possibili compensazioni; inoltre, l'avvenuta approvazione del progetto senza aver ottenuto il previo consenso neppure sulle sue linee principali ha aumentato il rischio di *lock in* nelle trattative.

In generale, due considerazioni ci sembrano inoltre utili. Da un lato, non crediamo che il *quantum* di potere assegnato ai livelli inferiori di governo debba scaturire da una regola universale: così come, in tema di federalismo, è preferibile lasciare alle competenze centrali le materie su cui più rilevanti sono gli *spillovers* spaziali (e le economie di scala), la scelta del sito andrebbe collocata a un livello superiore ove a contare siano soprattutto considerazioni di tipo tecnologico. Dall'altro, ci sembra ineludibile una tendenza a un maggiore coinvolgimento del territorio nelle scelte di localizzazione, conseguenza della maggiore sensibilità che, coerentemente con quanto accaduto in tutti i paesi avanzati, si è sviluppata per le questioni ambientali e di salute e dell'elevato grado di congestionamento raggiunto dal territorio, che comporta che ogni nuova opera debba interloquire con quanto già esistente.

Quale che sia il grado di coinvolgimento del territorio, ci pare comunque rilevante il miglioramento di una serie di condizioni di contesto istituzionale, evitando situazioni di accentuata frammentazione (ad esempio, nel caso italiano, sarebbe utile superare la eccessiva numerosità dei comuni) e tramite forme stringenti, quanto a tempi e modalità, di coordinamento istituzionale. Con tempi e modalità precisi potrebbero anche sperimentarsi momenti di coinvolgimento della cittadinanza in quanto tale, ad esempio lungo la falsariga del cosiddetto *débat public* previsto in Francia. Si tratta di strumenti potenzialmente utili – anche e soprattutto in situazioni di sfiducia nelle istituzioni (locali) e al fine di trasmettere informazioni affidabili a, e recepire le istanze dei, i cittadini – che però richiedono in ogni caso di funzionare in un contesto di trasparenza e affidabilità (ad esempio in tema di certificazione delle informazioni trasmesse dal proponente l'opera) e

³⁵ Per la ricostruzione della vicenda si rinvia a CULLINO R. e FABRIZI C. (2011).

secondo regole che assicurino certezza dei tempi e degli impegni presi. Il fatto che oggi però non si disponga al momento di una valutazione dei loro benefici rispetto a un più tradizionale modello in cui la rappresentazione delle istanze della cittadinanza avvenga per il tramite degli amministratori locali regolarmente eletti (e che sono indotti a tenere in conto l'intero arco di esigenze dei loro amministrati, non solo quelle specificamente connesse con la singola opera infrastrutturale), ci induce a suggerire un approccio sperimentale alla questione.

Concludendo a proposito di *Nimby* e di contesto istituzionale, non possiamo però soprattutto tacere del fatto che molto frequentemente il *Nimby* in senso proprio si sovrappone infatti a situazioni in cui certi investimenti sono proposti in maniera estemporanea, con scarsa e poco attenta valutazione dei costi e dei benefici e dei volumi di utilizzo previsti. Inoltre, i rischi percepiti dalle comunità locali sono spesso gonfiati da una diffusa sfiducia istituzionale, non solo verso le scelte di programmazione e localizzazione delle opere, ma anche – e soprattutto – nei confronti di chi è tenuto a garantire il rispetto di standard sanitari e ambientali e vigilare sulla correttezza dei comportamenti durante il funzionamento dell'impianto. Una più chiara attribuzione dei compiti ordinari delle diverse istituzioni e una loro maggiore efficacia, anche indipendentemente dalle decisioni in tema di localizzazione di un'opera, spesso potrebbero prevenire la sfiducia dei cittadini e l'incancrenirsi dei fenomeni di *Nimby*.

BIBLIOGRAFIA

- BAJARI P. - MCMILLAN R. - TADELIS S., «Auctions versus Negotiations Procurement: An Empirical Analysis», *Journal of Economic Behavior and Organization*, vol. 35, no. 2, 2008, pages 372-399.
- CASINI L., «L'inchiesta pubblica. Analisi comparata», *Rivista Trimestrale di Diritto Pubblico*, n. 1, 2007.
- .-, «La partecipazione nelle procedure di localizzazione delle opere pubbliche. Esperienza di diritto comparato», in MACCHIATI A. - NAPOLITANO G. (a cura di), *È possibile realizzare le infrastrutture in Italia?*, Bologna, Il Mulino, 2010.
- CASSESE S., «La partecipazione dei privati alle decisioni pubbliche. Saggio di diritto comparato», *Rivista Trimestrale di Diritto Pubblico*, n. 1, 2007.
- CORSETTI A. - FERRARA R. - FRACCHIA F. - OLIVETTI RASON N., *Diritto dell'ambiente*, Roma-Bari, Laterza, 2008.
- CLARKE E., «Multipart Pricing of Public Goods», *Public Choice*, no. 2, 1971, pages 19-33.
- CULLINO R. - FABRIZI C., «Senza la base non si può cantare: la nuova linea ferroviaria Torino-Lione nell'esperienza italiana e francese», Banca d'Italia, *Le infrastrutture in Italia: dotazione, programmazione, realizzazione*, 2011.
- DEWATRIPONT M. - TIROLE J., «Modes of Communication», *Journal of Political Economy*, vol. 113, no. 6, 2005, pages 1217-37.
- FRANCHINI C., «La figura del commissario straordinario prevista dall'art. 20 del d.l. n. 185 del 2008», *Giornale di Diritto Amministrativo*, n. 5, 2009.
- GHIRARDATO P. - MARINACCI M., «Ambiguity Made Precise: A Comparative Foundation», *Journal of Economic Theory*, no. 102, 2002, pages 251-289.
- GOEREE J.K. - OFFERMAN T., «Competitive Bidding in Auctions with Private and Common Values», *The Economic Journal*, vol. 113, no. 489, 2003, pages 598-613.
- GROVES T., «Incentives in Teams», *Econometrica*, no. 41, 1973, pages 617-631.
- IAIONE C., «La localizzazione degli impianti e delle infrastrutture strategiche: una rassegna critica dei procedimenti in vigore», in ASTRID, *Programmazione, decisione e localizzazione degli impianti e delle infrastrutture strategiche. Proposte di riforma delle regole e delle procedure*, 2010.
- ISAE, *Gli ostacoli normativi alla realizzazione delle grandi opere pubbliche*, 2008.
- ITALIADECIDE, *Le politiche pubbliche per le infrastrutture di interesse nazionale*, 2009.
- KUNREUTHER H. - KLEINDORFER P., «A Sealed-Bid Auction Mechanism for Siting Noxious Facilities», *American Economic Review*, no. 76, 1986, pages 295-299.
- LABOCCETTA A.M., «Lavori relativi ad infrastrutture strategiche ed insediamenti produttivi. Commento agli artt. 161-194 del d.lgs. 12 aprile 2006, n. 163», in SAITTA F. (a cura di), *Il nuovo codice dei contratti pubblici di lavori, servizi e forniture*, Commentario sistematico, Padova, Cedam, 2008.

- LAURENT-LUCCHETTI J. - LEROUX J., *Choosing and Sharing*, mimeo, 2009.
- MERLONI F., «Infrastrutture, ambiente e governo del territorio», *Le Regioni*, 2007.
- MYERSON R. - SATTERTHWAITTE M., «Efficient Mechanisms for Bilateral Trading», *Journal of Economic Theory*, no. 28, 1983, pages 265-281.
- MINEHART D. - NEEMAN Z., «Effective Siting of Waste Treatment Facilities», *Journal of Environmental Economic and Management*, no. 43, 2002, pages 303-324.
- O' SULLIVAN A., «Voluntary Auctions for Noxious Facilities: Incentives to Participate and the Efficiency of Siting Decisions», *Journal of Environmental Economics and Management*, no. 25, 1993, pages 12-26.
- OCCHILUPO, R. - PALUMBO G. - SESTITO P. «Le scelte di localizzazione delle opere pubbliche: il fenomeno Nimby», *Questioni di Economia e Finanza*, Banca d'Italia, n. 91, 2011.
- PEREZ-CASTRILLO D. - WETTSTEIN D., «Choosing Wisely: A Multibidding Approach», *American Economic Review*, vol. 92, no. 5, 2002, pages 1577-1587.
- ROTHKOPF M.H., «Thirteen Reasons Why the Vickrey-Clarke-Groves Process is not Practical», *Operations Research*, vol. 55, no. 2, 2007, pages 191-197.
- TONETTI A., «La partecipazione nelle procedure di localizzazione di opere pubbliche: il caso italiano», in MACCHIATI A. - NAPOLITANO G. (a cura di), *È possibile realizzare le infrastrutture in Italia?*, Bologna, Il Mulino, 2010.
- VICKREY W., «Counterspeculation, Auctions and Competitive Sealed Tenders», *Journal of Finance*, no. 16, 1961, pages 8-37.

Financial Development and Economic Growth. A Theoretical and Empirical Overview

Filomena Pietrovito*

Università degli Studi del Molise

This article presents a review of the literature on finance and growth. Theoretical contributions show that the existence of financial institutions, motivated by the reduction of risk aversion and asymmetric information between borrowers and lenders, allows a better resources allocation which, in turn, leads to capital accumulation and technological innovation. Empirical analyses, by adopting different techniques able to measure the functions performed by financial systems, meant to identify the channels through which financial development enhances economic growth, and to attenuate endogeneity biases, suggest that policy makers should take into account the role played by financial systems in sustaining growth.

[JEL Classification: G0; O0].

Keywords: financial intermediaries; financial markets; financial development; economic growth.

* <filomena.pietrovito@unimol.it>. This paper is based on a portion of my dissertation at Università degli Studi del Molise. I would like to thank Alberto Franco Pozzolo for his valuable supervision and support. I would also like to acknowledge the comments of the members of my thesis committee: Giovanni Ferri, Sergio Ginebri and Francesco Nucci. Valuable suggestions from an anonymous referee of this journal are much appreciated. Remaining errors are my own. The usual disclaimer applies.

1. - Introduction

A financial system includes competitive financial intermediaries, such as banks, other financial institutions, financial markets and a central bank. Its key function consists of pooling financial resources from disparate individual savers that have a financial *surplus*, and ensuring that these savings are allocated to the most productive use. The differences in how well financial systems across countries provide this function determine the degree of financial development, whose role in contributing to long-term economic growth has been extensively investigated by economists.

The traditional development literature generally offers two opposing views of the relationship between finance and growth. The first view suggests that financial sectors develop, first of all, to satisfy the financial needs of firms operating in more developed countries, and then because wealthier economies are more able to afford a costly financial system (Robinson, 1952). The second view argues that differences in the degree of development of the financial system help to explain why countries grow at different rates because of the different ability in identifying and financing good investment opportunities (Bagehot, 1873; Schumpeter, 1912; Hicks, 1969).

The aim of this paper is to provide an overview of the most relevant findings from the theoretical models on finance and growth and to relate them to the empirical evidence. Since it would be too ambitious to provide a complete understanding of the existing debate, this paper only describes the most relevant contributions by emphasising their innovative aspects.¹

A common finding in the theoretical literature is that financial intermediaries and financial markets arise in order to: (i) provide risk-diversification services; (ii) reduce liquidity risk, and (iii) mitigate asymmetric information which entails either *adverse selection* or *moral hazard* problems. As emphasised by Demirgüç-Kunt - Levine (2008, page 1) «in emerging to ameliorate market frictions, financial arrangements change the incentives and constraints facing economic agents». Indeed, by fulfilling the mentioned services, financial institutions are able to increase the propensity to save and to invest, thereby stimulating long-run economic

¹ The growing body of theoretical and empirical research on finance and growth is extensively surveyed in the seminal paper by LEVINE R. (2005) and, more recently, by ANG J.B. (2008) and DEMIRGÜÇ-KUNT A. - LEVINE R. (2008). For a review of the different econometric methodologies used to assess the relationship between financial development and growth, see WACHTEL P. (2001) and BECK T. (2008).

growth.² The standard growth theory identifies two key channels through which financial development can influence growth: capital accumulation and technological innovation. The capital accumulation channel focuses on the ability of financial institutions to convey savings into the most productive sectors in order to fund investment projects. The technological innovation channel highlights the role of an efficient financial system in the adoption of expensive new technologies (Ang, 2008).

The role of financial development in contributing to long-run economic growth, as theoretically postulated, is extensively supported by the empirical literature. However, while the theoretical literature identifies specific functions to explain the link between finance and growth, due to data limitations, most empirical research focuses on size, activity and efficiency of financial intermediaries and markets. Also empirical research mostly progresses from the country-level down to the industry- and firm-level. While cross-country studies have developed various strategies to address biases due to reverse causality, the use of industry- and firm-level data, not only provides additional ways to deal with this issue, but also identifies specific channels through which financial development might enhance economic growth. Overall, the empirical evidence is consistent with the theoretical findings in assessing that «the services provided by the financial system exert a first-order impact on long-run economic growth» (Demirgüç-Kunt - Levine, 2008, page 1).

The remainder of this survey is organized as follows. Section 2.1 analyses how risk aversion and asymmetric information justify the existence of financial intermediaries and financial markets. Section 2.2 surveys the theoretical contributions that analyse the mechanisms through which financial systems are able to manage risk, while Sections 2.3 and 2.4 survey the literature on the mechanisms used to reduce, respectively, *ex-ante* and *ex-post* asymmetric information. On the empirical side, Section 3.1 provides a description of the main issues concerning the finance-growth nexus. Section 3.2 discusses the different proxies used to measure the level of financial development. Sections 3.3 and 3.4 analyse the solutions adopted to deal with the issues of causality and endogeneity, respectively. Section 4 concludes the survey.

² Almost all theoretical studies make comparisons between two environments: an autarkic environment, without a financial system, in which firms are not able on their own to invest in innovative projects and accumulate capital, and an environment that provides some financial services and shows higher performance in terms of growth rates.

2. - The Theoretical Channels

2.1 *Market Imperfection and Financial Institutions*

A key issue in the theoretical literature on finance and growth is how risk aversion and asymmetric information between borrowers and lenders can motivate the existence of financial institutions and of the mechanisms that allow them to channel resources into the most profitable and innovative projects. The ability of financial institutions and markets to mitigate market frictions can lead to more capital accumulation and technological innovation and thus foster long-run growth.

In an autarkic environment, in the absence of a financial system, when firms face wealth constraints in establishing investment projects, they need to raise funds from individual households that have a financial *surplus*. This can be achieved either by stipulating debt contracts or by selling equities, or through a combination of both. In general, financial contracts between lenders and borrowers are affected both by lenders' risk aversion and by asymmetric information about borrowers.³

Even in an environment where information on the borrowers is available, risk averse lenders might refuse to finance especially risky borrowers, such as new firms or projects with uncertain returns. This implies that savers are not willing to invest funds directly in more risky projects, even renouncing, on the other hand, to higher returns. In such a context, financial institutions may be set up to provide risk diversification services, thereby allowing direct financing even of the most risky projects. A second type of risk is liquidity risk which may arise because high-return investment require long-term commitment of capital, although risk-averse agents will be reluctant to commit funds for long periods, *i.e.* until the conclusion of a project. In an autarkic environment, liquidity risk would imply underinvestment of funds in more profitable projects. Therefore, in such a framework, financial institutions might emerge in order to mitigate the negative effects of liquidity needs on long-term investment, by allowing savers to withdraw their funds on demand or to sell their shares.

In order to provide risk diversification services, financial institutions reallocate the savings collected from several investors to high-return projects. In diversifying their asset portfolios, compared to individuals, financial institutions have the advantage of being able to reduce transaction costs, thereby profiting from economies of scale. Financial institutions improve capital allocation towards high-

³ See BOLTON P. - DEWATRIPONT M. (2005) for a detailed analysis of contract theory.

return, high-risk projects, thereby promoting capital accumulation and technological changes.

To mitigate liquidity risk, financial institutions may transform illiquid assets into liquid liabilities (Diamond - Dybvig, 1983). On the one hand, financial intermediaries are able to pool savings for long periods, while ensuring investors to withdraw their savings on demand when liquidity needs occur.⁴ On the other hand, financial markets allow savers to sell their securities to other investors if they need to access their savings.

A second justification for the existence of financial institutions is asymmetric information between borrowers and lenders. In a financial contract, asymmetric information occurs when the lender has insufficient information about the borrower. This asymmetry creates misinformation on which project to finance and its expected returns. In debt and equity contracts two types of asymmetric information can arise, since firms may behave opportunistically before and after the transaction takes place.

The presence of asymmetric information before the transaction takes place implies an *adverse selection* (or *hidden information*) problem. This refers to the situation where the lender does not have sufficient information about the risk-return characteristics of a project. In other words, «the lender is in the same position as a buyer of a second-hand car: it does not know perfectly the “quality” of the project it invests in» (Bolton - Dewatripont, 2005, page 57). Since the “quality” of the project is not public information, the lender may not be able to distinguish between “good” and “bad” projects, which may result in an inefficient allocation of resources due to “good” projects rejection. In order to improve the allocation of savings, the lender needs to acquire information about the entrepreneur’s trustworthiness, before the transaction takes place. Since to do so is both time consuming and costly, it is optimal to delegate the acquisition of this information to competitive financial intermediaries, such as banks, or to financial markets (Boyd - Prescott, 1986).

However, financial intermediaries may still face problems of adverse selection since they are not able, without a costly activity, to distinguish *ex-ante* between “good” and “bad” entrepreneurs. For instance, if financial intermediaries fix the interest rate to be paid by observing the distribution of risk among borrowers, it

⁴ In this case, the financial intermediary bridges between lenders and borrowers, collecting savings for short periods and issuing credit for long periods. The transformation of risk lies in both the deadline transformation and in the fact that the saver has as counterpart a financial intermediary which is less risky than the individual borrower.

may be that “good” entrepreneurs do not apply for loans due to the relatively high interest rate, while “bad” ones find it convenient to ask for funding.

One solution that financial intermediaries may adopt to mitigate adverse selection is to screen borrowers. By this means, financial intermediaries create some incentive to discourage “bad” entrepreneurs from declaring themselves as “good” and applying for the low-risk contract. This is significant since, if “bad” borrowers pretend to be “good”, there will be underinvestment in good projects and the overall economic growth rate will be negatively affected.

If screening does not distinguish between “good” and “bad” entrepreneurs, credit rationing might be a solution. In particular, it is used when «the demand for credit exceeds supply even though the banks are free to increase interest rates» (Freixas - Rochet, 1997, page 194). In order to explain the existence of credit rationing, Stiglitz - Weiss (1983) construct a model in which the bank’s expected profits are not monotonically increasing with the interest rate on the loan. Indeed, when the bank rises the interest rate, safer borrowers consider the cost of external finance to be too high as compared to the returns from their projects, and they drop out the credit market. Therefore, at higher interest rates only riskier borrowers remain in the credit market. In order to mitigate these adverse selection problems, banks are forced to decrease the interest rate below the equilibrium which makes the demand for credit higher than its supply, resulting in credit rationing.

When financial markets are transparent and efficient they play a crucial role in improving *ex-ante* information acquisition since their prices imply relevant information about listed firms. In this case, financial markets can help households to decide about investments without the intervention of financial intermediaries.

The second type of asymmetric information occurs *ex-post*, *i.e.* after the transaction takes place, when the lender cannot see what the borrower does and what is the outcome from the project. In this case, asymmetric information involves *moral hazard* (or *hidden actions*) problems. For instance, after obtaining a loan, the manager of a firm can decide to undertake only high-risk projects. In addition to lacking information on the development of the project, investors may also need information about the returns from the project, which is private information. This is particularly important in the case of debt contracts where the amount to be repaid depends on the project outcomes. In this case, the entrepreneur might decide to declare a loss to the *financier* in order to avoid repaying all of the loaned funds.

As it is in the case of *ex-ante* asymmetric information, it is better to delegate the reduction of *moral hazard* problems to a financial institution in order to

avoid cost duplication (Townsend, 1979; Diamond, 1984; Williamson 1986). For instance, in the “delegated monitoring” model developed by Diamond (1984), financial intermediaries have some cost advantage over direct lending in terms of monitoring the outcomes of entrepreneurs’ projects.⁵

The solution to the *moral hazard* problem requires financial intermediaries to devote some resources to monitoring how the entrepreneur is managing the project. Indeed, it is crucial for efficient resource allocation that financial institutions monitor firms and induce managers to maximize the firm’s value. Of course, the absence of financial arrangements that enhance corporate governance may hamper the mobilization of savings from disparate agents and also keep capital from flowing to profitable investment (Stiglitz - Weiss, 1983).

After a project is concluded, the lender may invest some resources to provide incentives to the borrower so she can truthfully report its outcome. This solution is described in the standard “costly-state verification models” which are based on the assumption that firms’ returns are unobservable (Townsend, 1979; Diamond, 1984; Gale - Hellwing, 1985; Williamson, 1986). Since repayment of equity contracts is based on firm performance which is observable with the payment of a fee, equity markets are not active in the standard environment and only debt is used. Debt implies repayments contingent on firm performance only in the event of bankruptcy and, therefore, verification costs are minimized (Hart, 1995).

Financial markets can also improve over corporate control, after a transaction has occurred, in two ways: first, by facilitating a hostile takeover and second, by structuring managerial compensation such that it is conditioned on the firm’s stock price performance (Holmstrom - Tirole, 1993). By linking stock market performance with manager’s compensation, it is possible to align the interests of managers and owners (Diamond - Verrecchia, 1982; Jensen - Murphy, 1990).

In sum, financial intermediaries and markets, by managing risk and mitigating asymmetric information, improve the allocation of savings and stimulate capital accumulation and technological progress, thereby promoting long-run economic growth.

⁵ Building on the Diamond’s idea, GALETOVIC A. (1996) proposes a model in which financial intermediaries emerge when increased production specialization is a necessary condition for sustained growth. In this case, financial intermediaries enable duplication of the monitoring effort, which is increasing in specialization, to be avoided. Therefore, “it is specialization in the real sector which causes the emergence of intermediaries, institutions that specialize in monitoring” (GALETOVIC A., 1996, page 551).

2.2 *Risk Diversification and Liquidity Provision*

The basic idea behind the models presented in this section is that high-return projects imply high levels of risk and long-term commitment of capital. Capital accumulation and technological innovation generally involve a high risk-return profile and require the investment of capital for more than one period before any profit is made. As already highlighted, since individual savers dislike risk, firms cannot obtain direct funding from external sources for capital accumulation and innovative projects, and require the risk amelioration services of both financial intermediaries and financial markets in the form of risk diversification and liquidity provision.

In this view, and focusing on the technological innovation channel, Saint-Paul (1992) constructs an endogenous model based on the idea that technological innovation can be achieved only through greater division of labour which implies higher specialization of resources and hence higher risk. The economy described in this model is exposed to demand shocks which may result in a lack of demand for some goods. In the absence of financial markets, entrepreneurs and consumers will diversify risk on their own, by choosing more flexible, but less productive technologies (*i.e.* technologies that can be used to produce several types of goods). Financial markets mitigate the negative effects of demand shocks by diversifying portfolios and enabling investment in more specialized and more productive technologies. In sum, an economy with a developed financial market will be able to achieve higher levels of growth enabled by risk diversification.

Along the same lines as Saint-Paul (1992), but focusing on the capital accumulation channel, Acemoglu - Zilibotti (1997) construct a model in which intermediation cost is endogenously determined by the risk-diversification efforts of agents. The main finding of this study is that, in the absence of financial intermediaries, savers are not willing to hold high-return, but risky assets. Therefore, capital accumulation and growth may be affected positively if financial institutions emerge to provide risk-diversification services and enable more productive use of funds.

The second strand of the literature surveyed below emphasises liquidity provision as the channel between financial development and real economic growth. The relevance of liquidity provision for economic growth is emphasised by Hicks (1969) and Diamond - Dybvig (1983). In particular, Hicks (1969) argues that the financial revolution occurred before the Industrial Revolution, in order to provide the necessary liquidity to adopt technological changes that require illiquid capital investment.

Building on the Diamond - Dybvig's setup, Bencivenga - Smith (1991) develop a model of endogenous growth in which they find that the role of banks is to affect the composition of savings in such a way as to stimulate capital accumulation. Their model has two important results. First, an economy with a financial sector will invest more of its savings in capital goods (more illiquid, but productive assets) and less in unproductive but liquid assets, than an economy without a financial sector. Second, in an economy without financial intermediaries, savers are forced to liquidate their investments when liquidity needs arise, whereas an economy populated by banks eliminates premature liquidation of illiquid assets. In fact, while individuals face uncertainty about their future liquidity needs, banks, by applying the law of large numbers, may predict the probability of withdrawals. Consequently, an economy with a financial sector can grow more than an autarkic economy because in the latter all investment is self-financed and subject to premature liquidation if liquidity needs arise.

In addition to the role of financial intermediaries, the ability of financial markets to increase liquidity is considered in the literature. For instance, by taking account of the technological channel, Levine (1991) develops an endogenous growth model where financial markets arise to reduce the negative impact of productivity and liquidity shocks on economic growth. First, financial markets allow individuals to invest in large number of firms and diversify away productivity shocks, thereby increasing the proportion of resources invested in firms. Second, stock markets allow investors to sell their shares to other investors if liquidity shocks occur, thereby avoiding premature liquidation of capital and increasing the resources maintained in the firm. The result is that stock markets contribute to overall economic growth.

So far, financial market liquidity has referred to the possibility of withdrawing funds on demand or selling shares to other investors in order to satisfy liquidity needs. Bencivenga *et al.* (1995) use an alternative measure of financial market liquidity, *i.e.* reduction of transaction costs based on commission, fee, tax and/or time spent on the financial transaction. Therefore, they define liquid markets as those having low transaction costs, and illiquid markets as those with high transaction costs. Since the most productive capital investments imply a long-term commitment of capital, the use of these technologies in an economy requires illiquid (long-term) investment. As a consequence, because of liquidity needs, ownership of capital-in-process needs to be transferred to secondary markets through the payment of the aforementioned transaction costs. Bencivenga *et al.* (1995) argue that a reduction in transaction costs, by facilitating transactions among in-

vestors and by improving market liquidity, stimulates investment in long-term technologies and economic growth.

In comparing the ability of financial intermediaries and financial markets to provide insurance to households against idiosyncratic liquidity shocks, Fecht *et al.* (2008) construct a model in which they show that there is a trade-off between risk-sharing and growth. On the one hand, financial intermediaries provide for more risk-sharing among depositors while investing less in productive technology.⁶ On the other hand, financial markets in which households can invest directly by paying a cost, provide less risk-sharing while promoting investment in productive technology. As a consequence, bank-oriented economies will grow more slowly than more market-oriented economies.

In sum, the literature highlights that financial development improves upon an autarkic situation because it allows firms to undertake riskier, but more productive and longer-term projects. By channelling resources into the highest returns, financial institutions are able to play a crucial role in capital accumulation and technological innovation, which are considered the main channels of long-run economic development (Greenwood - Smith, 1997). Therefore, the risk amelioration service provided by financial institutions can be considered an important determinant of economic growth.

2.3 *Credit Rationing and Screening*

A relevant strand in the literature assumes that a well developed financial system dominates over an autarkic environment because it is able to mitigate *ex-ante* asymmetric information between borrowers and lenders. In all the studies surveyed in this section, better financial systems are more successful at evaluating projects whose returns are difficult to estimate, through screening or rationing strategies. Consequently, a common finding is that financial development promotes growth through an increase in capital accumulation and more efficient investment in intangible assets.

For instance, Boyd - Prescott (1986) focus on the screening role of banks to show how these institutions, by generating signals about the economic situation, could promote the withdrawal of funds from “bad” projects and their investment in more profitable ones. By examining the role of banks in altering aggregate in-

⁶ FECHT F. *et al.* (2008) do not consider an overlapping generation model and, therefore, competitive banks maximize the utility of depositors alive at each date, without taking account of capital accumulation for future generations.

vestment patterns, Boyd - Prescott (1986) highlight an important link between the real and the financial sectors.

In a seminal paper, Bencivenga - Smith (1993) analyse the consequences, for real rates of economic growth, of *ex-ante* informational frictions that give rise to credit rationing. They show that a policy action addressed to increasing the probability of obtaining a loan (*i.e.* to reduce credit rationing) has a positive effect on capital accumulation and, hence, on long-run growth rates.

Building on Bencivenga - Smith (1993); Bose - Cothren (1996), in addition to the credit rationing strategy, consider the screening technology as a way to determine the borrower type (*i.e.* probability of defaulting). Whether a screening or rationing regime prevails depends on the marginal product of capital. If the marginal product of capital is high, then the returns from investing are high. As a consequence, it is costly to ration credit (*i.e.* to not fund the project) and hence the screening solution prevails. However, if the marginal product of capital is low, then the rationing strategy will dominate. Moreover, Bose - Cothren (1996) find that the growth rate of capital (and output) is affected by the cost of screening. More specifically, the long-run economic growth rate will increase only if the cost of screening falls and a threshold level of sophistication of the financial sector is reached.⁷

Greenwood - Jovanovic (1990) describe an economy where each agent can invest her capital in either a safe, but low return technology, or in a riskier, but more profitable one.⁸ The asymmetric information in this model is based on the fact that all that an agent can costlessly observe is the realized composite rate of return (*i.e.* the composite technology shock) on his own project (Greenwood - Jovanovic, 1990). In such an environment, one reason why financial intermediation may arise endogenously is to collect and process information on the overall return from the project, coordinate productive activity, and spread risk across projects. These services require a lump-sum fee to be paid to intermediaries. Hence, Greenwood - Jovanovic (1990) show that only economies with high growth rates are able to afford intermediation costs and to develop a financial structure. After randomly selecting a finite number of high return-risk projects, intermediaries can compute the average net realized rate of return on these projects. Therefore, financial intermediaries can improve upon an autarkic situation

⁷ Similar results, based on a neoclassical framework, are presented in BOSE N. - COTHREN R. (1997).

⁸ More specifically, the second technology is characterized by a composite technology shock comprising a common component and a specific shock component.

by allocating resources to the highest return projects. As is well known, this increases the economic growth rate.

Alongside investment in capital goods, *ex-ante* asymmetric information may also affect innovative projects. In this case, the friction is related to the probability of success of the project which depends on the entrepreneur's actions. Therefore, as well as channelling funds into the most profitable firms, financial institutions may emerge to identify those firms with the highest probability of success in undertaking either product or process innovations (King - Levine, 1993*a*). The contributions in this strand of the literature are based on the Schumpeterian "creative destruction" view (Schumpeter, 1912) and on the R&D setup proposed in Aghion - Howitt (1992).

For instance, in King - Levine (1993*a*) the financial system plays a key role in evaluating and financing innovative investment projects.⁹ These authors find that a decrease in the cost of evaluating and financing entrepreneurs, and an increase in the efficiency of the financial sector due, for instance, to a decrease in the time cost of evaluation, lead to a higher equilibrium growth rate. Hence, in this model, the financial system plays an active role in promoting productivity growth, by mitigating *ex-ante* asymmetric information.

In comparing the performance of financial markets and intermediaries in financing technological innovation, Allen - Gale (1999) construct a diversity of opinion model, in which each investor has her own belief about the best choice to be made and its potential returns. The authors argue that in more uncertain environments, such as those characterized by innovative industries, markets have considerable advantage over banks. Indeed, in financial markets each investor decides whether or not to finance an innovation based on her beliefs. Consequently, at least some projects will be financed because investors will "agree to disagree". On the other hand, in an environment characterized by homogeneity of opinions, financial intermediaries are at an advantage in reducing information costs, since only the manager needs to be informed, but the disadvantage is that they may not be able to fund innovative projects when a high number of investors disagrees with the manager (Diamond, 1984).¹⁰

⁹ KING R.G. - LEVINE R. (1993*b*) do not distinguish between intermediaries and markets so that their model can be used to explain the link between finance and growth in both market-based and bank-based systems.

¹⁰ In DIAMOND D.W. (1984), however, investors agree to finance projects and in this case of converging beliefs, financial intermediaries act as delegated monitors economizing on information costs.

This brief review of the theoretical literature highlights the importance of financial institutions in reducing information costs and free-rider problems that arise in an autarkic environment. In fact, the financial system collects information on projects before contract loans are stipulated, thereby improving the use of savings. Clearly, only the most productive investments will be undertaken, and less profitable projects will not be funded. The fact that only highly profitable investments are financed allows the overall economy to grow at a faster rate.

2.4 *Monitoring and Costly-State Verification*

In an environment where investors are able to exert corporate control by inducing managers to maximize firm value, direct financing will lead to an efficient allocation of resources to the most productive projects. However, when shareholders are not able to control managers, the intervention of financial intermediaries is required in order to align the goals of managers and investors (Shleifer - Vishny, 1997).¹¹

The literature surveyed in this section considers the need for a financial institution to monitor the borrower's efforts so that the highest return is achieved. The theoretical literature highlights the relevance of *ex-post* information, especially for innovative projects, which imply the highest levels of riskiness. For instance, De la Fuente - Marin (1996) develop an endogenous growth model where the probability of success in innovation depends on the efforts of an unobservable entrepreneur. In this framework, banks may use a costly monitoring technology to provide more precise information at rising costs. Thus, monitoring encourages the channelling of funds to risky innovative activities by spurring growth. On the other hand, growth may stimulate financial development since the optimal level of monitoring positively depends on capital accumulation.

Galetovic (1996) develops an R&D model where each research firm is characterized by a degree of specialization, based on the number of workers that participate in production. Since research firms do not make profits, they need to seek external credit. Without financial intermediaries, it would be necessary for each worker to spend some resources to verify the firm's output. Therefore, in such a framework, financial intermediaries may emerge endogenously in order to prevent, as in Diamond (1984), the duplication of monitoring efforts. The

¹¹ Some economists argue that this solution shows some weaknesses due to the fact that banks can disproportionately increase their power and induce managers to act in their interest (see, for example, BOOT A.W.A. - TAKOR A., 2000).

conclusions from this study rely on the fact that financial intermediaries endogenously emerge, depending on the degree of the firm's specialization, and work to reduce monitoring costs, thereby sustaining economic growth.

In the R&D model developed by Blackburn - Hung (1998) the outcome of the innovation project is private information and investors need to spend some resources to acquire information, using costly monitoring technology. The main implication is that monitoring imposes a fixed cost on R&D since the innovator's maximization problem is subject to the investor's incentive constraint. In other words, investors will fund innovative projects only if the expected repayment, net of monitoring costs, equals the return on a safe investment. The main result of this model is that financial development occurs when a financial institution reduces the fixed costs of R&D, by eliminating duplication of monitoring costs, which results in an increase of the expected profits of R&D firms and encourages more innovating firms to enter the market by promoting overall economic growth.

In Sussman (1993) and in Harrison *et al.* (1999) the link between financial development and real growth is analysed in both directions, in a costly-state verification model. Both studies find that economic growth helps to increase the profits and activities of banks, thereby inducing new banks to enter the industry. Since in this model the greater the distance between the bank and the firm being financed, the higher the monitoring costs, an increase in the number of banks will reduce the average distance and the monitoring costs.¹² At the same time, higher levels of banking activity will lead to more investment projects and greater growth of the economy.

In a recent contribution, Greenwood *et al.* (2010a) develop a costly-state verification model in which the efficiency of the monitoring activity, *i.e.* the probability of detecting malfeasance, depends upon the resources devoted to the monitoring technology.¹³ Greenwood *et al.* (2010a) demonstrate that when the efficiency of the auditing activity increases due to technological innovation in the financial sector, then the cost of information production decreases. This increases

¹² In other words, when new banks enter the industry, the local specialization of each bank increases and the intermediation cost falls.

¹³ The ideas contained in GREENWOOD J. *et al.* (2010a) originate from a previous study developed by ROUSSEAU P.L. (1998) that describes a mechanism based on the role of innovation in the interest rate setting decisions of competitive intermediaries. More specifically, an increased ability to monitor borrowers allows intermediaries to lower interest rates on loans. This increases both the size and the efficiency of the financial sector since: (i) only higher-quality borrowers will seek for intermediated finance and (ii) competition among lenders will eliminate temporary profits.

financial intermediation efficiency for several reasons: (i) the spread between the firm's internal rate of return and the lending rate decreases; (ii) the spread between the return on loans received by the intermediary and what it pays to savers is reduced, and (iii) in the long-run, unproductive projects will no longer be financed and capital will be directed towards more profitable projects. Therefore, overall economic performance will improve as a result of increased technological innovation and capital investment.¹⁴

The mechanisms described above emphasise the advantage of financial institutions, especially banks, in acquiring *ex-post* information about the entrepreneur's efforts and the project's returns. The role of banks is fundamental for several reasons: first, banks are able to avoid cost duplication by providing direct financing collected from several agents; second, banks are specialized in both screening and monitoring activities, allowing them to benefit from economies of specialization; finally, banks can make use of the information on borrowers acquired during a long relationship.

3. - The Empirical Evidence

3.1 *Main Issues*

The role of financial development in contributing to long-term growth, theoretically postulated, has been extensively tested empirically. Evidence on the relationship between financial development and economic growth dates back to Schumpeter (1912) who argues that financial system development spurs the long-run economic growth rate through identification and financing of good growth opportunities. However, subsequent studies support different perspectives on the link between financial development and economic growth. For instance, Robinson (1952, page 80) argues that «where enterprise leads, finance follows», meaning that financial development is a consequence and not a cause of economic growth since it responds to an increase in the demand for financial services and institutions. On the other hand, Goldsmith (1969); McKinnon (1973) and Shaw (1973), along Schumpeterian lines, emphasise a proactive role of financial development and growth, which stems from an increase in capital stock and TFP. In sum, earlier empirical studies do not achieve consensus on the direction of causality between finance and growth because it is not clear whether financial develop-

¹⁴ For a quantitatively investigation of the impact of financial development on growth using the aforementioned costly-state verification model, see GREENWOOD J. *et AL.* (2010*b*).

ment is a catalyst for economic growth or whether economic growth stimulates the supply of financial services.

As convincingly argued by Levine (2005), the relationship between finance and growth is affected by a data problem involving the proxies used to measure financial development. Indeed, while the theoretical literature links specific functions of the financial system to economic growth, data limitations make it difficult to construct accurate measures of these functions.

An additional and critical issue concerning the relationship between finance and growth is whether the causal effect goes only from financial development to growth or goes in both directions.¹⁵ As highlighted by Rajan - Zingales (1998), there may be two main sources of endogeneity. First, there might be some omitted variables that affect the levels of both financial development and economic growth, determining a correlation between them. For instance, the propensity of households to save is an important determinant of both long-run economic growth, because the more people save, the more capital the economy can accumulate, and of financial development, because the more people save, the higher the demand for financial services, so that the financial sector becomes more developed. Second, financial development may affect economic growth simply because it anticipates future growth. For example, if the stock market capitalizes the present value of growth opportunities, and financial institutions are more willing to lend in the belief that sectors will grow in the future, the development of the financial sector could be due simply to higher demand for loans from firms taking advantage of future growth opportunities. In all these cases, the estimated coefficient of financial development in a growth regression would suffer from an upward bias and this would overestimate the impact of financial development on growth.

The following sections review the more recent empirical works, highlighting their contribution to addressing the main problems concerning the finance-growth nexus. Specifically, various proxies for financial development have been used to measure financial intermediaries and financial market development in an attempt to reconcile theoretical functions and data availability. Also, the most recent literature addresses explicitly the causality issue by: (i) adopting more disag-

¹⁵ From a theoretical point of view, several contributions suggest that, not only does financial development have a positive effect on growth, but it also could enable the provision of funds to growing countries (GURLEY J.G. - SHAW E.S., 1955; GREENWOOD J. - JOVANOVIC B., 1990; SAINT-PAUL G., 1992; BLACKBURN K. - HUNG V.T.Y., 1998).

gregated data and (ii) using interaction terms.¹⁶ Finally, recent contributions deal with the issue of endogeneity by: (i) proxing financial development by its predetermined value or using the instrumental variable approach and (ii) adopting more advanced econometric methodologies, such as panel data and time-series techniques.

3.2 Proxies for Financial Development

While the theoretical literature links specific functions of the financial system to economic growth, empirical researchers are forced to use variables measuring the size, activity and efficiency of financial institutions and markets, due to data limitations (Beck, 2008). The main source of data on financial intermediaries and financial market development is the World Bank's Financial Development and Structure Database, described in Beck *et al.* (2009).

This section describes the most widely used indicators of financial development distinguishing those measuring: (i) financial "depth"; (ii) relative importance of commercial banks over the central bank; (iii) credit to the private sector, including liabilities; (iv) size, liquidity and efficiency of the stock market, and (v) the ease with which financing can be obtained (indirect measures of financial development). In order to measure the size of the financial sector, Goldsmith (1969); King - Levine (1993*a,b*) and Levine *et al.* (2000) use liquid liabilities of banks and other financial intermediaries, over GDP. However, this measure is inappropriate for measuring the functioning of the overall financial system, since it concentrates on the size of the system, without taking account of its activity.

To measure the relative abilities of commercial banks and the central bank to allocate savings, King - Levine (1993*a,b*), as well as Levine *et al.* (2000) construct an indicator which equals the deposit money bank domestic assets over deposit money banks assets plus central bank domestic assets (total credit). Although this indicator is an improvement on liquid liabilities, since it captures who is allocating the credit (the central bank or a commercial bank), it has some limitations since: (i) it does not distinguish who is using credit (government or firm) and (ii) banks are not the only financial services providers.

¹⁶ The interaction term consists in interacting the country level financial development indicator with industry or firm characteristics, in order to identify the specific channel through which financial development affects economic growth. The interaction effect helps to explain causality since it focuses on the independent effect of financial development on growth, as well as on the theoretical mechanism underlying this relationship.

In order to capture other financial aspects related to the degree of financial intermediation, King - Levine (1993*a,b*) add two more indicators of credit provided to the private sector. The first measures the total amount of credit allocated to private business over total domestic credit. The second measures the credit issued to private enterprises divided by GDP. Since these measures include only credit to the private sector, they are more appropriate for evaluating capital accumulation and technological changes and, through these channels, economic growth. By regressing three different measures of economic growth on the above described proxies for financial development, King - Levine (1993*a,b*) find a strong and positive relationship between the predetermined measures of financial development and subsequent growth.

Furthermore, Levine *et al.* (2000) use the value of credits issued by financial intermediaries to the private sector divided by GDP. This indicator improves on the measures of private credit used by King - Levine (1993*a,b*) because it includes only the credit issued by banks and other financial intermediaries and does not include the credit issued by monetary authority and government agencies. Therefore, the measure proposed by Levine *et al.* (2000) represents an improvement in: (i) more comprehensively measuring the banking sector and (ii) more carefully distinguishing who is conducting the intermediation and where the funds are flowing. Levine *et al.* (2000), consistent with the previous literature, find that the exogenous component of financial intermediary development is positively associated with economic growth.

So far, this section has discussed the degree of financial development by considering only financial intermediaries. A significant portion of financial services is provided outside the banking system, and there is a group of studies claiming that the stock market provides different services from the banking sector, which may influence investment decisions and real growth. For example, Demirgüç-Kunt - Levine (1996) show a positive relation between *per-capita* GDP and stock market development, measured by turnover *ratio* which proxies for stock market efficiency and equals the value of the shares traded in domestic exchanges divided by the total value of listed shares. Levine - Zervos (1998), in addition to turnover *ratio*, use indicators that measure stock market capitalization and value traded. Stock market capitalization is measured by the value of the listed shares divided by GDP, and proxies for the size of the stock market; value traded equals the value of the domestic shares traded in domestic exchanges divided by GDP, and proxies for stock market liquidity. Their results are consistent with theoretical

models which emphasise that the stock market facilitates long-run growth (Levine, 1991; Bencivenga *et al.*, 1995).

Finally, indirect measures of financial development are used in the empirical literature and improve on past proxies because they are forward looking indicators of financial efficiency. For example, Rajan - Zingales (1998), together with the more traditional indicators of financial development presented above, adopt the accounting standards of the particular country which «reflects the potential to obtain finance rather than the actual finance raised» (Rajan - Zingales, 1998, page 571). They provide a country index created by rating the annual reports of at least three firms in each country studied. At regional level, Guiso *et al.* (2004*b*), in order to show that local financial development matters in an integrated country such as Italy, adopt a different indirect measure of financial development that is based on the idea that «a good indicator of financial development would be the ease with which individuals in need of external finance can access them and the premium they have to pay for these funds» (Guiso *et al.*, 2004*b*, page 933). To construct such a measure Guiso *et al.* (2004*b*) use the Bank of Italy Survey of Households Income-Wealth which asks households whether they have been denied credit or have been discouraged from applying for it.

3.3 Causality

In order to assess the causality between financial development and economic growth, several studies use industry- and firm-level data as well as cross-country data. Moreover, at the microeconomic level, the interaction effect is used to identify the specific channels underlying the relationship between finance and growth.

In cross-country analyses the average level of financial development over a long period of time is regressed on the average economic growth rate (King - Levine, 1993*b*; Levine - Zervos, 1998; Levine, 1999; Levine *et al.*, 2000; Beck *et al.*, 2000). This enables an evaluation of the average influence of the determinants of economic growth. For instance, King - Levine (1993*b*), using data for 80 countries for the period 1960-1989, find that different indicators of financial development are strongly and positively associated with real *per-capita* GDP growth, capital accumulation, and TFP growth. Similarly, Levine - Zervos (1998), for a sample of 47 countries for the period 1976-1993, find a positive impact of stock market liquidity on the economic growth rate, capital accumulation and productivity growth. Nevertheless, they find that bank development is still a significant determinant of economic growth meaning that, consistent with Allen - Gale

(1999), competitive intermediaries and financial markets are complementary. However, Rousseau - Wachtel (2011) in an attempt to examine the robustness of the “now-classic” findings concerning the cross-country relationship between financial development and growth, find that the finance-growth nexus is not as strong in more recent data as it was in original studies using data from the 1960s to the 1980s. The explanations rely on the incidence of domestic banking and financial crisis on the financial liberalisation process and on the omission of the role of equity market.

The justification for adopting a more microeconomic perspective relies on the fact that it enables researchers to control more adequately for causality between finance and growth since it relates industry- or firm-level growth to country-level financial development indicators and other determinants. This microeconomic perspective sheds light on the specific channels through which financial development enhances economic performance and enables analysis of the characteristics of industries and firms that benefit most from improvements in the country's financial system.

One way to identify the specific channel through which financial development affects economic growth is to test whether the coefficient of financial development varies across industries or firms with different characteristics, utilizing the interaction terms (Beck, 2008). Based on this idea, the seminal paper by Rajan - Zingales (1998) analyses to what extent the effect of financial development on industry value added growth varies with its level of external financial dependence. In order to test this accelerator effect, Rajan - Zingales (1998) interact demand for external finance, measured as the *ratio* between investment minus cash flows and investment, for all US manufacturing industries, with alternative indicators of national financial development. They find that the interaction term is positive and significantly different from zero, meaning that financial development fosters economic growth through the ability of financial institutions to channel savings to industries that are more reliant on external finance.

The interaction effect developed by Rajan - Zingales (1998) has been used to test the capital allocation hypothesis, which consists in assessing whether financial development plays a key role in allocating capital among firms and industries with better growth opportunities. For instance, Fisman - Love (2004) test whether two countries growing (or declining) in similar industries during the same period, exhibit similar patterns of growth if both have well developed financial institutions. By using different proxies for financial development, Fisman - Love (2004) find that the interaction term between financial development and similarities in growth

opportunities is negative and significantly different from zero even after controlling for a set of variables. This means that financial development helps countries to align their growth opportunities to realised growth rates.

Bekaert *et al.* (2007) examine the role of financial development in enabling countries to take advantage of growth opportunities using industry-level data. Their main assumption is that countries' growth opportunities are reflected in the growth potential of the mix of industries in the price to earnings *ratios*. The results in Bekaert *et al.* (2007) reveal that: (i) growth opportunities predict growth in both developed and emerging markets; (ii) domestic banking development is important for exploiting growth opportunities, while stock market development is not, and (iii) growth opportunities are better predictors of growth in more financially open markets.¹⁷

Another methodology that can be considered a variant of the interaction effect is developed in Wurgler (2000) to assess whether, in more financially developed countries, capital is allocated among those industries with higher growth opportunities, and is withdrawn from declining industries. Specifically, Wurgler (2000) estimates an elasticity for each country which measures the extent to which investment growth rates react to value-added growth rates. Wurgler (2000) finds that, for almost all the countries in the sample, elasticity is positive and significantly different from zero, meaning that funds flow from declining to growing industries. Moreover, to find the determinants of international differences in the elasticity of investment to value added, he regresses each elasticity on the relevant country's degree of financial development and finds that it strongly and positively affects the responsiveness of investment to value added growth. Using the estimated investment elasticity to value added as a measure of the efficiency of capital allocation, Pang - Wu (2009) find that countries with better financial development show higher investment elasticities in industries more reliant on external finance.

Altogether, the aforementioned analyses support the theoretical view that financial markets channel funds into most productive uses, *i.e.* to firms and industries with good growth opportunities, even if they do not have enough internal funds available.

At firm-level, for instance, Love (2003) uses a structural investment model to investigate whether financial development affects firms' investments through its

¹⁷ In an earlier study, BEKAERT G. *et AL.* (2005) show that equity market liberalisation, which gives foreign investors the opportunity to invest in domestic equity securities and domestic investors the right to transact in foreign equity securities, has a positive impact on a country's growth rate even after controlling for growth opportunities.

ability to reduce financing constraints. More specifically, Love (2003) conjectures that financial development reduces the impact of financing constraints (measured by cash flow) on investment. The contribution of this work is huge since it shows that the decrease in financial constraints, due to improvements in financial markets, allows easier access to external funds for firms with good investment opportunities and this improvement in capital allocation, in turn, enhances growth.

The interaction effect is used in several empirical works to show that larger firms have easier access to external finance and, therefore, are less affected by higher financial development than smaller firms (Bernake *et al.*, 1998; Love, 2003; Guiso, Jappelli, Padula, Pagano, 2004; Beck *et al.*, 2008*a,b*). Moreover, other channels through which financial development affects firms' economic performance are: reduction of financial constraints from investing in profitable growth opportunities (Demirgüç-Kunt - Maksimovic, 1998; Love, 2003), ability to finance product or process innovation (Benfratello *et al.*, 2008), and reduction in the financial, legal and corruption barriers faced by firms (Beck *et al.*, 2005).

In sum, several researchers emphasise the accelerator effect of financial development on country, industry or firm growth by adopting the methodology developed by Rajan - Zingales (1998). In practice, the empirical evidence reveals that, in countries with well developed financial systems, savings are more likely to be channelled into industries and firms that are more dependent on external finance, face higher growth opportunities and are more financially constrained.

3.4 *Endogeneity*

3.4.1 First Attempts to Address Endogeneity

At country-level, the studies by King - Levine (1993*a,b*) and by Levine - Zervos (1998) represent the first relevant attempts to disentangle the causal relationship between finance and growth. In particular, King - Levine (1993*b*) provide evidence on the causality between finance and growth by regressing the level of economic growth, physical capital accumulation and economic efficiency improvements, averaged over 30 years, on the initial (in 1960) level of financial development. Using an indicator of financial development at the beginning of the period, King - Levine (1993*b*) find a positive and strong relation between finance and growth, even after controlling for other influencing variables. Levine - Zervos (1998) adopt the same strategy as King - Levine (1993*b*) and find also that the initial level of stock market liquidity has a positive effect on growth, capital ac-

cumulation and productivity. This implies that the initial level of financial development, therefore, is a good predictor of subsequent economic growth.

However, as stressed by Rajan - Zingales (1998), even after proxing for financial development at its initial level, the causality problem could persist if there are omitted variables that affect both economic growth and the initial level of financial development, or if financial development affects economic growth because it anticipates future growth. Therefore, a limitation of the previously cited works is that they do not completely address the endogeneity problem.

3.4.2 Instrumental Variables for Financial Development

A more appropriate way to deal with endogeneity is to isolate the exogenous component of financial development on economic growth through instrumental variables. It is well known that good instruments need to be strongly correlated with financial development indicators, but uncorrelated with unobserved factors influencing economic performance and not included in the model specification. This property guarantees that the only effect that the instrumental variable can exert on economic growth is through financial development. Widely accepted exogenous factors explaining differences in countries' level of financial development are: legal origin, measures of creditors' rights, quality of law enforcement and accounting standards, as demonstrated by the law and finance literature.¹⁸

The law and finance literature is based primarily on the seminal paper by La Porta *et al.* (1998) which provides evidence on the correlation between a country's legal tradition and the level of legal protection allowed to creditors and the efficiency of law enforcement. This work complements and confirms some of the findings in La Porta *et al.* (1997), which assess that countries with poor investor protection have significantly smaller debt and equity markets.¹⁹

The aforementioned studies lead to the conclusion that a country's legal origin, as well as its investors' standards of protection and law enforcement, can be used as instruments for its financial architecture since they are based on exogenous factors, such as occupation and colonization processes, and influence the way in

¹⁸ A detailed survey of the law and finance literature is provided by BECK T. - LEVINE R. (2005). Unlike the "intermediary-based" and the "market-based" views, which respectively emphasize the role of banks and financial markets in promoting financial development, the "legal view" claims that it is the legal rights of investors that matter with financial development.

¹⁹ Consistent with the findings of La PORTA R. *et AL.* (1997, 1998); BECK T. *et AL.* (2003) find further support for the view that the legal origin of a country matters for the development of financial intermediaries and explains some of the differences among financial intermediary sectors.

which financial institutions are determined. While the law and finance theory predicts that the legal and regulatory environment of a country influences the development of its financial infrastructure, the finance and growth literature uses legal and regulatory determinants as instrumental variables to extract the exogenous component of finance in growth and to control for simultaneity bias.

At country-level there are several important contributions in the law and finance and in the finance and growth literatures, such as Levine (1998, 1999) which use instrumental variable techniques to extract the exogenous component of the connection between financial development and long-run growth. The general results of Levine's studies confirm the view that a country's overall legal and regulatory environment affects its rate of economic growth through banking sector development. Therefore, part of the influence of finance on economic performance is captured by the exogenous determinants of financial development.

A further contribution to the view that legal and regulatory determinants matter for financial intermediary development and positively affect economic growth, is provided by Levine *et al.* (2000). More specifically, they use legal origin, indicators of creditor rights, contract enforcement and accounting standards as instruments for financial development. They find that there is a strong connection between the exogenous determinants of financial development and real *per capita* GDP growth. The instrumental variable results indicate an economically large impact of financial development on growth. For instance, «the estimated coefficient suggests that if Argentina had enjoyed the level of financial intermediary development of the average developing country during the 1960-1995 period it would have experienced about one percent point faster real *per capita* GDP growth *per annum* over this period» (Levine *et al.*, 2000, page 35).

At regional level, Guiso *et al.* (2004a), in order to assess whether local financial development enhances the probability that an individual starts a new business, favours the entry of new firms, increases competition and promotes growth, construct an indicator of financial development based on the ease with which individuals can access external funds and the premiums they have to pay. Since this indicator could be correlated with some unobserved determinants of entrepreneurship, Guiso *et al.* (2004a) need also to address the issue of reverse causality.²⁰ For this reason, regional financial development is instrumented with variables that capture the exogenous variation in the supply of credit in 1936: number of branches

²⁰ For instance, «if demand for financial development generates its own supply, the regions with the best economic prospects might have the most financially developed banking system», (GUIISO L. *et al.*, 2004b, page 936).

per million inhabitants in the region, fraction of branches owned by local banks, number of saving banks per million inhabitants in the region, and number of cooperative banks per million inhabitants in the region.²¹ By instrumenting financial development with these indicators, Guiso *et al.* (2004a) show that the differences in 1936 regional regulations would affect the availability of credit 60 years later.

Finally, using firm-level data, Beck *et al.* (2005) provide evidence about the relevance of financial, legal and corruption obstacles for firm growth rates. They find that the extent to which the financial and legal systems and the level of corruption influence firm growth, depends on size: smaller firms are the most constrained by all obstacles. Nevertheless, Beck *et al.* (2005) identify a possible source of endogeneity, which rests on the fact that firms that do not grow are more likely to complain about the legal and financial institutions, and have a high perception of obstacles. In order to control for this potential simultaneity bias, Beck *et al.* (2005) instrument financial, legal and corruption obstacles with country indicators of level of financial development, efficiency of the legal system, and corruption. By interacting the instrumental variables with firm size, they confirm that small firms are more constrained than larger ones.

To summarize, the law and finance literature, together with the literature on finance and growth, shows that the legal protection of investors and law enforcement matter for cross-country differences in financial development and, in turn, that financial development explains differences in economic performance.

3.4.3 Panel Data Techniques

Although the pure cross-sectional instrumental variable approach is used extensively to disentangle the reverse causality problem, it has several limitations. For instance, in contrast to the legal view, Rajan - Zingales (2003) argue that the role of the legal environment in promoting economic growth is overemphasised since, for instance, it is not a precondition for a well developed financial system that a country was originally a British colony. The pure cross-sectional instrumental variable approach takes account only of the endogeneity of financial development, thereby treating the other explanatory variables as exogenous. This implies that the potential endogeneity of the other explanatory variables may lead to inappropriate inferences about the coefficient of financial development (Demirgüç-

²¹ In another study, GUIO L., JAPPELLI T., PADULA M., PAGANO M. (2004) investigate the determinants of regional financial development and find that in high-social-capital areas (with high levels of trust towards others), the use and availability of financial instruments is higher than in low-social-capital areas (with low levels of trust towards others).

Kunt - Levine, 2008). Moreover, pure cross-sectional instrumental variable regressions do not control explicitly for unobserved country-specific, fixed-effects, *i.e.* country-specific characteristics that are likely to influence country growth rates. This implies that the fixed-effects are likely to be part of the error term and their correlation with the explanatory variables will result in biased coefficients. Finally, pure cross-sectional instrumental variable regressions consider only cross-country variations of data, but do not take account of time-series variability.

In the finance and growth literature, Levine *et al.* (2000) and Beck *et al.* (2000) adopt the Generalised Method of Moment (GMM) dynamic panel estimator, introduced by Arellano - Bond (1991) and Arellano - Bover (1995), to address the statistical problems characterizing pure cross-sectional studies. Although both studies methodologically are based on the same techniques, they make different contributions to the literature. Specifically, the first assesses whether there is a link between financial development and long-run economic growth that is not driven by a simultaneity bias; the second provides evidence on the development of financial intermediaries and the sources of growth (physical capital accumulation, private saving rates, TFP). Levine *et al.* (2000) and Beck *et al.* (2000) use a panel of countries for the period 1960-1995, thereby considering the cross-country (between-country) and time-series (within-country) dimensions of all the variables included in the regressions.

The results obtained by Levine *et al.* (2000) and Beck *et al.* (2000) using dynamic panel data techniques, show a positive and strong link between financial development and both real *per-capita* GDP growth and TFP growth. Moreover, the results obtained using the methodology described are similar to those obtained using the Ordinary Least Squares (OLS) estimator. Therefore, in these studies, the impact of financial development on economic growth is not driven by simultaneity bias, nor omitted variables or omitted country-specific effects.

Beck - Levine (2004), based on Levine - Zervos (1998), show that both stock market and bank development matter for economic growth. They argue that, even though Levine - Zervos (1998) adopt the initial value of stock market and bank development to control for potential endogeneity bias, they do not completely address this issue. Indeed, «if the contemporaneous behaviour of the explanatory variable matters for current growth, (one runs) the risk of grossly mis-measuring the true explanatory variable by using initial value, which could bias the coefficient estimates. Using proper instruments for the contemporaneous values of the explanatory variables is therefore preferable to using initial values». (Beck - Levine, 2004, page 431). Therefore, Beck - Levine (2004) make some progress on this

issue by using the GMM dynamic panel estimator. The general results found by Beck - Levine (2004) support the view that stock market and bank development have a positive impact on the process of growth, which is not due to simultaneity bias, omitted variables or unobserved country-specific effects.

Finally, the GMM dynamic panel estimator is adopted by Rioja - Valev (2004) to show that the relationship between financial development and growth is not linear, but depends on the level of financial development. More specifically, based on a panel of 74 countries for the period 1960-1995, they show that the highest influence of financial development on economic growth is in countries with intermediate levels of financial development.²²

The panel data methodology improves on the pure cross-sectional estimators by addressing some relevant econometric issues. Nevertheless, the panel approach does not allow us to account for much heterogeneity across countries. In fact, by including country-specific fixed effects, it takes account of the intercept heterogeneity across-countries, but does not permit heterogeneity in the slope coefficients. An attempt to overcome this problem is provided by Favara (2003) who uses a procedure specifically designed to estimate panel data with varying slope coefficients and finds that the effect of financial development differs considerably across countries.

3.4.4 Time-Series Techniques

From a methodological point of view, a further improvement over pure cross-country analyses is given by the time-series approach, which has been extensively used in the empirical literature on finance and growth since the seminal work by Gupta (1984). The main differences between cross-country and time-series approaches rely on frequency of data and causality (Beck, 2008).

First, while the cross-country approach aims to estimate the average influence of financial development on economic growth, over a long time period, the time-series approach relies on higher-frequency data. In other words, time-series do not impose homogeneity in the relationship between finance and growth between countries, as in cross-country studies, but they allow to estimate a country-by-country coefficient based on country heterogeneity (Demetriades - Hussein, 1996).²³

²² Another empirical work assessing the non-linearity between financial development and growth is the one by TADESSE S. (2002). In particular, he finds that bank-based systems, *i.e.* financial systems in which the role of banks prevails, outperform in financially underdeveloped countries and in countries dominated by smaller firms, whereas market-based systems are more effective in financially developed countries and in countries dominated by larger firms.

²³ The use of data averaged over time may also mask the evolution of key variables and their causal relationships (ARESTIS P. - DEMETRIADES P.O., 1997).

Second, the pure cross-country approach does not properly address the issue of endogeneity, even though many researchers often include instrumental variables to deal with this issue, as described in Section 3.4.2. Indeed, as argued by Ahmed (1998) and Ericsson *et al.* (2001), the instrumental variable approach is not adequate to account for possible reverse causality if data are averaged over decades. On the other hand, the time-series approach relies on the concept of Granger causality (Granger, 1969). Specifically, a financial development indicator Granger-causes GDP if, controlling for lagged values of GDP, lagged values of financial development provide statistically significant information about the current value of GDP. Moreover, most studies test the bi-directional causality between finance and growth by adopting the vector autoregression (VAR) method. In the time-series context, researchers also address the issues of nonstationarity and cointegration of GDP and financial development series. The existence of cointegration has implications for the way in which causality is analysed. Indeed, in presence of cointegration the relationship is estimated through the vector error correction model (VECM). The main advantage of cointegration test is that «it provides evidence on the existence of a stable long-run linear relationship between economic and financial development, which is itself interesting from a theoretical perspective» (Demetriades - Hussein, 1996, page 391).

Several studies attempt to highlight the heterogeneity of the impact of finance on growth between countries, in a time-series context. Most studies distinguish between developed and developing countries. In a seminal paper, Jung (1986), using VAR in levels on *post-1960* annual time-series for financial and real variables, finds that the causality goes from financial development to growth in less developed countries and from growth to financial development in developed countries. Demetriades - Hussein (1996), using both VAR and VECM approaches, study 16 developing countries and find little support to the view that finance is a catalyst in the process of economic development. Instead, they find that the relationship between financial development and economic growth is bi-directional and country-specific since, in some countries, economic growth causes financial development. Similarly, Luintel - Kahn (1999), using the VAR technique, find consistent evidence for bi-directional causality across a smaller sample including 10 developing countries.

On the other hand, Artesis - Demetriades (1997) assess the relationship between finance and growth in developed countries finding that, even in this case, the relationship of interest is country-specific. Rousseau - Wachtel (1998) apply the VAR approach to five industrialized countries over the period between 1870-1929 and

find strong uni-directional links from finance to growth. They also estimate a VECM for each country and report evidence of an economically important long-run relationship between the two sectors. The results suggest that financial intermediation spurs output and that output does not directly feed back into intermediation. In a recent study, Hassan *et al.* (2011) employ VAR models to estimate whether financial development Granger-causes economic growth and *vice-versa*. The innovation of this study relies on the fact that, instead of analysing each country, Hassan *et al.* (2011) use geographic regions and income groups as defined by the World Bank. They find a two-way causality relationship between finance and growth for all regions but Sub-Saharan and East Asia & Pacific, and one-way causality from growth to finance Sub-Saharan Africa and East Asia & Pacific.

Attempts have also been made by researchers to examine the specific channel through which financial development contribute to economic growth in a time-series context. For example, Xu (2000) uses a VAR approach to estimate the effect of financial development on both investment and output in 41 countries between 1960 and 1993. The results show a strong evidence that financial development is a catalyst for growth, especially through the capital accumulation channel. Rousseau - Wachtel (2000) estimate panel VAR for 47 countries with annual data over the period 1980-1995. They find strong support for the argument that deep and liquid equity markets have a significant and persistent impact on economic performance. Bell - Rousseau (2001) use the VECM approach for *post-independence* India and find that financial intermediaries played a more important role in promoting investment than in increasing TFP, and interpret this as evidence for the presence of a factor accumulation channel. More recently, Rousseau - Vuthiapadadorn (2005) consider 10 Asian countries from 1950 to 2000. They use both VAR and VECM and find a strong uni-directional link from finance to investment for most of these countries, which is consistent with the factor accumulation channel. On the other hand, there is less support for a causal link from finance to the level of output. Altogether these results are consistent with the view that capital accumulation stimulates growth.

Even though the time-series models have been extensively adopted in the finance-growth literature, most studies are subject to omitted variable problems. As convincingly argued by Ang (2008), «the variables in the time-series models are always kept to a minimum in order to preserve degrees of freedom», (Ang, 2008, page 51). Moreover, the Granger causality does not attempt to control for omitted variable bias by directly including other variables or by controlling with instrumental variables, as in cross-country panel regressions, but by including a rich lag structure (Beck, 2008).

4. - Conclusions

The theoretical contributions on finance and growth highlight the mechanisms through which competitive financial intermediaries, such as banks and stock markets, are likely to influence economic development by addressing both savings and investment decisions. The common finding emerging from almost all the theoretical contributions is that banks and stock markets have a comparative advantage over an autarchic environment in mitigating risk and acquiring and revealing information on the projects to be financed. Therefore, financial institutions play a key role in channelling funds from savers, which are risk averse and have asymmetric information, into the most innovative and profitable investment projects. In that way, they stimulate physical capital accumulation and technological innovations, which are two standard channels of growth.

Although there seems to be some consensus in the debate about the existence of a positive influence going from finance to growth, some theoretical contributions show the existence of a two-way relationship (Greenwood - Jovanovic, 1990; Saint-Paul, 1992; Blackburn - Hung, 1998). According to this strand of literature, financial development not only makes countries grow faster, but it is also influenced by economic growth since as the economy develops, more financial services are required.

On the other hand, the empirical evidence tests whether the theoretical mechanisms are confirmed by country-, industry- and firm-level data, and suggests that both financial intermediaries and markets matter for growth, even when controlling for potential simultaneity bias. These studies make crucial advances to the work on causal inferences and provide valuable results about how finance contributes to economic performance. Specifically, studies adopting disaggregated data demonstrate that an increase in the credit provided by banks and other financial institutions to the private sector, and an increase in the size, liquidity and efficiency of financial markets, are likely to reduce financing constraints, especially for small firms (Guiso *et al.*, 2004a; Beck *et al.*, 2005).

In sum, both theoretical and empirical evidence find a positive and strong link between financial development and economic growth, suggesting that policy makers need to take into account the role of the financial system to sustain growth.

However, from an empirical point of view, there are different areas that need more research. The most relevant questions are related to the proxies used for financial development. Indeed, while the theory identifies different services that financial institutions are likely to provide, such as mitigation of risk and asym-

metric information, the proxies for financial development in the empirical contributions do not measure these services accurately. They measure instead the amount of credit provided by banks and other financial institutions, and the liquidity, size and efficiency of financial markets. More work is needed to reduce the gap between the theoretical role of financial institutions and the empirical indicators of financial development. Note that more recent research adopts more appropriate measures of financial development, such as the ease with which households can access external finance (World Bank, 2008).

Another concern is about the instrumental variables adopted for financial development. As is well known, in the empirical literature on finance and growth a country's legal origin, legal protection and law enforcement are used to instrument financial development. However, some recent contributions show that there are many other determinants that can be used to instrument financial development. For instance, Guiso (2004*b*) suggest that social capital, *i.e.* the trust that each individual displays towards others, plays an important role in determining the ease with which households can access credit, use cheques and invest in stocks. Consequently, more research is needed on the determinants of financial development.

Given the important role of the financial system in providing financial services to the firms in a country and, therefore, in stimulating economic growth, it would be interesting to conduct more research to identify the specific factors that influence the degree of development of financial systems and the mechanisms underlying the relationship of interest.

BIBLIOGRAPHY

- ACEMOGLU D. - ZILIBOTTI F., «Was Prometheus Unbound by Chance? Risk, Diversification, and Growth», *Journal of Political Economy*, vol. 105, 1997, pages 709-775.
- AGHION P. - HOWITT P., «A Model of Growth Through Creative Destruction», *Econometrica*, vol. 60, 1992, pages 323-351.
- AHMED S., «Comment on “The Legal Environment, Banks, and Long-Run Economic Growth”», *Journal of Money, Credit, and Banking*, vol. 30, 1998, pages 614-620.
- ALLEN F. - GALE D., «Diversity of Opinion and Financing of New Technologies», *Journal of Financial Intermediation*, vol. 8, 1999, pages 68-89.
- ANG J.B., «A Survey of Recent Developments in the Literature of Finance and Growth», *Journal of Economic Surveys*, vol. 22, 2008, pages 536-576.
- ARELLANO M. - BOND S., «Some Tests of Specification for Panel Data, Monte Carlo Evidence Application to Employment Equations», *Review of Economic Studies*, vol. 58, 1991, pages 277-297.
- ARELLANO M. - BOVER O., «Another Look at the Instrumental - Variable Estimation of Error Components Models», *Journal of Econometrics*, vol. 68, 1995, pages 29-52.
- ARESTIS P. - DEMETRIADES P.O., «Financial Development and Economic Growth: Assessing the Evidence», *Economic Journal*, vol. 107, 1997, pages 783-799.
- BAGEHOT W. (ed.), *Lombard Street*, Irwin, Homewood, IL, 1873.
- BECK T., «The Econometrics of Finance and Growth», Washington DC, The World Bank, *Policy Research Working Paper*, no. 4608, 2008.
- BECK T. - DEMIRGÜÇ-KUNT A. - LAEVEN L. - LEVINE R., «Finance, Firm Size and Growth», *Journal of Money, Credit and Banking*, vol. 40, 2008a, pages 1379-1405.
- BECK T. - DEMIRGÜÇ-KUNT A. - LEVINE R., «Law and Finance, Why Does Legal Origin Matter?», *Journal of Comparative Economics*, vol. 31, 2003, pages 653-675.
- BECK T. - DEMIRGÜÇ-KUNT A. - LEVINE R., «Financial Institutions and Markets Across-Countries and Over Time. Data and Analysis», Washington DC, The World Bank, *Policy Research Working Paper*, no. 4943, 2009.
- BECK T. - DEMIRGÜÇ-KUNT A. - MAKSIMOVIC V., «Financial and Legal Constraints to Firm Growth, Does Size Matter?», *Journal of Finance*, vol. 60, 2005, pages 137-177.
- BECK T. - DEMIRGÜÇ-KUNT A. - LEVINE R., «Financing Patterns Around the World: Are Small Firms Different?», *Journal of Financial Economics*, vol. 89, 2008, pages 467-487.
- BECK T. - LEVINE R., «Stock Markets, Banks, and Growth, Panel Evidence», *Journal of Banking and Finance*, vol. 28, 2004, pages 423-442.
- BECK T. - LEVINE R., «Legal Institutions and Financial Development», in MÉNARD C. - SHIRLEY M.M. (eds.), *Handbook of New Institutional Economics*, Dordrecht, Springer, 2005.
- BECK T. - LEVINE R. - LOAYZA N., «Finance and the Sources of Growth», *Journal of Financial Economics*, vol. 58, 2000, pages 261-300.

- BEKAERT G. - HARVEY C.R. - LUNDBLAD C., «Does Financial Liberalization Spur Growth?», *Journal of Financial Economics*, vol. 7, 2005, pages 3-55.
- BEKAERT G. - HARVEY C.R. - LUNDBLAD C. - SIEGEL S., «Global Growth Opportunities and Market Integration», *Journal of Finance*, vol. 62, 2007, pages 1081-1137.
- BELL C. - ROUSSEAU P.L., «Post-Independence India: A Case of Finance-Led Industrialization?», *Journal of Development Economics*, vol. 65, 2001, pages 153-175.
- BENCIVENGA V.R. - SMITH B.D., «Financial Intermediation and Endogenous Growth», *Review of Economics Studies*, vol. 58, 1991, pages 195-209.
- .,-,-, «Some Consequences of Credit Rationing in an Endogenous Growth Model», *Journal of Economic Dynamics and Control*, vol. 17, 1993, pages 97-122.
- BENCIVENGA V.R. - SMITH B.D. - STARR R.M., «Transactions Costs, Technological Choice, and Endogenous Growth», *Journal of Economic theory*, vol. 67, 1995, pages 53-177.
- BENFRATELLO L. - SCHIANTARELLI F. - SEMBENELLI A., «Banks and Innovation, Microeconomic Evidence on Italian Firms», *Journal of Financial Economics*, vol. 90, 2008, pages 197-217.
- BERNANKE B. - GERTLER M. - GILCHRIST S., «The Financial Accelerator in a Quantitative Business Cycle Framework», in TAYLOR J. - WOODFORD M. (eds.), *Handbook of Macroeconomics*, 1998, pages 1341-1393.
- BLACKBURN K. - HUNG V.T.Y., «A Theory of Growth, Financial Development, and Trade», *Economica*, vol. 65, 1998, pages 107-124.
- BOLTON P. - DEWATRIPONT M. (eds.), *Contract Theory*, Cambridge, MIT Press, 2005.
- BOOT A.W.A. - THAKOR A., «Can Relationship Banking Survive Competition?», *Journal of Finance*, vol. 55, 2000, pages 679-713.
- BOSE N. - COTHREN R., «Equilibrium Loan Contracts and Endogenous Growth in the Presence of Asymmetric Information», *Journal of Monetary Economics*, vol. 38, 1996, pages 363-376.
- .,-,-, «Asymmetric Information and Loan Contracts in a Neoclassical Growth Model», *Journal of Money, Credit, and Banking*, vol. 29, 1997, pages 423-439.
- BOYD J.H. - PRESCOTT E.C., «Financial Intermediary - Coalitions», *Journal of Economics Theory*, vol. 38, 1986, pages 211-232.
- DE LA FUENTE A. - MARIN J.M., «Innovation, Bank Monitoring, and Endogenous Financial Development», *Journal of Monetary Economics*, vol. 38, 1996, pages 269-301.
- DEMETRIADES P.O. - HUSSEIN K.A. «Does Financial Development Cause Economic Growth? Time-Series Evidence from Sixteen Countries», *Journal of Development Economics*, vol. 51, 1996, pages 387-411.
- DEMIRGÜÇ-KUNT A. - LEVINE R. «Stock Market Development and Financial Intermediaries, Stylized Facts», World Bank, *Economic Review*, vol. 10, 1996, pages 291-321.
- .,-,-, «Finance, Financial Sector Policies and Long-Run Growth», Washington DC, The World Bank, *Policy Research Working Paper*, no. 4469, 2008.

- DEMIRGÜÇ-KUNT A. - MAKSIMOVIC V., «Law, Finance, and Firm Growth», *Journal of Finance*, vol. 53, 1998, pages 2107-2137.
- DIAMOND D.W., «Financial Intermediation and Delegated Monitoring», *Review of Economic Studies*, vol. 51, 1984, pages 393-414.
- DIAMOND D.W. - DYBVIK P.H., «Bank Runs, Deposit Insurance, and Liquidity», *Journal of Political Economy*, vol. 91, 1983, pages 401-419.
- DIAMOND D.W. - VERRECCHIA R.E., «Optimal Managerial Contracts and Equilibrium Security Prices», *Journal of Finance*, vol. 37, 1982, pages 275-287.
- ERICSSON N.R. - IRONS J.S. - TRYON R.W., «Output and Inflation in the Long Run», *Journal of Applied Econometrics*, vol. 16, 2001, pages 241-253.
- FAVARA G., «An Empirical Reassessment of the Relationship Between Finance and Growth», *International Monetary Fund, Working Paper*, no. 03/123, 2003.
- FECHT F. - HUANG K.X.D. - MARTIN A., «Financial Intermediaries, Markets and Growth», *Journal of Money, Credit and Banking*, vol. 40, no. 4, 2008, pages 701-720.
- FISMAN R. - LOVE I., «Financial Development and Intersectoral Allocation, A New Approach», *Journal of Finance*, vol. 54, 2004, pages 2785-2805.
- FREIXAS X. - ROCHET J. (eds.), *Microeconomics of Banking*, Cambridge, MA, MIT Press, 1997.
- GALE D. - HELLWIG M., «Incentive-Compatible Debt Contracts, the One-Period Problem», *Review of Economic Studies*, vol. 52, 1985, pages 647-663.
- GALETOVIC A., «Specialization, Intermediation and Growth», *Journal of Monetary Economics*, vol. 38, 1996, pages 549-559.
- GOLDSMITH R.W. (ed.), *Financial Structure and Development*, New Haven, CT, Yale University Press, 1969.
- GRANGER C., «Investigating Causal Relations by Econometric Models and Cross-Spectral Methods», *Econometrica*, vol. 37, 1969, pages 424-438.
- GREENWOOD J. - JOVANOVIC B., «Financial Development, Growth, and the Distribution of Income», *Journal of Political Economy*, vol. 98, 1990, pages 1076-1107.
- GREENWOOD J. - SANCHEZ J.M. - WANG C., «Financing Development: The Role of Information Costs», *American Economic Review*, vol. 100, 2010a, pages 1875-1891.
- , «Quantifying the Impact of Financing Development on Economic Development», *National Bureau of Economic Research, Working Paper*, no. 15893, 2010b.
- GREENWOOD J. - SMITH B., «Financial Markets in Development, and the Development of Financial Markets», *Journal of Economic Dynamics and Control*, vol. 21, 1997, pages 145-181.
- GUIO L. - JAPPELLI T. - PADULA M. - PAGANO M., «Financial Market Integration and Economic Growth in the EU», *Economic Policy*, vol. 19, 2004, pages 523-577.
- GUIO L. - SAPIENZA P. - ZINGALES L., «Does Local Financial Development Matter?», *Quarterly Journal of Economics*, vol. 119, 2004a, pages 929-969.
- , «The Role of Social Capital in Financial Development», *American Economic Review*, vol. 94, 2004b, pages 526-556.

- GUPTA K.L. (ed.), *Finance and Economic Growth in Developing Countries*, London, Croom Helm, 1984.
- GURLEY J.G. - SHAW E.S., «Financial Aspects of Economic Development», *American Economic Review*, vol. 45, 1955, pages 515-538.
- HARRISON P. - SUSSMAN O. - ZEIRA J., «Finance and Growth, Theory and Evidence», Federal Reserve Boards, *Working Paper*, no. 99-35, 1999.
- HART O. (ed.), *Firms Contract and Financial Structure*, London, Oxford University Press, 1995.
- HASSAN K.A. - SANCHEZ B. - YU J.S., «Financial Development and Economic Growth: New Evidence from Panel Data», *The Quarterly Review of Economics and Finance*, vol. 51, 2011, pages 88-104.
- HICKS J. (ed.), *A theory of Economic History*, Oxford, Clarendon Press, 1969.
- HOLMSTROM B. - TIROLE J., «Market Liquidity and Performance Monitoring», *Journal of Political Economy*, vol. 101, 1993, pages 678-709.
- JENSEN M. - MURPHY K., «Performance Pay and Top Management Incentives», *Journal of Political Economy*, vol. 98, 1990, pages 225-263.
- JUNG W.S., «Financial Development and Economic Growth: International Evidence», *Economic Development and Cultural Change*, vol. 34, 1986, pages 333-346.
- KING R.G - LEVINE R., «Finance, Entrepreneurship and Growth, Theory and Evidence», *Journal of Monetary Economics*, vol. 32, 1993a, pages 513-542.
- .-,-.-, «Finance and Growth, Schumpeter Might be Right», *Quarterly Journal of Economics*, vol. 108, 1993b, pages 717-738.
- LA PORTA R. - LOPEZ-DE-SILANES F. - SHLEIFER A. - VISHNY, R., «Legal Determinants of External Finance», *Journal of Finance*, vol. 52, 1997, pages 1131-50.
- .-,-.-,-.-, «Law and Finance», *Journal of Political Economy*, vol. 106, 1998, pages 1113-1155.
- LEVINE R., «Stock Markets, Growth, and Tax Policy», *Journal of Finance*, vol. 46, 1991, pages 1445-1465.
- .-, «The Legal Environment, Banks, and Long-Run Economic Growth», *Journal of Money, Credit, and Banking*, vol. 30, 1998, pages 596-613.
- .-, «Law, Finance, and Economic Growth», *Journal of Financial Intermediation*, vol. 8, 1999, pages 36-67.
- .-, «Finance and Growth, Theory, Evidence, and Mechanisms», in AGHION P. - DURLAUF S. (eds.), *Handbook of Economic Growth*, Amsterdam, Netherlands, North-Holland, 2005.
- LEVINE R. - LOAYZA N. - BECK T., «Financial Intermediation and Growth, Causality and Causes», *Journal of Monetary Economics*, vol. 46, 2000, pages 31-77.
- LEVINE R. - ZERVOS S., «Stock Markets, Banks, and Economic Growth», *American Economic Review*, vol. 88, 1998, pages 537-558.

- LOVE I., «Financial Development and Financing Constraint, International Evidence From the Structural Investment Model», *Review of Financial Studies*, vol. 16, 2003, pages 765-791.
- LUIINTEL K.B. - KHAN M., «A Quantitative Reassessment of the Finance-Growth Nexus: Evidence from a Multivariate VAR», *Journal of Development Economics*, vol. 60, 1999, pages 381-405.
- MCKINNON R.I., *Money and Capital in Economic Development*, Washington DC, Brookings Institution, 1973.
- PANG J. - WU H., «Financial Markets, Financial Dependence, and the Allocation of Capital», *Journal of Banking & Finance*, vol. 33, 2009, pages 810-818.
- RAJAN R.G. - ZINGALES L., «Financial Dependence and Growth», *American Economic Review*, vol. 88, 1998, pages 559-586.
- .-.-. *Saving Capitalism from the Capitalists*, New York, Crown Business Division of Random House, 2003.
- RIOJA F. - VALEV N., «Does One Size Fit All? A Re-examination of the Finance and Growth Relationship», *Journal of Development Economics*, vol. 74, 2004, pages 429-447.
- ROBINSON J., *The Rate of Interest and Other Essays*, London, Macmillan, 1952.
- ROUSSEAU P.L., «The Permanent Effects of Innovation on Financial Depth: Theory and US Historical Evidence from Unobservable Components Models», *Journal of Monetary Economics*, vol. 42, 1998, pages 387-425.
- ROUSSEAU P.L. - VUTHIPADADORN D., «Finance, Investment, and Growth: Time Series Evidence from 10 Asian Economies», *Journal of Macroeconomics*, vol. 27, 2005, pages 87-106.
- ROUSSEAU P.L. - WACHTEL P., «Financial Intermediation and Economic Performance: Historical Evidence from Five Industrial Countries», *Journal of Money, Credit, and Banking*, vol. 30, 1998, pages 657-78.
- .-.-. «Equity Market and Growth, Cross Country Evidence on Timing and Outcomes 1980-1995», *Journal of Banking and Finance*, vol. 24, 2000, pages 1933-1957.
- .-.-. «What is Happening to the Impact of Financial Deepening of on Economic Growth», *Economic Inquiry*, vol. 49, 2011, pages 276-288.
- SAINT-PAUL G., «Technological Choice, Financial Markets and Economic Development», *European Economic Review*, vol. 36, 1992, pages 763-781.
- SCHUMPETER J.A., *The Theory of Economic Development*, Cambridge MA, Harvard University Press, 1912.
- SHAW E.S. (ed.), *Financial Deepening in Economic Development*, London, Oxford University Press, 1973.
- SHLEIFER A. - VISHNY R.W., «A Survey of Corporate Governance», *Journal of Finance*, vol. 52, 1997, pages 737-783.

- STIGLITZ J. - WEISS A., «Incentive Effects of Terminations, Applications to Credit and Labor Markets», *American Economic Review*, vol. 73, 1983, pages 912-927.
- SUSSMAN O., «A Theory of Financial Development», in GIOVANNINI A. (ed.), *Finance and Development, Issues and Experience*, Cambridge University Press, 1993.
- TADESSE S., «Financial Architecture and Economic Performance, International Evidence», *Journal of Financial Intermediation*, vol. 11, 2002, pages 429-454.
- TOWNSEND R.M., «Optimal Contracts and Competitive Markets with Costly State Verification», *Journal of Economic Theory*, vol. 21, 1979, pages 265-293.
- WACHTEL P., «Growth and Finance: What Do We Know and How Do We Know It?», *International Finance*, vol. 4, 2001, pages 335-362.
- WILLIAMSON S.D., «Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing», *Journal of Monetary Economics*, vol. 18, 1986, pages 159-179.
- WORLD BANK, *Finance for All? Policies and Pitfalls in Expanding Access*, Washington, DC, The World Bank, 2008.
- WURGLER J., «Financial Markets and the Allocation of Capital», *Journal of Financial Economics*, vol. 58, 2000, pages 187-214.
- XU Z., «Financial Development, Investment, and Economic Growth», *Economic Inquiry*, vol. 38, 2000, pages 331-344.

PAROLE CHIAVE

Immigrazione

Tommaso Frattini*

Università degli Studi di Milano, LdA, CReAM e IZA

In this paper I discuss the economic effects of immigration for destination countries, with a focus on labour market adjustments. First, I show the potential effects of immigration on the labour market in a stylized one-output model, and discuss some extensions. Then I focus on empirical estimations, critically presenting strengths and weaknesses of the empirical strategies that are commonly used in the literature. In the last part I review recent empirical evidence on the effects of immigration in Italy.

[JEL Classification: F22; J31; J61].

Keywords: immigration; labour market impact of immigration; wages.

* <tommaso.frattini@unimi.it>, Dipartimento di Economia, Management e Metodi Quantitativi. Desidero ringraziare Carlo Devillanova, Francesco Fasani, Elena Meschi e Sauro Mocetti per commenti e suggerimenti su versioni preliminari di questo testo. Tutti gli errori ed omissioni sono ovviamente di mia esclusiva responsabilità.

1. - Introduzione

Da sempre le migrazioni hanno fatto parte della storia umana, con individui e popolazioni che si spostano o per sfuggire a disastri naturali e persecuzioni politiche o religiose (rifugiati e profughi), o per il desiderio di migliorare le proprie condizioni di vita (migranti economici). Negli anni recenti tuttavia la diminuzione drastica nei costi dei trasporti e l'aumento delle disuguaglianze di reddito tra paesi (si veda per esempio Milanovic, 2005) hanno portato a un aumento consistente nel numero di persone che annualmente decidono di trasferirsi in un altro paese¹.

TAV. 1

IMMIGRATI NEL MONDO, 2010

	<i>stock</i>	<i>% popolazione</i>	<i>% rifugiati tra popolazione immigrata</i>
Africa	19.263.183	1,86	13,33
Asia	61.323.979	1,47	17,73
Europa	69.819.282	9,53	2,28
America Latina e Caraibi	7.480.267	1,27	7,08
Nord America	50.042.408	14,23	1,45
Oceania	6.014.693	16,78	0,99
Paesi meno sviluppati	86.232.341	1,52	16,21
Paesi più sviluppati	127.711.471	10,32	1,86
Mondo	213.943.812	3,10	7,64

La prima colonna riporta lo *stock* di immigrati per area. La seconda colonna riporta lo *stock* di immigrati in percentuale della popolazione totale nell'area. La terza colonna riporta la percentuale di rifugiati tra la popolazione immigrata.

Paesi meno sviluppati: Africa, Asia (eccetto il Giappone), America Latina, Melanesia, Micronesia, e Polinesia.

Paesi più sviluppati: Australia, Europa, Giappone, Nord America e Nuova Zelanda.

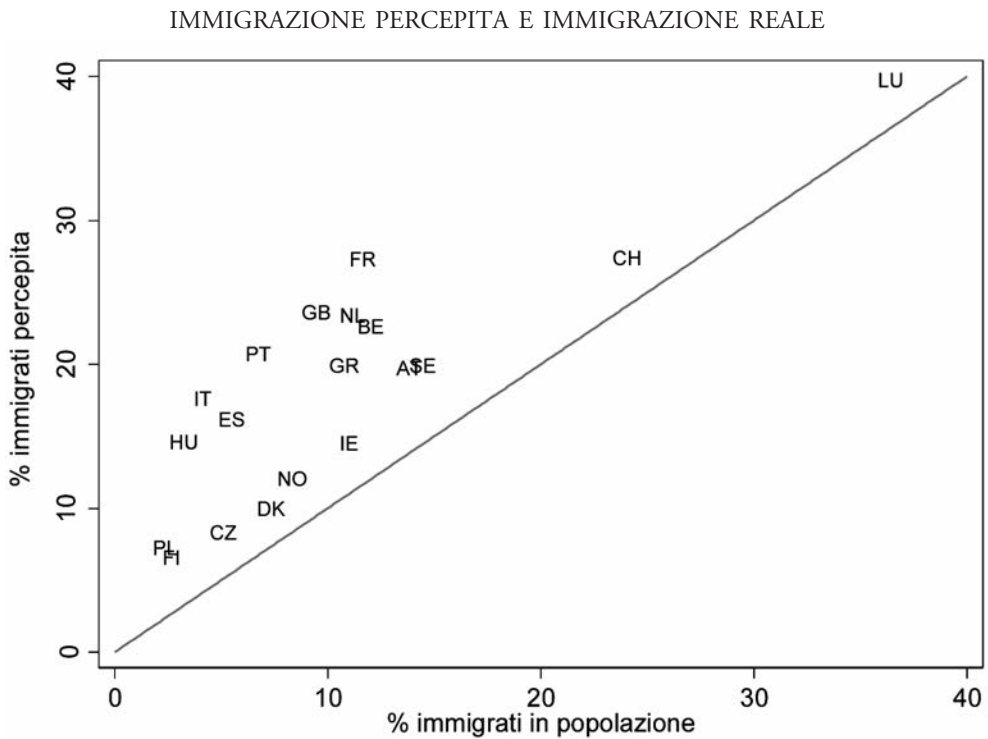
Fonte: elaborazione su dati UNITED NATIONS database, *POP/DB/MIG/Stock/Rev.2008*.

Come si vede nella Tavola 1, nel 2010 quasi 214 milioni di persone vivevano in un paese diverso da quello in cui erano nate. In altre parole, oltre il 3% della popolazione mondiale è costituita da migranti. L'immigrazione è un fenomeno che interessa tutte le aree del mondo, anche se la maggior parte dei migranti, sia in numero assoluto che in percentuale della popolazione, è concentrato nei paesi più sviluppati: oltre il 10% della popolazione nelle aree più ricche del mondo è costituita da migranti, mentre nei paesi più poveri gli immigrati costituiscono

¹ In quanto segue mi occupo esclusivamente delle migrazioni internazionali, che avvengono fra stati, tralasciando invece le migrazioni interne.

solo l'1,5% della popolazione. Inoltre, non sorprendentemente, le motivazioni economiche sono alla base della migrazione per una frazione più alta dei migranti nei paesi ricchi, mentre la percentuale di migranti per motivi non economici è otto volte superiore fra gli immigrati nei paesi poveri (16%) rispetto a quelli nei paesi ricchi (2%).

GRAF. 1



Fonti: Immigrazione percepita: EUROPEAN SOCIAL SURVEY, 2002; Immigrazione reale: OECD, Database on Immigrants in OECD Countries, basato su dati dei censimenti della popolazione nazionale intorno all'anno 2000.

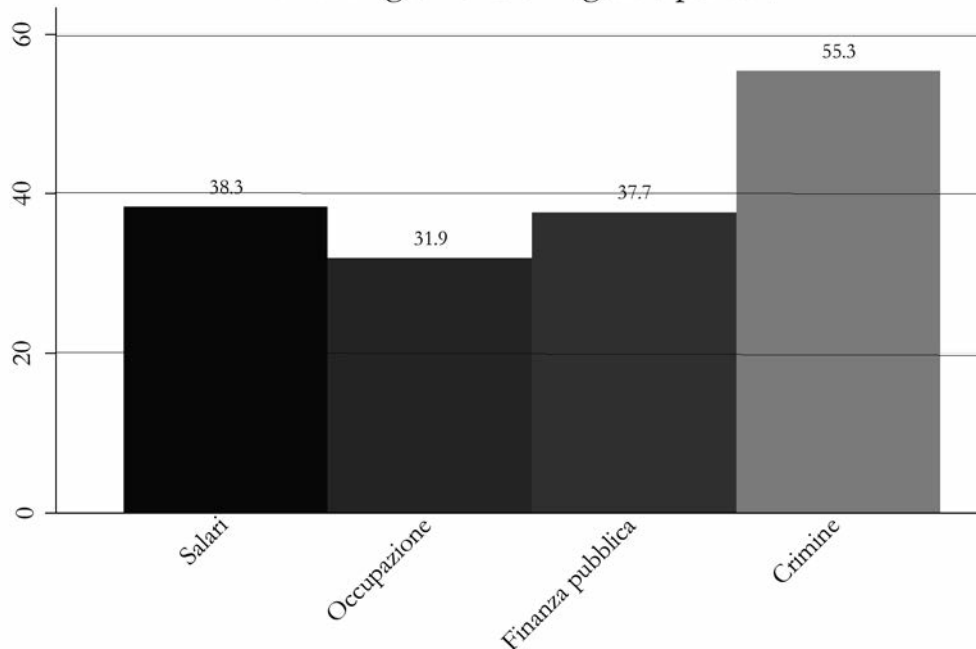
L'aumento dell'immigrazione e la difficoltà di buona parte dei cittadini dei paesi di destinazione ad abituarsi alla convivenza con persone provenienti da paesi diversi, e portatrici quindi di diversi usi e costumi, oltre che spesso caratterizzati da tratti somatici che li distinguono dai cittadini nativi, hanno suscitato, e continuano a suscitare, timori che la migrazione internazionale possa portare a un abbassamento della qualità della vita per i cittadini autoctoni. È significativo, ad esempio, che i cittadini dei paesi di destinazione tendano ovunque a sovrastimare

la dimensione della popolazione straniera nel proprio paese, come mostriamo nel Grafico 1, basato su dati della *European Social Survey 2002* e del DIOC (*Database on Immigrants in OECD Countries*). Se le stime dei cittadini dei paesi di destinazione fossero accurate, i paesi dovrebbero giacere lungo la linea inclinata di 45 gradi che divide il grafico: in tutti i paesi invece le stime sono sistematicamente superiori al dato reale, come testimoniato dal fatto che tutti i punti nel grafico sono al di sopra della bisettrice. Se nel caso di paesi con una lunga esperienza migratoria, come la Gran Bretagna, la Francia, o il Belgio, questa sovrastima può almeno in parte essere dovuta al fatto che i nativi classificano erroneamente come “immigrati” anche i figli di cittadini stranieri nati nel paese di destinazione, per paesi di più recente immigrazione, come la Spagna o l’Italia, nei quali i cosiddetti “immigrati di seconda generazione” sono assai poco numerosi, la sovrastima sembra decisamente imputabile ad un errore di percezione.

GRAF. 2

CONSEGUENZE PERCEPITE DELL’IMMIGRAZIONE

L’immigrazione e’ negativa per ... ?



I dati si riferiscono ai cittadini nativi di Austria, Belgio, Svizzera, Germania, Danimarca, Spagna, Finlandia, Francia, Gran Bretagna, Grecia, Irlanda, Italia, Lussemburgo, Olanda, Norvegia, Portogallo, Svezia.

Fonte: EUROPEAN SOCIAL SURVEY, 2002.

Che l'immigrazione sia fonte di preoccupazione tra i cittadini europei è confermato anche dai dati riportati nel Grafico 2, che si basano sulle risposte dei cittadini dei paesi dell'Europa occidentale alle domande poste nella *European Social Survey* del 2002. Quasi quattro cittadini europei su dieci si dicono d'accordo o molto d'accordo con l'affermazione che l'immigrazione causi un abbassamento dei salari medi, uno su tre è convinto che gli immigrati tendano a sottrarre posti di lavoro ai cittadini nativi, il 38% crede che gli immigrati costituiscano un costo per le finanze pubbliche, mentre addirittura più di un cittadino su due ritiene che i problemi legati alla criminalità nel proprio paese siano peggiorati a causa dell'immigrazione.

Quanto sono fondati questi timori? Che conseguenze ha l'immigrazione per i cittadini dei paesi ospiti? Trovare una risposta a queste domande è fondamentale per poter gestire al meglio le complesse problematiche che l'immigrazione pone per i paesi di destinazione, ed è importante, anche per poter meglio organizzare politiche di integrazione sociale e lavorativa degli immigrati, che i decisori politici siano in grado di distinguere i dati reali dal sentire comune. Il timore di conseguenze avverse sul mercato del lavoro o di maggiori costi per le finanze pubbliche è infatti spesso alla base delle richieste di politiche migratorie più restrittive (si vedano ad esempio Facchini e Mayda, 2009 e 2011).

La ricerca economica degli ultimi trent'anni ha affrontato questo tipo di argomenti, sia dal punto di vista teorico che empirico, grazie anche alla crescente disponibilità di *dataset* rappresentativi per un ampio numero di paesi. I risultati di questi studi, a volte controintuitivi, permettono di fare luce su alcuni di questi temi, spesso sfatando dei luoghi comuni assai diffusi che non risultano però confermati da un'analisi rigorosa dei dati.

Lo scopo di questo articolo è di presentare gli strumenti di base utilizzati dagli economisti per affrontare queste questioni e le principali conclusioni raggiunte. L'immigrazione è un fenomeno complesso, che trasforma l'economia e la società tanto dei paesi di origine che dei paesi di destinazione dei migranti sotto molteplici aspetti, che sono studiati con prospettive diverse e spesso complementari da diverse scienze sociali. In questo articolo forniremo una panoramica dell'analisi economica dell'immigrazione, ma anche la letteratura in ambito giuridico, demografico, sociologico e antropologico è assai ricca e interessante. Limiti di spazio – e di competenza dell'autore – impediscono però di affrontare i contributi delle altre scienze sociali in questa sede.

Ci soffermeremo soprattutto sugli effetti dell'immigrazione sul mercato del lavoro, per una duplice ragione. Innanzitutto, è questo il campo sul quale si è tra-

dizionalmente concentrata la ricerca economica, e per il quale esistono il maggior numero di lavori sia teorici che empirici. In secondo luogo, il presentare gli strumenti adottati per lo studio del mercato del lavoro ha anche uno scopo “didattico”, poiché le tecniche di analisi sviluppate originariamente per studiare gli effetti dell’immigrazione su salari e occupazione sono poi state adottate anche per lo studio di altre conseguenze economiche e sociali dell’immigrazione. Presenteremo innanzitutto, nella sezione 2, un semplice modello teorico che ci guiderà nello strutturare il ragionamento sui canali attraverso i quali l’immigrazione può avere conseguenze sul mercato del lavoro, e su quali fattori ne influenzino gli effetti. La sezione 3 espone invece le principali tecniche di analisi empirica, discutendone i vantaggi relativi e gli ambiti di applicazione, mentre nella sezione 4 forniremo una breve rassegna dei contributi più importanti nella letteratura internazionale. Nella sezione 5, infine, ci concentriamo sull’Italia, presentando una rassegna quanto più aggiornata e completa possibile degli studi condotti in ambito economico per il nostro paese, al quale è anche dedicato una specifica appendice che presenta alcuni fatti chiave sull’immigrazione. La sezione 6 conclude.

Prima di procedere, è opportuno sottolineare che le migrazioni possono avere conseguenze significative anche per i paesi di origine. Si pensi al cosiddetto fenomeno del *brain drain*, ossia l’emigrazione di persone altamente qualificate, che può comportare una significativa riduzione del capitale umano nei paesi di origine, con conseguenze potenzialmente dannose per lo sviluppo economico dei paesi più poveri, dove il capitale umano è più scarso. Allo stesso tempo, la possibilità di emigrare può far aumentare gli incentivi all’investimento in capitale umano (*brain gain*), mentre la mobilità internazionale dei lavoratori qualificati può creare delle esternalità positive al paese di origine, favorendo ad esempio l’adozione di nuove tecnologie. O si pensi all’importanza delle rimesse dei migranti verso i paesi di origine, che rappresentano per alcuni paesi una sostanziale frazione del PIL: secondo i dati della Banca Mondiale le rimesse degli emigrati sono pari nel 2010 allo 0,74% del PIL mondiale, e contribuiscono per il 5,67% del PIL nei paesi poveri. Per esempio, le rimesse costituiscono l’11% del PIL nelle Filippine e in Bangladesh, il 22% del PIL ad Haiti e il 24% del PIL della Moldova. Comprendere le principali determinanti di questi flussi, capire chi ne siano i principali destinatari nei paesi di origine, e studiare in che modo le rimesse vengano utilizzate è un’importante area di ricerca per gli scienziati sociali. In questa sede ci concentreremo solo sui paesi di destinazione, mentre rimandiamo a Docquier e Rapoport (2012) per una rassegna della letteratura su *brain drain* e a Yang (2011) per una rassegna della letteratura economica sulle rimesse.

2. - Immigrazione e mercato del lavoro: teoria

Gli immigrati sono spesso visti come una minaccia per i salari e l'occupazione dei lavoratori dei paesi di destinazione. La teoria economica ci può fare da guida nell'analisi della questione dal punto di vista teorico.²

2.1 *Modello base*

Consideriamo un'economia stilizzata, in cui si produce un solo bene con una tecnologia a rendimenti di scala costanti (ossia una tecnologia in cui se si raddoppiano i fattori di produzione raddoppia anche il prodotto), utilizzando come fattori di produzione capitale, lavoro qualificato, e lavoro non qualificato. I lavoratori possono essere nativi o immigrati. Assumiamo che nativi e immigrati con lo stesso tipo di qualifiche siano perfettamente interscambiabili nella produzione. Facciamo inoltre due ipotesi semplificatrici, che rimuoveremo più avanti:

- 1) l'offerta di capitale è infinitamente elastica, ossia il capitale si muove immediatamente e senza costi tra paesi: questo implica che il prezzo del capitale, cioè il tasso di interesse, è fissato sul mercato mondiale e non dipende dalle condizioni della nostra economia
- 2) l'offerta di lavoro tanto dei lavoratori qualificati che dei lavoratori non qualificati è rigida, ossia tutti sono disposti a lavorare per qualsiasi livello di salario.

Assumiamo che ci siano due periodi. Nel periodo 1, nell'economia esistono solo lavoratori nativi. Nel periodo 2 invece la nostra economia riceve un flusso migratorio che immaginiamo per semplicità interamente composto da lavoratori non qualificati³. L'aumento dell'offerta di lavoro non qualificato porta a un eccesso di offerta per tale tipo di lavoro, al livello di salario iniziale. L'aumento dell'offerta di lavoro non qualificato permette alle imprese in questa economia di soddisfare la loro domanda di quel tipo di lavoro a un salario inferiore rispetto al periodo 1. La diminuzione dei salari dei lavoratori non qualificati fa a sua volta aumentare la domanda di lavoro non qualificato cosicché alla fine del periodo 2 l'economia ritorna a un nuovo equilibrio in cui tutti i lavoratori non qualificati, sia immigrati che nativi, sono occupati, ma a un salario inferiore a quello del periodo 1. L'immigrazione ha quindi un effetto negativo sui salari dei lavoratori

² La trattazione teorica si basa su DUSTMANN C., GLITZ A. e FRATTINI T. (2008) cui si rimanda, insieme a BORJAS G.I. (1999), per maggiori dettagli. Per un approccio più tecnico si vedano anche, ad esempio, ALTONJI J.G. e CARD D. (1991) e BORJAS G.J. (2009).

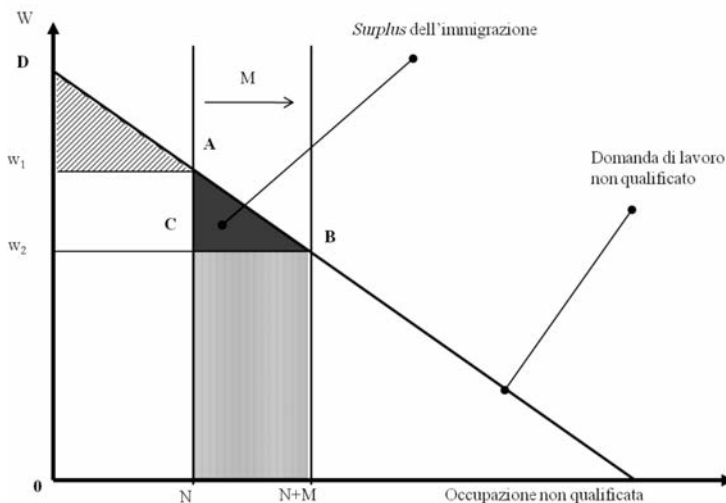
³ Il ragionamento che sviluppiamo vale specularmente a parti invertite nel caso di immigrazione qualificata.

non qualificati. Che cosa succede invece sul mercato del lavoro qualificato? In termini relativi l'offerta di lavoro qualificato è diventata più scarsa, il che spinge il salario dei lavoratori qualificati verso l'alto. In questa economia quindi l'immigrazione non qualificata porta a una diminuzione dei salari dei lavoratori non qualificati e a un aumento dei salari dei lavoratori qualificati. La somma algebrica dei due effetti è teoricamente positiva – il cosiddetto “*surplus* dell'immigrazione” – poiché la remunerazione che va ai lavoratori immigrati è solo una frazione dei beni e servizi che hanno contribuito a produrre con il loro lavoro.

Abbiamo illustrato graficamente questo ragionamento nel Grafico 3, che rappresenta il mercato del lavoro non qualificato. Sull'asse verticale abbiamo riportato i salari dei lavoratori non qualificati e sull'asse orizzontale il numero di occupati. Nel periodo pre-migrazione, l'economia è in equilibrio al punto *A*, dove tutti i lavoratori nativi (*N*) sono impiegati con salario w_1 . In questo equilibrio iniziale il prodotto totale dell'economia è rappresentato dall'area *D-A-N-0*, e ai lavoratori non qualificati va una quota di prodotto rappresentata dall'area $w_1-A-N-0$, con il rimanente triangolo *D-A- w_1* suddiviso tra gli altri fattori di produzione, capitale e lavoro qualificato. Nel periodo 2 entrano nell'economia *M* lavoratori immigrati, che provocano uno spostamento della curva di offerta di lavoro non qualificato da *N* a *N+M*, e un conseguente slittamento lungo la curva di domanda fino al nuovo equilibrio *B* dove i salari sono diminuiti a w_2 .

GRAF. 3

EFFETTO DELL'IMMIGRAZIONE SUL MERCATO DEL LAVORO NON QUALIFICATO



Nel nuovo equilibrio, la quota di prodotto totale che va ai lavoratori nativi non qualificati è diminuita di un ammontare rappresentato dalla superficie del rettangolo ($w_1-A-C-w_2$), che – data la nostra ipotesi di offerta di capitale infinitamente elastica – è stata interamente trasferita ai lavoratori qualificati. Ma l’immigrazione non ha un semplice effetto distributivo: l’aumento dell’offerta di lavoro non qualificato fa aumentare le dimensioni dell’economia di un ammontare rappresentato dall’area $A-B-(N+M)-N$. Gli immigrati, tuttavia, ricevono solo parte di questo prodotto aggiuntivo, rappresentato dall’area $C-B-(N+M)-N$, creando un *surplus*, dato dall’area $A-B-C$, del quale si appropriano i lavoratori qualificati. In questa economia l’immigrazione ha quindi sia un effetto distributivo, con una perdita per i lavoratori non qualificati a favore dei lavoratori qualificati, sia un effetto positivo sul prodotto totale, che è più che sufficiente a compensare la perdita subita da un gruppo di lavoratori.

Più in generale il reddito *pro capite* della popolazione autoctona aumenta come conseguenza della migrazione, ma i guadagni della migrazione sono distribuiti in modo diseguale: i salari medi aumentano grazie al *surplus* dell’immigrazione, ma i salari dei lavoratori che competono con gli immigrati diminuiscono. È importante sottolineare che la condizione necessaria perché, in un’economia come questa, l’immigrazione abbia effetti sui salari è che la composizione dei lavoratori immigrati sia diversa da quella dei nativi, ovvero che l’immigrazione provochi dei cambiamenti nell’offerta relativa di lavoro qualificato e non qualificato. Se l’immigrazione non altera l’offerta di lavoro relativa, i salari rimangono invariati e l’unico effetto dell’immigrazione è di aumentare le dimensioni dell’economia (effetto di scala) senza toccare i salari relativi (effetto distributivo).

Il modello illustrato sopra è estremamente stilizzato e basato su una serie di ipotesi semplificatrici, ma permette di cogliere le dinamiche principali dell’effetto dell’immigrazione sul mercato del lavoro. Vediamo che cosa succede rimuovendo alcune di queste ipotesi.

Innanzitutto, abbiamo ipotizzato (ipotesi 1) che l’offerta di capitale sia infinitamente elastica. Questa ipotesi, non implausibile per piccole economie aperte come l’Italia nel medio-lungo periodo, garantisce che i possessori di capitale non traggano alcun vantaggio dall’immigrazione, e che i vantaggi vengano interamente goduti dai lavoratori che non competono con gli immigrati (i lavoratori qualificati nel nostro esempio). Se invece assumiamo che l’offerta di capitale sia almeno in parte rigida, allora il tasso di interesse si può muovere in risposta ai cambiamenti nell’offerta di lavoro. Ciò implica che una parte del *surplus* dell’immigrazione andrà ai proprietari del capitale invece che ai lavoratori qualificati, e gli effetti sui

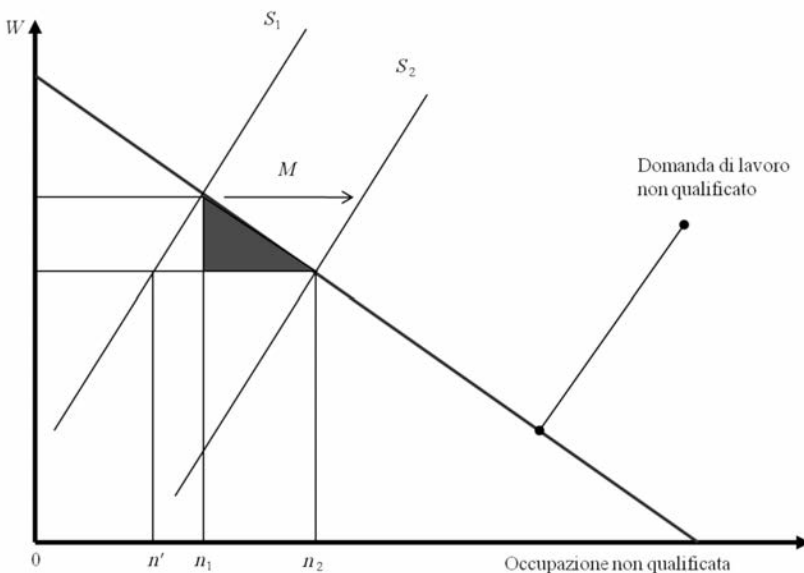
salari medi dei nativi potrebbero essere positivi o negativi, a seconda dell'elasticità dell'offerta di capitale; rimane in ogni caso vero che l'immigrazione non qualificata farà diminuire i salari dei lavoratori non qualificati e aumentare quelli dei lavoratori qualificati. Tanto maggiore è l'elasticità dell'offerta di capitale, tanto minori saranno gli effetti negativi sui salari medi dei nativi.

La seconda ipotesi semplificatrice che abbiamo introdotto è che l'offerta di lavoro sia perfettamente rigida, ossia che sia rappresentata da una linea verticale nel Grafico 3. Se rimuoviamo questa assunzione, e assumiamo che l'offerta di lavoro sia almeno in parte elastica, l'immigrazione avrà effetti non solo sui salari, ma anche sull'occupazione dei nativi. Un aumento dell'offerta di lavoro, infatti, porterà alcuni lavoratori nativi non essere più disposti a lavorare al nuovo salario di equilibrio più basso.

Abbiamo rappresentato questa situazione nel Grafico 4, dove indichiamo con n_1 l'occupazione prima dell'immigrazione, con n_2 l'immigrazione totale dopo l'immigrazione, e con n' il numero di nativi occupati dopo l'immigrazione. Il segmento $n' - n_1$ rappresenta i lavoratori nativi che rimangono disoccupati. Da notare che quanto più alta è l'elasticità dell'offerta di lavoro (ossia, quanto più orizzontale è la curva di offerta), tanto minore è il *surplus* dell'immigrazione

GRAF. 4

EFFETTO DELL'IMMIGRAZIONE CON OFFERTA DI LAVORO ELASTICA



2.2 Estensioni

Il modello che abbiamo presentato è alla base di gran parte delle analisi empiriche sugli effetti dell'immigrazione sul mercato del lavoro, e formalizza inoltre ciò che molti immaginano l'immigrazione possa comportare per il mercato del lavoro dei paesi di destinazione, come indicano anche i dati presentati nel Grafico 2.

Tuttavia tale modello, ipotizzando tecnologia fissa e un'economia che produce un solo bene, trascura altri meccanismi attraverso i quali l'economia può adeguarsi ai cambiamenti nell'offerta di lavoro, che non comportano cambiamenti nei salari o nei livelli di occupazione.

Un primo importante meccanismo è la possibilità per l'economia di modificare la composizione del prodotto totale. Consideriamo il caso di un'economia che produce diversi beni, che vengono commerciati internazionalmente, così che il loro prezzo viene determinato sul mercato internazionale, ed è indipendente dalle scelte produttive e di consumo interne. Supponiamo per semplicità che l'economia sia composta da due soli settori: un settore "avanzato" che utilizza prevalentemente manodopera qualificata, e un settore "tradizionale" che utilizza prevalentemente manodopera non qualificata. Come reagisce un'economia di questo tipo all'immigrazione di lavoratori non qualificati? L'effetto immediato dell'immigrazione è, come nel caso precedente, di deprimere i salari dei lavoratori non qualificati e di far crescere quelli dei lavoratori qualificati. Questo cambiamento nei salari relativi porterà a una diminuzione dei profitti nel settore avanzato e a un aumento dei profitti nel settore tradizionale. L'aumento della profittabilità nel settore tradizionale spingerà a un aumento della produzione in questo settore, con conseguente aumento della domanda di lavoro non qualificato, il che tenderà a sua volta a spingere verso l'alto i salari dei lavoratori non qualificati. Nel lungo periodo quindi il processo di spostamento della produzione dal settore avanzato a quello tradizionale riporterà i salari al livello dell'equilibrio pre-immigrazione.

In un'economia multisettoriale, dunque, cambiamenti nell'offerta relativa di lavoro possono essere riassorbiti, per lo meno nel lungo periodo, attraverso cambiamenti nelle quantità di beni prodotte da ciascun settore, senza alterare i salari o il tasso di occupazione d'equilibrio. Leamer e Levinsohn (1995) denominano questo meccanismo, mutuato dalla teoria del commercio internazionale, e in particolare dal teorema di Heckscher-Ohlin, "ipotesi di insensibilità del prezzo dei fattori".

Un'ulteriore meccanismo attraverso il quale l'economia può adattarsi ai cambiamenti nell'offerta relativa di lavoro indotti dall'immigrazione, senza alterare né i livelli di occupazione e i salari di equilibrio, né la composizione settoriale

della produzione, è il cambiamento tecnologico. L'economia può cioè continuare, anche nel periodo *post*-migrazione, a produrre gli stessi beni, ma utilizzando più intensamente i fattori produttivi la cui offerta relativa è stata aumentata dall'immigrazione. A livello aneddotico, tale situazione può essere esemplificata dall'industria del vino in Australia e in California. Entrambi i paesi sono produttori ed esportatori di vino, e condividono diverse similarità culturali e istituzionali, ma le loro industrie vinicole sono organizzate in maniera radicalmente diversa. Mentre l'industria vinicola californiana è ad alta intensità di lavoro manuale, potendo contare su un'abbondante offerta di lavoro non qualificato garantita dall'immigrazione messicana, l'industria del vino australiana è altamente meccanizzata, poiché la manodopera non qualificata in Australia è assai più scarsa, e dunque più cara, a causa anche delle politiche migratorie restrittive.⁴

Anche in un modello ad un solo settore e a tecnologia fissa, tuttavia, si può introdurre un altro margine di aggiustamento dell'economia ai cambiamenti nell'offerta di lavoro indotti dall'immigrazione, se si assume che immigrati e nativi, anche a parità di livello di istruzione, non siano perfetti sostituti nella produzione. Consideriamo ad esempio il caso di immigrazione non qualificata. Pur facendo aumentare l'offerta relativa di lavoro non qualificato, è probabile che il tipo di competenze possedute dagli immigrati sia diverso da quelli dei nativi con lo stesso livello di istruzione. In particolare, gli immigrati avranno probabilmente competenze linguistiche inferiori a quelle dei nativi, mentre avranno capacità manuali simili a quelle dei lavoratori autoctoni. In questo caso, gli immigrati hanno quindi un vantaggio comparato in occupazioni che richiedono principalmente lavoro manuale, mentre i nativi hanno un vantaggio comparato in occupazioni che richiedono maggiori competenze comunicative. Immigrati e nativi tenderanno quindi a specializzarsi in tipi di occupazione diverse: i lavoratori nativi non qualificati lasceranno le occupazioni che richiedono lo svolgimento di compiti più manuali per specializzarsi in occupazioni dove le competenze linguistiche e di comunicazione sono più importanti. A loro volta, la domanda di lavoro per professioni ad alta intensità di compiti comunicativi crescerà grazie all'aumento nell'offerta di lavoro manuale, ad esse complementare, spingendo verso l'alto i salari in queste occupazioni. La specializzazione di immigrati e nativi in diverse occupazioni è quindi un ulteriore canale attraverso il quale l'economia può as-

⁴ Si vedano ad esempio QUISPE-AGNOLI M. e ZAVODNY M. (2002); GANDAL N., HANSON G.H. e SLAUGHTER M.J. (2004); DUSTMANN C. e GLITZ A. (2011); GONZALEZ L. e ORTEGA F. (2011) e LEWIS E.G. (2011) per analisi più approfondite e non puramente aneddotiche sul tema.

sorbire i flussi migratori, senza portare necessariamente ad una flessione nei salari dei nativi, che possono al contrario crescere proprio grazie all'immigrazione.⁵

3. - L'impatto dell'immigrazione sul mercato del lavoro: metodi di stima

La sezione precedente ha messo in luce quali sono i meccanismi teorici attraverso i quali l'immigrazione può avere effetti sul mercato del lavoro dei paesi di destinazione. Ma come si può studiare la questione dal punto di vista empirico? Supponiamo di voler analizzare l'effetto dell'immigrazione sui salari dei lavoratori autoctoni. È chiaro che una semplice correlazione tra l'evoluzione dei salari e l'evoluzione della popolazione immigrata nel tempo non fornisce alcuna informazione riguardo alla relazione causale tra queste due grandezze, poiché l'evoluzione di entrambe le variabili può essere influenzata da altri fattori, come il ciclo economico.

In questa sezione descriveremo le strategie sviluppate nella letteratura per risolvere i problemi di stima, che sono state poi applicate anche allo studio degli effetti dell'immigrazione su altri fenomeni. Forniremo poi una breve rassegna della letteratura più significativa nelle sezioni successive.

3.1 *L'approccio spaziale*

L'approccio più diffuso in letteratura è motivato dal seguente esperimento mentale. Si supponga di avere due mercati del lavoro, A e B, perfettamente uguali tra loro. Supponiamo ora che l'immigrazione abbia luogo, e che tutti gli immigrati vengano inviati nel mercato del lavoro A. L'effetto dell'immigrazione sui salari e sull'occupazione potrebbe in questo caso essere misurato confrontando i salari e il tasso di occupazione in A con quelli in B. In questo esempio, B serve come mercato del lavoro controfattuale: rappresenta come sarebbe stato A in assenza di immigrazione.

Questo tipo di intuizione può essere estesa a più di due mercati del lavoro, ed implementata empiricamente, con modelli di regressione dove la variabile dipendente sono i salari medi dei lavoratori residenti in una determinata area e la variabile indipendente di interesse è la concentrazione di immigrati in quell'area. I diversi mercati del lavoro vengono quindi identificati come diverse unità geografiche. Nel caso italiano, una possibilità sarebbe ad esempio di definirli su base provinciale o regionale.

⁵ Si vedano, tra gli altri, PERI G. e SPARBER C. (2009, 2011a); AMUEDO-DORANTES C. e DE LA RICA S. (2011) e D'AMURI F. e PERI G. (2010) per analisi di questo meccanismo.

Se l'obiettivo è quello di stimare l'effetto causale dell'immigrazione, l'implementazione di questo approccio richiede però una serie di accorgimenti.

3.1.1 Effetti permanenti

L'approccio spaziale fornisce una stima dell'effetto causale dell'immigrazione sul mercato del lavoro solo se la distribuzione degli immigrati tra le diverse aree geografiche è casuale e indipendente dalle condizioni del mercato del lavoro nella rispettiva regione. Tuttavia, le condizioni dei mercati del lavoro locali non sono tipicamente identiche (ad esempio la Lombardia ha salari e tassi di occupazione più elevati rispetto alla Calabria), e gli immigrati scelgono autonomamente, una volta arrivati nel paese di destinazione, l'area geografica in cui stabilirsi: in genere, è probabile che gli immigrati sceglieranno le aree che offrono le migliori prospettive economiche. Questo può portare ad una correlazione positiva (o negativa, se per qualche motivo storico gli immigrati tendessero a stabilirsi in prevalenza in aree economicamente più svantaggiate) tra il livello dei salari locali e la concentrazione di immigrati, anche in assenza di effetti reali dell'immigrazione sui lavoratori autoctoni.

Dal punto di vista econometrico, il problema è dunque di rimuovere i c.d. "effetti fissi", cioè le caratteristiche regionali permanenti che differenziano i mercati del lavoro locali a prescindere dall'immigrazione. Se sono disponibili dati longitudinali, ossia dati sui salari regionali in diversi anni, questo può essere fatto in due modi, concettualmente equivalenti: o mettendo in relazione i *cambiamenti* nella concentrazione regionale di immigrati nel tempo con i *cambiamenti* nel livello dei salari; oppure includendo nella regressione variabili *dummy* per ciascuna regione. Entrambe le strategie permettono di eliminare tutte le differenze tra le diverse regioni che sono costanti nel tempo. Il ricercatore stimerà quindi equazioni tipo:

$$(1.a) \quad \Delta y_{it} = \alpha + \beta \Delta m_{it} + \tau_t + \Delta u_{it}$$

o

$$(1.b) \quad y_{it} = \alpha + \beta m_{it} + \varphi_i + \tau_t + u_{it}$$

dove y_{it} sono i salari dei lavoratori nativi nella regione i all'anno t , m_{it} è il rapporto tra immigrati e nativi, τ_t sono variabili *dummy* annuali, e φ_i sono variabili *dummy* per ciascuna regione i . Le formulazioni (1.a) e (1.b) sono equivalenti, e permet-

tono entrambe di stimare il parametro di interesse β , che misura la correlazione tra l'immigrazione e i salari regionali.

3.1.2 Simultaneità

Una volta eliminati gli effetti fissi, è realmente possibile interpretare il parametro β delle equazioni (1.a) e (1.b) come l'effetto causale dell'immigrazione sui salari? Tornando al nostro esperimento mentale, supponiamo che ci siano due periodi e che le condizioni economiche siano identiche in entrambe le regioni all'inizio del periodo 1, ma che alla fine del periodo 1 uno *shock* positivo colpisca la regione B. All'inizio del periodo 2 gli immigrati entrano nell'economia e scelgono la regione in cui andare a risiedere, dopo avere osservato lo *shock*. Ovviamente, è probabile che scelgano di stabilirsi in B piuttosto che in A.

La direzione di causalità tra l'afflusso di immigrati e i salari non è quindi necessariamente chiara, anche una volta tenuto conto delle differenze permanenti tra regioni o, equivalentemente, anche mettendo in relazione la *crescita* dei salari con la *crescita* dell'immigrazione. Gli immigrati possono infatti essere attratti dalle regioni che stanno godendo di un periodo di particolare crescita economica al momento della loro immigrazione. In questo caso, non solo il flusso migratorio potrebbe portare a cambiamenti nei salari regionali, ma le diverse dinamiche salariali regionali potrebbero direttamente riorientare la destinazione dei flussi migratori. La selettività nelle scelte residenziali dei migranti porterebbe quindi a una stima distorta verso l'alto degli effetti dell'immigrazione sul mercato del lavoro. In particolare, qualsiasi impatto depressivo dell'immigrazione sui salari potrebbe essere mascherato dal fatto che i flussi migratori sono particolarmente sostenuti nelle regioni in cui il loro effetto è compensato da *shock* economici positivi.

Un modo per risolvere questo problema empiricamente si basa sul fatto che una delle determinanti principali delle scelte residenziali dei migranti è la presenza in una certa area di concittadini emigrati in passato: i nuovi immigrati tendono a stabilirsi in aree con elevata concentrazione di immigrati della stessa origine o gruppo etnico, come ampiamente dimostrato in letteratura (si vedano per esempio Bartel, 1989 e Munshi, 2003). Poiché è improbabile che le scelte residenziali degli immigrati nel passato, se sufficientemente lontano nel tempo, siano correlate con gli *shock* economici contemporanei, esse possono aiutare a risolvere il problema di simultaneità e a identificare gli effetti dei flussi migratori sui salari. Molti degli studi empirici sull'effetto dell'immigrazione sul mercato del lavoro hanno seguito questo approccio (per esempio Altonji e Card, 1991; Card, 2001; Card e Lewis, 2007; Dustmann, Fabbri e Preston 2005; Dustmann, Frattini e Preston, 2012).

L'idea in questo caso è di decomporre il flusso migratorio regionale in una parte dovuta all'esistenza di *network* migratori non correlati con le condizioni economiche contemporanee, e in un'altra parte dovuta a tutti gli altri fattori, e di utilizzare per la stima solo la prima parte. Questa tecnica è chiamata *regressione a variabili strumentali*, e gli insediamenti storici dei migranti sono in questo caso lo "strumento".

È importante sottolineare come l'ipotesi che gli insediamenti di immigrati passati siano correlati con i movimenti dei salari solo attraverso la loro correlazione con i flussi migratori presenti è un'assunzione identificativa non verificabile. Tale assunzione potrebbe essere non valida, per esempio, nel caso di *shock* economici locali persistenti e nel caso gli insediamenti passati non siano misurati in un periodo sufficientemente lontano nel tempo (si veda per esempio Pischke e Velling, 1997). La correlazione tra gli insediamenti passati e quelli contemporanei può invece essere direttamente osservata nei dati e può quindi essere testata.

3.1.3 Errore di misurazione

Un ulteriore problema nelle stime ottenute con l'approccio spaziale, messo in luce in particolare da Aydemir e Borjas (2011), è dovuto alla scarsa qualità dei dati tipicamente disponibili ai ricercatori. La maggior parte delle ricerche sugli effetti dell'immigrazione sul mercato del lavoro sono basate su indagini campionarie, o su un sottoinsieme dei dati censuari. In questi casi, e in particolare in paesi nei quali la densità di immigrati nella popolazione non è particolarmente elevata, le misure di concentrazione locale dell'immigrazione possono risultare imprecise a causa della ridotta dimensione campionaria. Inoltre, le conseguenze di eventuali errori di misurazione nella concentrazione regionale degli immigrati si aggravano quando le stime sono basate sui *cambiamenti* nel corso del tempo, come suggerito sopra per eliminare il problema delle differenze regionali permanenti. Si può dimostrare che la presenza di errori di misurazione non sistematici nei regressori porta ad un'attenuazione dell'effetto stimato (ovvero, è più difficile trovare una relazione significativa nei dati anche se, nella realtà, la relazione è presente). In questo caso, e riferendoci ancora all'equazione (1.a) o (1.b), la presenza di un errore di misurazione in m_{it} porterebbe a stime di β distorte verso lo zero, anche nel caso in cui un effetto sia realmente presente.

Una possibile soluzione per correggere la distorsione nelle stime introdotta dall'errore di misurazione è la stessa utilizzata per risolvere il problema di simultaneità: la stima a variabili strumentali. Gli strumenti suggeriti per risolvere il problema di simultaneità – ossia le scelte residenziali passate degli immigrati –

possono essere utilizzati anche come strumenti per correggere la distorsione indotta dall'errore di misurazione. Altri esempi di possibili strumenti in questo caso sono misure di concentrazione degli immigrati provenienti da altre fonti, anche misurati con errore, purché l'errore di misurazione sia indipendente da quello presente nella variabile di interesse.

3.1.4 Emigrazione dei nativi

Un'assunzione chiave dell'approccio spaziale è che i mercati del lavoro locali siano economie chiuse, cosicché l'arrivo di un flusso migratorio in uno di essi fa automaticamente aumentare l'offerta di lavoro locale. Se però gli individui sono liberi di muoversi tra le diverse regioni, allora gli effetti dell'immigrazione in una regione si possono diffondere anche ad altre aree geografiche, attenuando quindi l'impatto locale. In particolare, se l'immigrazione ha conseguenze negative sui salari dei lavoratori con un particolare tipo di competenze, è probabile che almeno una parte di questi lavoratori tenda a trasferirsi in un'altra regione, dove non deve subire la concorrenza dei lavoratori immigrati. L'emigrazione dei residenti dalle aree a maggiore intensità migratoria porterà a una dispersione dell'effetto dell'immigrazione su tutto il territorio nazionale, e a una conseguente sottostima degli effetti negativi dell'immigrazione sul mercato del lavoro locale. La letteratura sul tema, sviluppatasi principalmente in ambito statunitense, contiene opinioni differenti riguardo la serietà di questo problema: i lavori di Filer (1992); Frey (1995, 1996) e Borjas (1999, 2006) per esempio, indicano che il fenomeno della migrazione dei nativi in risposta all'immigrazione dall'estero è rilevante, mentre gli studi di Card e DiNardo (2000) o Card (2001, 2009*a*) non trovano nessuna evidenza a supporto di questa tesi. Peri e Sparber (2011*b*) mostrano come i risultati di questo tipo di lavori dipendano crucialmente dal tipo di modelli di regressione utilizzati, indicando in particolare che alcune specificazioni portano per costruzione a risultati distorti. La loro analisi, basata su specificazioni che non soffrono di questo problema, non mostra nessuna evidenza di uno spiazzamento dei nativi in seguito all'immigrazione. Hatton e Tani (2005), uno dei pochi lavori sul tema in paesi europei, trova qualche evidenza di emigrazione dei nativi in risposta all'immigrazione internazionale per il Regno Unito. Per l'Italia, Bruecker, Fachin e Venturini (2010) mostrano che, nel complesso, l'immigrazione internazionale tende a sostituire quella interna. L'analisi di Mocetti e Porello (2010) indica però che vi è una notevole eterogeneità negli effetti. Da un lato, c'è un effetto di spiazzamento per i lavoratori nativi meno qualificati: in particolare l'immigrazione internazionale avrebbe parzialmente sostituito i flussi migratori interni di lavoratori

poco qualificati dalle regioni meridionali a quelle settentrionali. D'altra parte, la loro analisi dimostra anche che l'immigrazione internazionale è associata positivamente alla migrazione interna di lavoratori qualificati, con un impatto più forte nelle aree maggiormente urbanizzate e per i giovani.

È da sottolineare, infine, che il problema sarà tanto più grave quanto più piccole sono le unità spaziali utilizzate nell'analisi: se anche fosse vero che i nativi cambiano residenza a seguito dell'arrivo di immigrati, è probabile che tali spostamenti avvengano entro un'area relativamente ridotta. È probabile quindi che un'analisi che utilizzi regioni sufficientemente ampie permetta di internalizzare i movimenti dei nativi.

3.2 L'approccio di correlazione per celle di istruzione-esperienza

La maggioranza degli studi empirici sull'effetto dell'immigrazione sul mercato del lavoro ha utilizzato l'approccio di correlazione spaziale che abbiamo descritto nella sezione 3.1. Tuttavia, in considerazione soprattutto dei problemi derivanti dalla possibilità che le migrazioni interne dei nativi possano diffondere su tutto il territorio nazionale gli effetti locali dell'immigrazione, Borjas (2003) ha proposto un metodo alternativo di stima, che non richiede di considerare i mercati del lavoro locali come economie separate.

Il punto di partenza di Borjas è la considerazione, ripresa da Card e Lemieux (2001) che, anche all'interno dello stesso gruppo di istruzione, i lavoratori non sono perfetti sostituti, ma che un'altra importante dimensione del capitale umano è costituita dall'esperienza sul mercato del lavoro. All'interno di ciascun livello di istruzione, i lavoratori sarebbero perciò perfetti sostituti con lavoratori con la stessa esperienza lavorativa, mentre solo parzialmente sostituibili con lavoratori che, pur avendo lo stesso livello di istruzione formale, abbiano un diverso livello di esperienza. Di conseguenza, i lavoratori nativi competono sul mercato del lavoro soltanto con gli immigrati con lo stesso livello di istruzione e di esperienza.

Borjas suddivide quindi il mercato del lavoro nazionale in "celle" definite in base al livello di istruzione e agli anni di esperienza lavorativa, e identifica l'effetto dell'immigrazione sui salari dei nativi regredendo i salari medi in ciascuna cella sul livello di immigrazione nella stessa cella. Il problema più ovvio da affrontare utilizzando questo tipo di strategia empirica è che, per esempio, i lavoratori con istruzione universitaria e quindici anni di esperienza lavorativa hanno salari più elevati rispetto ai lavoratori che si sono appena affacciati sul mercato del lavoro dopo avere solo completato la scuola dell'obbligo, indipendentemente dal livello

di immigrazione in ciascuna cella. Come nel caso dell'approccio spaziale, le differenze permanenti nei livelli salariali delle diverse celle possono essere rimosse utilizzando dati longitudinali e inserendo nella specificazione empirica delle variabili *dummy* per ciascuna cella. La tipica equazione stimata in un approccio di correlazione per celle è dunque del tipo:

$$(2) \quad y_{ijt} = \beta m_{ijt} + \varphi_i + \psi_j + \varphi_i \times \psi_j + \tau_t + u_{ijt}$$

dove l'indice i denota il livello di istruzione, l'indice j l'esperienza e l'indice t l'anno; y_{ijt} sono i salari medi nella cella ij nell'anno t , φ_i , ψ_j e τ_t sono variabili *dummy* per, rispettivamente ciascun livello di istruzione, esperienza e anno; l'interazione $\varphi_i \times \psi_j$ cattura le differenze permanenti tra le diverse celle. Il parametro di interesse β in questo caso misura quindi la correlazione tra l'immigrazione in ciascuna cella e i salari dei nativi nella stessa cella.

Il principale vantaggio dell'approccio di correlazione per celle rispetto all'approccio spaziale risiede nel fatto che, mentre gli individui possono scegliere la regione di residenza, essi non possono modificare (quantomeno nel breve periodo) il proprio livello di istruzione ed esperienza. Mentre nell'approccio spaziale la possibilità che i lavoratori nativi cambino la propria regione di residenza come conseguenza dell'immigrazione può inficiare la validità della strategia empirica, l'approccio di correlazione per celle è dunque immune da questa debolezza.

3.2.1 Simultaneità

Anche l'approccio di correlazione per celle, tuttavia, può risentire di problemi di simultaneità. Se da una parte è vero che anche gli immigrati sono esogenamente assegnati ad una cella di istruzione-esperienza, è d'altra parte probabile che l'immigrazione sia più elevata nelle celle dove i salari stanno crescendo più velocemente; *ceteris paribus*, saranno infatti i lavoratori con i profili più richiesti nel paese di destinazione che troveranno più vantaggiosa la migrazione e che perciò tenderanno ad emigrare di più. Non è quindi chiaro in che misura le stime di β in (2) possano essere interpretate come causali, poiché potrebbero essere distorte verso l'alto dall'autoselezione dei migranti. Mentre nell'approccio spaziale il problema di simultaneità è risolto utilizzando come variabile strumentale la distribuzione regionale degli immigrati nel passato, nell'approccio di correlazione per celle non esistono invece (o non sono ancora state individuate in letteratura) variabili strumentali adeguate.

3.2.2 Sostituibilità tra immigrati e nativi

Un'ipotesi fondamentale alla base dell'approccio di correlazione per celle è che, all'interno di uno stesso gruppo di istruzione-esperienza, immigrati e nativi siano perfetti sostituti. Se quest'ipotesi non è valida, infatti, il parametro stimato da regressioni del tipo (2) non è più interpretabile e l'effetto dell'afflusso di immigrati con un certo tipo di istruzione ed esperienza sui salari dei nativi a loro simili risulterà attenuato.

Questo punto è stato sottolineato in due recenti studi di Manacorda, Manning e Wadsworth (2012) e Ottaviano e Peri (2012), che hanno messo in dubbio la validità di questa assunzione per il Regno Unito e gli Stati Uniti, rispettivamente. Poiché i nativi e gli immigrati con livelli di istruzione ed età simili hanno competenze diverse, spesso lavorano in settori e occupazioni diverse, svolgendo diversi tipi di mansioni, questi autori sottolineano che l'ipotesi di perfetta sostituibilità è molto forte e andrebbe testata empiricamente, piuttosto che assunta per valida. Entrambi gli studi mostrano infatti che gli immigrati non sono sostituti perfetti per i nativi, anche quando posseggono lo stesso livello di istruzione formale e di esperienza lavorativa, ma al contrario in parte integrano le loro competenze, risultando complementari nella produzione, cosicché il loro effetto sui salari dei nativi è sostanzialmente inferiore rispetto a quanto sarebbe nel caso di perfetta sostituibilità.

Il risultato di imperfetta sostituibilità tra immigrati e nativi non è unanimemente accettato in letteratura: Jaeger (2007); Aydemir e Borjas (2007); Borjas, Grogger e Hanson (2008, 2010, 2011) stimano ad esempio un'elasticità di sostituzione infinita tra immigrati e nativi con le stesse caratteristiche osservabili negli Stati Uniti; d'altra parte le analisi di Card (2009*a* e 2009*b*) e Raphael e Smolensky (2009) raggiungono risultati diversi per gli Stati Uniti, mentre ad esempio D'Amuri, Ottaviano e Peri (2010) confermano anche per la Germania l'imperfetta sostituibilità tra immigrati e nativi.

In assenza di un consenso su quale sia la reale elasticità di sostituzione tra immigrati e nativi con lo stesso tipo di istruzione ed esperienza (che può fra l'altro variare tra paesi, periodi storici, e in funzione della composizione della popolazione immigrata), vista la crucialità del valore di tale elasticità per la validità delle stime ottenute con l'approccio di correlazione per celle, e considerando che tale elasticità può generalmente essere direttamente stimata, sembra quindi opportuno che il ricercatore che voglia adottare questo approccio di stima si preoccupi preventivamente di verificare se le assunzioni alla sua base siano valide per l'episodio migratorio che sta considerando.

3.2.3 Effetto parziale ed effetto totale

Anche qualora si verificasse che l'ipotesi di perfetta sostituibilità è valida, è importante chiarire che cosa esattamente può essere identificato mediante l'approccio di correlazione per celle. In particolare, è importante sottolineare la distinzione tra effetto *parziale* ed effetto *totale* dell'immigrazione sui salari. Con "effetto parziale" ci riferiamo all'impatto diretto dell'immigrazione in una cella di istruzione-esperienza sui salari dei nativi nella stessa cella, tenendo costante l'offerta di lavoro nelle altre celle. L'"effetto totale" per i lavoratori nativi in ciascun gruppo di istruzione-esperienza è invece dato dalla somma dell'effetto parziale dell'immigrazione di lavoratori con le stesse caratteristiche e degli effetti indiretti derivanti dall'immigrazione di lavoratori con diverse caratteristiche. Come evidenziato dallo stesso Borjas (2003) e sottolineato ad esempio da Ottaviano e Peri (2012), la stima di un'equazione del tipo (2) permette di identificare solo l'effetto parziale dell'immigrazione, presumibilmente negativo perché derivante da un aumento della competizione diretta sul mercato del lavoro, mentre non permette di cogliere l'effetto totale, che può invece anche essere positivo grazie alle complementarità tra il capitale umano di diverso tipo.

Sta nella differente interpretazione che è possibile dare al parametro β , a seconda che esso sia derivato dalla stima dell'equazione (1) o (2) la differenza concettuale tra l'approccio spaziale e quello di correlazione per celle: mentre il primo permette di identificare l'effetto totale dell'immigrazione sui salari dei nativi, con il parametro β che rappresenta la media degli effetti totali nelle diverse regioni, il secondo si limita a identificare l'effetto parziale dell'immigrazione dei lavoratori di uno specifico tipo sui salari dello stesso tipo di lavoratori nativi, con il parametro β che rappresenta in questo caso la media degli effetti parziali fra i diversi tipi di lavoratori.

3.3 L'approccio strutturale

Alcuni dei primi studi sull'effetto dell'immigrazione sul mercato del lavoro (per esempio Grossman, 1982; Borjas, 1987; Borjas, Freeman e Katz, 1996, 1997) descrivevano il mercato del lavoro utilizzando delle specifiche forme funzionali, utilizzate per stimare l'elasticità di sostituzione tra immigrati e nativi. Sulla base dei parametri stimati ricavano poi l'effetto dell'immigrazione sui salari dei nativi. Questo approccio, basato su un modello strutturale in cui i parametri vengono assunti come dati rispetto all'immigrazione e che ricostruisce la situazione controfattuale – le condizioni sul mercato del lavoro in assenza di immigrazione – attraverso la simulazione è stato poi progressivamente abbandonato

a favore di approcci di stima più diretti e meno basati su assunzioni parametriche, come quelli presentati sopra.

Tuttavia, negli ultimi anni, prendendo lo spunto proprio dall'approccio di correlazione per celle proposto da Borjas (2003), si è sviluppata una nuova generazione di studi che hanno ripreso e ampliato un approccio più strutturale all'analisi degli effetti dell'immigrazione sul mercato del lavoro. In particolare i lavori paralleli di Manacorda, Manning e Wadsworth (2012) per il Regno Unito e di Ottaviano e Peri (2012) per gli Stati Uniti hanno rilanciato, estendendolo, l'interesse per questo tipo di approccio. Un approccio di simulazione permette infatti di evitare la suddivisione del mercato del lavoro su base locale, e allo stesso tempo permette di affrontare esplicitamente i due principali punti deboli dell'approccio di correlazione per celle: la sostituibilità tra immigrati e nativi e l'identificazione dell'effetto parziale dell'immigrazione. Entrambi i lavori partono da una funzione di produzione CES (Elasticità di Sostituzione Costante) a due livelli – ormai *standard* nella maggior parte della letteratura recente che studia la reazione dei salari agli *shock* all'offerta e alla domanda di lavoro (si vedano per esempio Katz e Murphy, 1992 e Card e Lemieux, 2001); assumendo per semplicità che l'*output* totale dell'economia (Y) sia prodotto utilizzando solo lavoro qualificato (L_S) e lavoro non qualificato (L_U), la funzione di produzione può essere scritta come:

$$(3) \quad Y = A(\theta L_U^\rho + L_S^\rho)^{\frac{1}{\rho}}$$

dove A è un parametro esogeno che misura la produttività totale dei fattori, θ rappresenta l'efficienza relativa dei lavoratori non qualificati rispetto ai lavoratori qualificati, e ρ è un parametro che determina l'elasticità di sostituzione tra lavoro qualificato e non qualificato:

$$\sigma_Q = \frac{1}{1 - \rho}$$

La quantità totale sia di lavoro qualificato che di lavoro non qualificato è modellata come la combinazione di diversi tipi di lavoro, imperfettamente sostituibili tra loro, identificati dalla classe di età (o equivalentemente dall'esperienza lavorativa potenziale). Denotando con l'indice $i = U, S$ i diversi tipi di lavoro e con l'indice $e = 1, 2, \dots, E$ i diversi gruppi di età, l'aggregato di ciascun tipo di lavoro (L_i) può essere scritto come:

$$(4) \quad L_i = \left(\sum_e \alpha_{ie} L_{ie}^\eta \right)^{\frac{1}{\eta}}$$

dove gli α_{ie} misurano l'efficienza relativa dei diversi gruppi di età all'interno di ciascun tipo di lavoro e il parametro η determina l'elasticità di sostituzione tra i diversi gruppi di età:

$$\sigma_E = \frac{1}{1-\eta}$$

Rispetto alla funzione di produzione *standard* a due livelli, descritti in (3) e in (4), gli autori aggiungono però un terzo livello, dove all'interno di ciascun tipo di lavoro e gruppo di età (L_{ie}) gli immigrati (M) e i nativi (N) sono modellati come potenzialmente imperfettamente sostituibili:

$$(5) \quad L_{ie} = \left(N_{ie}^\delta + \zeta_{ie} M_{ie}^\delta \right)^{\frac{1}{\delta}}$$

dove ζ_{ie} rappresenta l'efficienza relativa di nativi e immigrati in ciascuna "cella" di istruzione ed età e l'elasticità di sostituzione tra immigrati e nativi è data da

$$\sigma_I = \frac{1}{1-\delta}$$

Una volta modellata in questa maniera l'economia del paese di destinazione, gli autori possono poi stimare i parametri del modello, e in particolare l'elasticità di sostituzione tra immigrati e nativi σ_I , l'elasticità di sostituzione tra i diversi gruppi di età σ_E e l'elasticità di sostituzione tra lavoro qualificato e non qualificato σ_Q . I parametri stimati vengono poi utilizzati per simulare l'effetto dell'immigrazione sui salari medi e sui salari dei diversi tipi di lavoratori immigrati e nativi.

L'approccio di simulazione basato su una CES a tre livelli è stato adottato anche, ad esempio, da D'Amuri, Ottaviano e Peri (2010) per la Germania, e da Elsner (2011) per studiare l'effetto dell'emigrazione dalla Lituania, e rappresenta una strategia interessante per studiare a livello nazionale l'effetto dell'immigrazione, andando oltre l'identificazione del solo effetto parziale dell'immigrazione sui salari dei lavoratori dello stesso tipo. Nonostante questo, però, anche l'approccio di simulazione presenta alcuni punti critici.

Innanzitutto, i risultati dipendono criticamente dalle assunzioni fatte riguardo la separabilità dei diversi fattori produttivi e l'ordine in cui i successivi livelli e sotto-livelli sono introdotti nella funzione di produzione. Inoltre, l'approccio di simulazione si basa sull'assunzione che gli immigrati e i nativi possano essere assegnati *ex-ante* a delle "celle" sulla base del loro livello di istruzione e alla loro esperienza lavorativa. Se però, come mostrato ad esempio per il Regno Unito da

Dustmann, Frattini e Preston (2012), almeno nei primi anni dopo l'arrivo nel paese di destinazione gli immigrati tendono ad essere occupati in occupazioni meno qualificate rispetto al loro livello di istruzione formale e di esperienza, le stime dell'elasticità di sostituzione tra immigrati e nativi all'interno di ciascuna cella possono essere distorte. Inoltre Dustmann e Preston (2012) mettono in luce come la stima di questa elasticità di sostituzione, dalla quale dipende crucialmente il risultato dell'analisi, sia sensibile alla lunghezza della variazione temporale utilizzata per il calcolo.

4. - Immigrazione e mercato del lavoro: evidenza empirica

Sono ormai moltissimi gli studi empirici sull'effetto dell'immigrazione sui mercati del lavoro dei paesi di destinazione. Non è questa la sede per presentare una rassegna esaustiva dei risultati ottenuti dalla letteratura nel corso degli anni e per i diversi paesi. Ci limiteremo invece a richiamare i risultati degli studi più importanti, suddivisi per metodologia, presentando poi nella sezione 5 gli studi per l'Italia. I lettori interessati possono trovare rassegne della letteratura più dettagliate in Friedberg e Hunt (1995); Gaston e Nelson (2002); Longhi, Nijkamp e Poot (2005, 2008); Dustmann, Frattini e Glitz (2008); Okkerse (2008); Kerr e Kerr (2011).

4.1 *Approccio spaziale*

Il primo e più famoso studio che utilizza l'intuizione dell'approccio spaziale per studiare gli effetti dell'immigrazione sul mercato del lavoro è quello di Card (1990). Card usa l'"esperimento naturale" fornito dall'afflusso imponente e improvviso di immigrati cubani a Miami tra il maggio e il settembre del 1980, a seguito del cosiddetto "esodo di Mariel" (in inglese "*Mariel boatlift*"), per studiare le conseguenze dell'immigrazione sul mercato del lavoro di Miami.⁶ Paragonando i cambiamenti nel mercato del lavoro di Miami prima e dopo l'esodo di Mariel con quelli in altre città americane comparabili, Card mostra che l'ondata migratoria che ha colpito la città della Florida non ha avuto conseguenze negative sui salari o sul tasso di occupazione e disoccupazione dei nativi.⁷ Altonji e Card

⁶ L'esodo di Mariel ha anche ispirato il famoso film di Brian De Palma "*Scarface*".

⁷ Esempi di altri studi che adottano una strategia simile sono HUNT J. (1992) e CARRINGTON W.J. e DE LIMA P.J.F. (1996), che utilizzano come "esperimenti naturali" gli episodi di migrazione di ritorno rispettivamente dall'Algeria verso la Francia e da Angola e Mozambico verso il Portogallo a seguito del processo di de-colonizzazione.

(1991) adottano invece per la prima volta l'approccio spaziale nella forma presentata nella sezione 3.1, stimando un'equazione del tipo (1) per studiare gli effetti dell'immigrazione negli USA tra il 1970 e il 1980, e tenendo conto dell'endogeneità nelle scelte residenziali degli immigrati utilizzando lo *stock* di immigrati in ciascuna area nel 1970 come strumento per il flusso di immigrati tra il 1970 e il 1980. I loro risultati mostrano che l'immigrazione ha conseguenze moderatamente negative sui salari dei lavoratori poco qualificati, ma non trovano effetti significativi sulla partecipazione alla forza lavoro e sul tasso di occupazione.

In un altro articolo molto influente, Card (2001) esamina l'impatto dell'immigrazione su salari e tassi di occupazione dei lavoratori in sei diversi gruppi occupazionali in 175 aree metropolitane negli Stati Uniti, utilizzando i dati del censimento del 1990. L'analisi di Card sovrappone in un certo senso l'approccio spaziale e quello per celle, poiché il mercato del lavoro è suddiviso su base regionale e professionale. L'endogeneità nelle scelte residenziali dei migranti è corretta utilizzando come strumento la cosiddetta "*supply-push component*", ossia il tasso di immigrazione atteso in ciascuna occupazione e regione se i nuovi immigrati fossero distribuiti sul territorio nazionale esattamente come i loro concittadini nel passato. I risultati della sua analisi mostrano ancora una volta che gli effetti sui salari relativi e sull'occupazione dei nativi sono estremamente limitati.

Un lavoro recente di Dustmann, Frattini e Preston (2012) analizza l'effetto dell'immigrazione sui salari medi e sulla distribuzione dei salari, utilizzando dati annuali su salari e immigrazione a livello regionale per il Regno Unito tra il 1997 e il 2005. Le loro stime mostrano che l'immigrazione ha avuto un effetto positivo sui salari medi, ma effetti differenziati lungo la distribuzione: mentre l'immigrazione ha compresso la crescita dei salari sotto il ventesimo percentile, ha invece avuto un effetto positivo sulla crescita salariale nella parte superiore della distribuzione del reddito, al di sopra del quarantesimo percentile. Questi risultati sono compatibili con le predizioni teoriche di un modello del tipo presentato nella sezione 2: nel periodo 1997-2005, infatti, gli immigrati erano più concentrati rispetto ai nativi sotto il primo quartile della distribuzione dei salari, esattamente dove gli autori dimostrano che la crescita salariale è stata rallentata, e meno concentrati nei quartili superiori, dove l'effetto sui salari è stato positivo. In ogni caso, gli effetti dell'immigrazione sui salari dei cittadini britannici sarebbero stati quantitativamente modesti: l'immigrazione avrebbe contribuito per al massimo 1/30 alla crescita dei salari.

4.2 *Approccio di correlazione per celle*

Come discusso in precedenza, è possibile che l'approccio di correlazione spaziale tenda a sottostimare l'effetto dell'immigrazione sui salari dei nativi, soprattutto a causa della possibilità che i movimenti dei nativi compensino i cambiamenti nell'offerta di lavoro dovuti all'immigrazione (vedi sezione 3.1). L'unità di osservazione corretta per studiare gli effetti dell'immigrazione sul mercato del lavoro sarebbe quindi l'intero paese, e il confronto andrebbe fatto tra celle di istruzione-esperienza, piuttosto che fra regioni (vedi sezione 3.2).⁸ Borjas (2003) propone per la prima volta l'approccio di correlazione per celle, e lo implementa utilizzando dati per gli USA dai censimenti decennali per gli anni 1960-1990 e i dati della *Current Population Survey* (CPS) 1998-2001. I risultati del suo studio implicano che un aumento dell'1% della quota di immigrati riduce i salari dei lavoratori nativi concorrenti dello 0,3-0,4%. In un lavoro successivo, Aydemir e Borjas (2007) applicano la stessa strategia empirica allo studio congiunto dell'immigrazione in Canada e USA, e dell'emigrazione dal Messico. Anche in questo caso, i loro risultati indicano che un cambio del 10% nell'offerta di lavoro è associato con un cambiamento nei salari di segno opposto del 3%-4%.

4.3 *Approccio strutturale*

Ottaviano e Peri (2012) prendono le mosse dal lavoro di Borjas, sottolineando però come questo permetta di identificare solo l'effetto parziale dell'immigrazione sui salari dei nativi, mentre per adottare un approccio che consideri tutto il territorio nazionale e stimi l'effetto totale dell'immigrazione sui salari vi sia bisogno di imporre una maggiore struttura ai dati. Il loro studio, basato sui dati del censimento USA per il periodo 1960-2000 e dell'*American Community Survey* (ACS) per il 2006, e quello parallelo di Manacorda *et al.* (2012) per il Regno Unito che utilizza i dati della *General Household Survey* (GHS) e *Labour Force Survey* (LFS) per il periodo 1975-2005, stimano l'elasticità di sostituzione tra immigrati e nativi in ogni cella di istruzione-esperienza, l'elasticità di sostituzione tra lavoratori con diversi livelli di esperienza e con diversi livelli di istruzione, e ricavano poi l'effetto complessivo dell'immigrazione sui salari dei nativi utilizzando queste stime per parametrizzare una funzione CES multilivello.

⁸ Da segnalare anche uno studio di FRIEDBERG R.M. (2001) sull'effetto dell'immigrazione di massa dalla Russia verso Israele tra il 1990 e il 1994, che utilizza variazioni nell'immigrazione a livello nazionale tra diversi settori per stimare le conseguenze sui salari, mostrando che l'immigrazione non ha avuto effetti negativi sui salari medi degli israeliani, e che anzi gli effetti potrebbero essere stati anche positivi.

Entrambi gli studi mostrano che, tenendo conto dell'imperfetta sostituibilità tra nativi e immigrati e considerando gli effetti incrociati dell'immigrazione in diversi gruppi di istruzione-esperienza, l'immigrazione ha avuto al più modesti effetti negativi per i nativi poco qualificati, mentre ha avuto (almeno nel caso degli USA) conseguenze positive sui salari medi dei nativi. Viceversa, i lavoratori che hanno sofferto maggiormente la competizione dei nuovi immigrati, subendo conseguenze negative per i propri salari, sarebbero stati i lavoratori stranieri arrivati in precedenza.

4.4 *Meta-analisi*

Nonostante la grande varietà di paesi analizzati, composizione della popolazione immigrata, periodi storici e tecniche di stima, la maggior parte delle analisi empiriche sugli effetti dell'immigrazione sul mercato del lavoro tendono a trovare conseguenze limitate, e in alcuni casi persino in media positive, dell'immigrazione sui salari dei lavoratori nativi (con la parziale eccezione degli studi che utilizzano l'approccio di correlazione per celle). Longhi, Nijkamp e Poot (2005) hanno condotto una meta-analisi su un campione di diciotto studi, riguardanti diversi paesi e che forniscono complessivamente 344 diverse stime del cambiamento percentuale nei salari dei lavoratori nativi indotto da un aumento di un punto percentuale nel rapporto immigrati/nativi. La loro analisi mette in relazione i risultati ottenuti da questi studi con diverse caratteristiche della ricerca come l'approccio utilizzato, il paese, le dimensioni del mercato del lavoro, il gruppo di lavoratori interessato, il tipo di immigrati, la definizione dei salari e tiene esplicitamente conto della qualità dello studio e del cosiddetto "*bias* di pubblicazione", ossia della tendenza degli autori e degli *editor* delle riviste a pubblicare risultati statisticamente significativi. I loro risultati indicano un effetto complessivamente assai modesto dell'immigrazione sui salari: l'aumento di un punto percentuale della proporzione di immigrati nella forza lavoro porta in media a una diminuzione dei salari dei nativi di solo lo 0,12%. In particolare, l'impatto negativo tende ad essere maggiore nei paesi dell'Unione Europea che negli Stati Uniti, mentre gli immigrati sono più in competizione tra loro che con i nativi. I risultati della meta-analisi sembrano quindi confermare la precedente conclusione: l'impatto dell'immigrazione sui salari, anche quando è statisticamente significativo, tende ad essere quantitativamente piccolo.

È opportuno infine ricordare che le metodologie di analisi delineate nella sezione 3, e in particolar modo l'approccio spaziale, sono state utilizzate per analizzare gli effetti dell'immigrazione su numerose altre variabili, economiche e non.

Una rassegna completa di questi studi è al di fuori delle possibilità di questo articolo, ma alcuni esempi di queste applicazioni vengono forniti nel prossimo paragrafo, che si concentra sull'impatto dell'immigrazione in Italia.

5. - L'impatto dell'immigrazione in Italia

L'Italia è diventata un paese di immigrazione solo recentemente, con un bilancio migratorio (cioè la differenza tra il numero di immigrati e il numero di emigrati) che è diventato positivo solo a partire dal 1973 (Einaudi, 2007), e la popolazione straniera ha iniziato ad essere una percentuale rilevante della popolazione totale solo a partire dagli anni Novanta (vedi Appendice 1). La trasformazione da paese di emigrazione a paese di immigrazione ha trovato l'Italia impreparata sotto diversi punti di vista, tra i quali anche la difficoltà di un'adeguata analisi quantitativa del fenomeno dovuta alla mancanza di dati completi, aggiornati e organici. Si pensi ad esempio al fatto che l'Indagine trimestrale dell'Istat sulle forze di lavoro, la principale fonte di informazioni sul mercato del lavoro italiano, non conteneva informazioni sulla nazionalità o sul paese di nascita fino a tutto il 2005. La carenza di dati è stata probabilmente una delle ragioni per cui la ricerca economica sugli effetti dell'immigrazione in Italia è stata scarsa fino a pochi anni fa, ed ha iniziato ad estendersi solo di recente.⁹

Gli studi di Gavosto, Venturini e Villosio (1999) e Venturini e Villosio (2006) sono stati tra i primi ad analizzare gli effetti dell'immigrazione sul mercato del lavoro italiano. I loro risultati indicano che i lavoratori immigrati tra la fine degli anni Ottanta e la metà degli anni Novanta sono stati complementari ai lavoratori italiani, favorendo in particolare la crescita dei salari dei lavoratori nativi non qualificati, senza avere effetti significativi sulle probabilità di transizione tra occupazione e disoccupazione e tra disoccupazione e occupazione.¹⁰

Queste ricerche si riferiscono però a un periodo nel quale l'immigrazione in Italia era ancora piuttosto limitata. Che cosa ci dice l'evidenza più recente? Gli studi sono numericamente limitati, ma sembrano indicare – in linea con la maggior parte della letteratura internazionale – che l'immigrazione non abbia avuto

⁹ Gli studi di taglio economico hanno affiancato una più sviluppata letteratura in ambito prevalentemente demografico e sociologico. Si vedano ad esempio: AMBROSINI M. (1999); COLOMBO A. e SCIORTINO G. (2002); ZINCONE G. (2000, 2001).

¹⁰ Si veda anche VENTURINI A. (1999), che studia l'effetto dell'impiego irregolare di lavoratori immigrati sull'occupazione dei nativi.

conseguenze significative per il mercato del lavoro italiano. Staffolani e Valentini (2010) usano i dati del *Work Histories Italian Panel* (WHIP), basato sui dati amministrativi dell'Inps, per studiare l'effetto dell'immigrazione *extra-UE* sui salari dei lavoratori italiani tra il 1995 e il 2004. Utilizzando un approccio simile a quello della correlazione per celle, ma definendo le celle in termini di età e genere, poiché i dati non riportano informazioni sul livello di istruzione degli individui, gli autori mostrano che l'immigrazione è associata positivamente sia con la crescita dei salari dei lavoratori qualificati che di quelli non qualificati. I risultati sono robusti a cambiamenti della definizione delle celle, anche se le stime potrebbero potenzialmente essere distorte verso l'alto per l'auto-selezione dei migranti (si veda al proposito la discussione nella sezione 3.2). Romiti (2011) utilizza gli stessi dati, ma adotta un approccio strutturale per stimare l'effetto dell'immigrazione sul mercato del lavoro italiano durante lo stesso periodo. Le sue stime dimostrano innanzitutto che immigrati e nativi non sono perfetti sostituti, anche all'interno della stessa area geografica e tipo di occupazione (*blue collar* o *white collar*). Le simulazioni ottenute con i parametri stimati mostrano poi che l'unico gruppo di lavoratori i cui salari sono stati negativamente colpiti dall'immigrazione sono gli immigrati già in Italia, mentre l'immigrazione non ha avuto effetti negativi per i nativi, pur facendo aumentare i salari relativi dei lavoratori nativi più qualificati. Nonostante l'utilizzo di approcci di stima diversi, i due studi arrivano dunque a conclusioni simili.¹¹

Sebbene ulteriori studi sugli effetti dell'immigrazione sul mercato del lavoro italiano sarebbero utili per fornire un quadro più stabile di risultati, l'evidenza disponibile finora indica che, tutt'al più, l'immigrazione internazionale ha avuto effetti positivi sui salari dei lavoratori italiani.¹² L'immigrazione potrebbe però avere avuto altri effetti, più indiretti ma altrettanto positivi, sul mercato del lavoro italiano. Barone e Mocetti (2011) dimostrano ad esempio che l'aumento dell'of-

¹¹ È da segnalare in questo contesto anche il lavoro di ACCETTURO A., BUGAMELLI M. e LAMORGESE A. (2012), che analizzano la reazione delle imprese all'immigrazione di lavoratori poco qualificati, concentrandosi sugli effetti sulla domanda di capitale invece che sulla domanda di lavoro. Il loro studio, basato su un *panel* di imprese manifatturiere italiane con almeno 50 addetti nel periodo 1996-2007 mostra che in Italia l'afflusso di immigrati poco qualificati avrebbe determinato un incremento dell'intensità di capitale delle imprese, con un effetto maggiore per le imprese più grandi e per quelle che operano in settori a più alta intensità di lavoratori istruiti.

¹² CINGANO F. e ROSOLIA A. (2010) riassumono lo stato dell'arte delle conoscenze sul ruolo dei lavoratori immigrati nel mercato del lavoro italiano con il loro titolo "Non sono concorrenti ma complementari".

ferta di lavoro domestico (*colfs*, badanti) dovuta all'immigrazione ha permesso alle donne italiane, e in particolar modo a quelle con alti livelli di istruzione, di sostituire il lavoro domestico con il lavoro sul mercato.¹³ La loro analisi, che utilizza un approccio spaziale ed è basata su dati dall'Indagine Istat sulle forze di lavoro e su dati amministrativi provenienti dalle anagrafi municipali per gli anni tra il 2006 e il 2008, rivela che mentre l'immigrazione non ha avuto effetti sulla probabilità di partecipazione femminile al mercato del lavoro (il c.d. *margin estensivo*), ha invece fatto aumentare in media il numero di ore lavorate (*margin intensivo*) dalle donne italiane. In particolare, il loro studio mostra che l'effetto medio deriva interamente dall'aumento nel numero di ore lavorate dalle donne laureate, mentre non ci sono cambiamenti significativi nelle ore lavorate dalle donne con livelli di istruzione inferiori. L'analisi suggerisce che questo effetto sia da attribuire alla riduzione del costo dei servizi domestici, dovuta ad una maggiore offerta trainata dall'immigrazione, e non all'aumento della domanda di lavoro femminile qualificato indotta dalla complementarità tra immigrati e nativi nella produzione.

Anche i risultati di Romiti e Rossi (2011) indicano che l'immigrazione ha avuto un effetto positivo sulla partecipazione femminile al mercato del lavoro, agendo però su un diverso margine, ossia sulle scelte di pensionamento. Utilizzando i dati dell'Indagine sui bilanci delle famiglie italiane di Banca d'Italia e dati anagrafici sul numero di residenti a livello regionale, adottano un approccio spaziale e mostrano che l'immigrazione ha contribuito a ritardare l'età di pensionamento delle donne, lasciando invariata quella degli uomini. L'effetto sarebbe da attribuire alla possibilità offerta dall'immigrazione di avere servizi di cura per gli anziani a basso costo, che riduce il costo-opportunità del lavoro, e rende più conveniente continuare a lavorare invece di andare anticipatamente in pensione per accudire i genitori anziani.

Gli effetti messi in luce da questi due studi sono compatibili con i risultati preliminari di Olivieri (2010) che mostrano, confermando anche per l'Italia risultati simili di Cortes (2008) per gli USA e Frattini (2008) per il Regno Unito, come l'immigrazione tra il 2005 e il 2009 abbia contribuito a ridurre la crescita dei prezzi dei servizi.¹⁴

¹³ CORTES P. e TESSADA J. (2011) e FARRÉ L., GONZALES L. e ORTEGA F. (2011) hanno condotto studi simili per gli USA e la Spagna.

¹⁴ Per quanto riguarda l'effetto dell'immigrazione sui prezzi, ACCETTURO A. *et AL.* (2012) mostrano che l'immigrazione ha un effetto positivo sui prezzi medi delle case a livello cittadino, ma tende a far diminuire i prezzi delle case nei quartieri dove si concentrano gli immigrati.

Se gli effetti sul mercato del lavoro sono stati, almeno dal punto di vista degli italiani, positivi, quali sono i costi fiscali dell'immigrazione? L'immigrazione di individui tendenzialmente poco qualificati dall'estero ha comportato una spesa per le finanze pubbliche italiane?¹⁵ Una risposta esaustiva a questo tipo di domande è metodologicamente complessa (si veda Rowthorn (2008) per una rassegna della letteratura e per una discussione), ma l'evidenza disponibile per l'Italia sembra indicare di no.

L'analisi di Devillanova (2011), basata sui dati dell'indagine Istat su *Redditi e condizioni di vita (IT-Silc)* per il 2007, dimostra che nell'anno considerato gli immigrati hanno pagato in media meno imposte degli italiani, ma hanno anche ricevuto minori trasferimenti. In particolare, gli italiani ricevono mediamente maggiori trasferimenti per benefici legati all'anzianità e alla disabilità e sono destinatari di maggiori spese sanitarie rispetto agli immigrati, che tendono invece a ricevere maggiori trasferimenti per indennità per carichi di famiglia (assegni familiari), aiuti pubblici per la casa e spesa per istruzione. La gran parte delle differenze nei trasferimenti ricevuti tra la popolazione italiana e quella immigrata è dovuta a differenze demografiche. In media, la popolazione italiana è più anziana di quella immigrata, da cui le maggiori spese per pensioni e sanità e le minori spese per istruzione. È da sottolineare però come, anche al netto della spesa pensionistica, non si rilevi un maggior costo fiscale per gli immigrati rispetto agli italiani. Lo studio si limita a "fotografare" la situazione per l'anno 2007, tuttavia altri studi condotti utilizzando dati provenienti dalla stessa indagine per anni precedenti portano a risultati simili (DeviIlanova 2009 e 2010).

Pellizzari (2012) studia le differenze nella probabilità di accesso al *welfare* tra immigrati e nativi, utilizzando dati amministrativi dall'archivio Inps per il 2005 e unendoli ad informazioni sull'ISEE (Indicatore della Situazione Economica Equivalente), dove la presenza nell'archivio ISEE è un indicatore del fatto che l'individuo abbia presentato domanda per qualche tipo di prestazione di *welfare* municipale. Il vantaggio principale di questi dati è che permettono di ottenere informazioni sui programmi di *welfare* locali, che non sono invece catturati dalle indagini campionarie tradizionalmente utilizzate in questo tipo di analisi. I suoi risultati mostrano che gli immigrati extraeuropei hanno una maggiore probabilità di avere fatto domanda per qualche forma di *welfare* rispetto agli italiani, principalmente a causa delle loro caratteristiche demografiche. Tuttavia, controllare per

¹⁵ Preoccupazioni di questo tipo sono spesso alla radice dell'opposizione a politiche migratorie meno restrittive (BOERI T., 2010).

la distribuzione geografica degli immigrati fa aumentare il differenziale nella probabilità di domanda di accesso a prestazioni di *welfare*. Gli immigrati tendono infatti ad essere concentrati nelle aree più ricche del Paese, dove maggiore è la domanda di lavoro poco qualificato. Di conseguenza, tendono ad essere occupati in lavori meno pagati rispetto a quelli dei nativi e dunque ad usufruire maggiormente del *welfare* municipale. I risultati di questo studio possono quindi anche fornire una spiegazione all'apparente contraddizione tra la percezione diffusa che gli immigrati sovra-utilizzino i servizi pubblici e i benefici sociali rispetto ai nativi, e i risultati di studi empirici che affermano il contrario. Se infatti in aggregato gli immigrati non rappresentano un costo per le finanze pubbliche maggiore di quello dei cittadini italiani, nelle aree dove sono maggiormente concentrati, che coincidono con le aree dove risiedono i nativi più ricchi, essi tendono invece a ricevere trasferimenti netti maggiori di quelli dei nativi.

L'immigrazione in Italia è spesso associata alla criminalità, e viene quasi dato per scontato nel discorso pubblico che flussi migratori più elevati implicino un aumento nel tasso di criminalità. Questa percezione è però confutata da un recente studio di Bianchi, Buonanno e Pinotti (2012). Lo studio analizza quantitativamente, basandosi su dati amministrativi sul numero di reati denunciati e sul numero di permessi di soggiorno per il periodo 1990-2003, la relazione esistente tra immigrazione e criminalità. Utilizzando un approccio di correlazione spaziale, gli autori mostrano che l'immigrazione è effettivamente correlata positivamente col numero di reati contro la proprietà e col tasso di criminalità complessivo. Tuttavia, la loro analisi a variabili strumentali, che tiene conto del fatto che gli immigrati tendono a vivere in province con maggiori tassi di criminalità, dimostra che l'unico effetto causale dell'immigrazione è di portare a un aumento nel numero di rapine. Poiché le rapine rappresentano meno del 2% del numero totale di reati, l'effetto sul tasso di criminalità complessivo è però insignificante. Non vi è dunque, ad oggi, alcuno studio che supporti l'ipotesi che l'immigrazione comporti un aumento nei tassi di criminalità.

6. - Conclusioni

Le migrazioni internazionali sono un fenomeno di crescente importanza nelle società moderne, e un ingrediente fondamentale dell'economia globalizzata. Tuttavia, paradossalmente, mentre nel corso dell'ultimo secolo i paesi del Nord del mondo si sono progressivamente aperti al flusso di beni e capitali, hanno con-

temporaneamente adottato politiche migratorie via via più restrittive (Hatton, 2007; Mayda, 2008). Alla base di queste scelte stanno probabilmente i timori di larga parte dell'opinione pubblica che l'immigrazione possa avere conseguenze negative per il benessere dei nativi nei paesi di destinazione (Facchini e Mayda, 2008).¹⁶ In questo articolo, abbiamo cercato di presentare un'introduzione all'analisi economica degli effetti dell'immigrazione sui paesi di destinazione, concentrandoci in particolare sugli effetti dell'immigrazione sul mercato del lavoro. Abbiamo innanzitutto mostrato come, all'interno di un modello stilizzato del mercato del lavoro, l'immigrazione porti a un aumento complessivo del benessere dei nativi, ma abbia importanti effetti distributivi che dipendono dalla composizione relativa della forza lavoro immigrata e nativa. Abbiamo poi discusso alcune estensioni del modello di base, che mettono in luce come il mercato del lavoro dei paesi di destinazione possa assorbire l'aumento nell'offerta di lavoro comportata dall'immigrazione senza che si debba necessariamente modificare il prezzo dei fattori. Nella seconda parte dell'articolo abbiamo invece presentato criticamente le diverse metodologie di stima empirica sviluppate nel corso degli ultimi decenni, analizzandone punti di forza e di debolezza. Alcune di queste strategie rivestono un ruolo particolarmente importante nella ricerca economica sulle migrazioni, perché sono state adottate con successo anche al di fuori dell'analisi delle conseguenze sul mercato del lavoro. La rassegna degli studi sugli effetti dell'immigrazione sull'Italia, con la quale concludiamo, ha permesso di mostrare alcuni degli altri ambiti nei quali questi strumenti sono stati utilizzati: il mercato del lavoro, le scelte di investimento delle imprese, l'offerta di lavoro femminile, le scelte di pensionamento, le conseguenze fiscali, il tasso di criminalità. Il quadro che emerge complessivamente è di un'immigrazione che ha avuto effetti assai limitati, quasi mai negativi e anzi spesso positivi, in ciascuno degli ambiti considerati. In questo senso, l'esperienza italiana è quindi in linea con l'esperienza della maggior parte degli altri paesi.

La ricerca di questi anni ha dunque mostrato che molti timori legati all'immigrazione sono infondati, e ha messo in luce come le conseguenze economiche dell'immigrazione per i cittadini dei paesi ospiti siano, in genere, quantitativamente modeste. Ciononostante, è indubbio che l'immigrazione ponga importanti questioni ai decisori politici dei paesi ospiti, ai quali spetta il compito di saper cogliere le opportunità che essa pone, e di saperne affrontare le problematiche che sorgono,

¹⁶ Si veda però, ad esempio, CARD D., DUSTMANN C. e PRESTON I. (2012) per un'analisi del ruolo dei fattori non economici.

ad esempio, nel processo di integrazione economica, lavorativa, e scolastica degli immigrati e dei loro figli. Il contributo della ricerca sociale potrà auspicabilmente essere in questi ambiti di valido supporto alle scelte dei decisori politici.

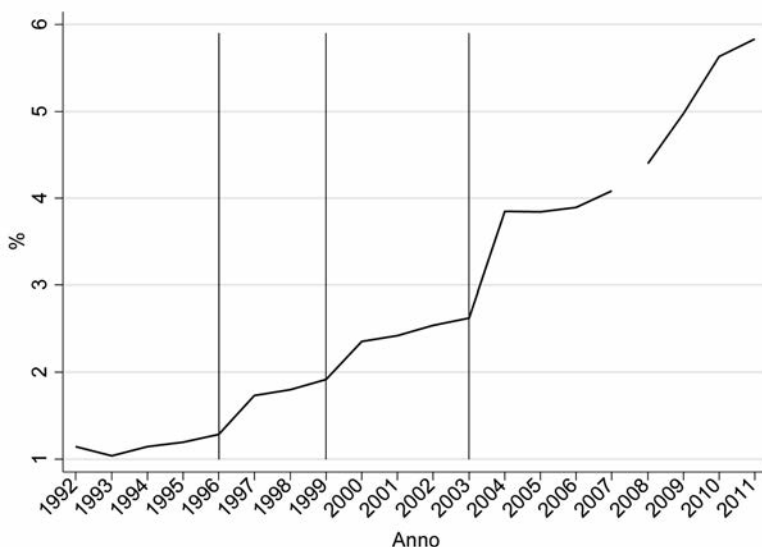
APPENDICE 1

L'immigrazione in Italia

Al 1 gennaio 2011 risultavano residenti in Italia 4.570.317 cittadini stranieri (di nazionalità non italiana), pari al 7,53% della popolazione totale, dei quali il 52% donne e il 22% minori (Fonte: Rilevazione ISTAT sulla “Popolazione residente comunale straniera per sesso ed anno di nascita”). Tale dato, basandosi sulla cittadinanza e non sul paese di nascita, può però sottostimare la numerosità della popolazione di origine straniera, in quanto non tiene conto dei cittadini naturalizzati italiani. Risultano inoltre esclusi dal conto i cittadini stranieri irregolarmente residenti in Italia. Secondo le stime di Blangiardo (2008), aggiornate al 1 gennaio 2008, erano presenti a quella data 651.000 cittadini stranieri non in regola col permesso di soggiorno, pari a circa il 15% della popolazione straniera totale.

GRAF. B1

EVOLUZIONE PERMESSI DI SOGGIORNO IN PERCENTUALE DELLA POPOLAZIONE, 1992-2011



Fonte: Elaborazione su dati ISTAT, Demo-Geodemo.

La politica migratoria italiana si caratterizza per il contrasto tra politiche di ingresso restrittive e frequente ricorso a sanatorie per regolarizzare *ex-post* gli stranieri irregolarmente residenti in Italia (si veda Fasani, 2010 per un'analisi del-

l'immigrazione irregolare in Italia). Il grafico B1 riporta l'evoluzione del numero di permessi di soggiorno rilasciati a cittadini stranieri non comunitari, misurati al 1 gennaio di ciascun anno ed espressi in percentuale della popolazione residente tra il 1992 e il 2011.¹⁷ Le barre verticali sono poste in corrispondenza del varo di provvedimenti di sanatoria: si nota come, in corrispondenza di ciascuna sanatoria, vi sia un salto nel numero di permessi di soggiorno. È anche interessante sottolineare come, fino alla sanatoria del 1995, gli immigrati regolarmente presenti in Italia fossero poco più dell'1% della popolazione totale.

Per quanto riguarda i paesi di provenienza, e concentrandoci solo sui paesi al di fuori dell'Unione Europea, per i quali abbiamo dati amministrativi dettagliati dai permessi di soggiorno¹⁸, la Tavola B1 mostra che un terzo degli immigrati residenti in Italia sono originari dell'Europa dell'Est, mentre oltre uno su cinque è nordafricano e uno su quattro asiatico. I latinoamericani sono il 10% del totale. Il paese più rappresentato è il Marocco (oltre il 14% del totale), seguito dall'Albania (poco sotto il 14%), dalla Cina (8%) e da Ucraina (6%) e Moldavia (4%).

TAV. B1

COMPOSIZIONE IMMIGRATI PER AREE DI ORIGINE
E PAESI PIÙ RAPPRESENTATIVI, 2011

aree di origine		dieci principali paesi di origine	
Europa Centro-orientale	32%	Marocco	14,2%
Nord Africa	22%	Albania	13,7%
Asia Centro-meridionale	12%	Cina	7,8%
Asia Orientale	12%	Ucraina	6,2%
America Latina	10%	Moldavia	4,0%
Africa Occidentale	8%	India	4,0%
Africa Orientale	1%	Filippine	3,9%
America Settentrionale	1%	Tunisia	3,3%
Asia Occidentale	1%	Egitto	3,1%
Africa Centro-Meridionale	1%	Bangladesh	2,9%
Oceania	0%		

Fonte: Elaborazione su dati ISTAT, Demo-Geodemo.

¹⁷ I dati dal 2008 in avanti non sono comparabili con quelli precedenti a causa dell'esclusione dei nuovi cittadini dell'Unione Europea (rumeni, bulgari), per i quali, dal 27 marzo 2007, non è più previsto il rilascio del documento di soggiorno, e a causa dell'inserimento nel computo dei permessi anche dei minori registrati sul permesso di un adulto, precedentemente esclusi.

¹⁸ Si noti che escludiamo quindi la Romania, dal quale provengono una importante frazione degli stranieri in Italia, ma che è un paese membro dell'UE dal 2007.

APPENDICE 2

Dati

Forniamo in questa appendice alcuni *link* a siti *web* e *database* che possono essere utili per reperire dati e informazioni sull'immigrazione in Italia e nel mondo.

- **United Nations Population Division**

Banca dati su *stock* e flussi di immigrati nel tempo, suddivisi per aree e paesi di origine e destinazione, età e genere, basandosi principalmente sui censimenti decennali di ciascun paese.

Trends in International Migrant Stock (dati *stock*): <http://esa.un.org/MigAge/>
International Migration Flows To and From Selected Countries (dati flussi):
<http://esa.un.org/MigFlows/MigrationFlows.html>

- **OECD International Migration Database**

Banca dati su flussi e *stock* della popolazione immigrata totale e della forza lavoro immigrata, e sulle naturalizzazioni, suddivisa per paesi di origine e di destinazione e per genere, raccolta con frequenza annuale dalla rete di corrispondenti nazionali dell'OCSE nell'ambito del Continuous Reporting System on Migration (SOPEMI). Le fonti variano tra paesi.

Il *dataset* è accessibile dal sito OECD.Stat: <http://stats.oecd.org>

La pubblicazione annuale dell'OCSE "International Migration Outlook" utilizza questi dati e comprende ogni anno un focus su un tema specifico. Per l'edizione 2011 si veda:

http://www.oecd.org/document/40/0,3746,en_2649_37415_48303528_1_1_1_37415,00.html

- **OECD Database on Immigrants in OECD Countries**

Banca dati che fornisce informazioni comparate su caratteristiche demografiche (età e genere), durata del soggiorno nel paese ospite, situazione lavorativa (*status* lavorativo, occupazione, settore), e istruzione di immigrati e nativi nei paesi OCSE. Informazioni basate su censimenti e dati amministrativi, integrate in alcuni casi da indagini campionarie sulle forze di lavoro. Le informazioni si riferiscono all'anno 2000.

Il *dataset* è accessibile dal sito OECD.Stat: <http://stats.oecd.org>

Un compendio dei dati è disponibile nella pubblicazione dell'OCSE "A Profile

of Immigrant Populations in the 21st Century: Data from OECD Countries”, disponibile su :

http://www.oecd.org/document/27/0,3746,en_2649_33931_40110299_1_1_1_1,00.html

- **World Bank - Global Bilateral Migration Database**

Matrici bilaterali su *stock* di immigrati in diversi paesi, disaggregate per genere, a intervalli decennali per gli anni 1960-2000. Informazioni basate sui censimenti decennali e su dati amministrativi: <http://data.worldbank.org/data-catalog/global-bilateral-migration-database>

La Banca Mondiale rende disponibili anche altri *dataset* sull'immigrazione per genere, età di immigrazione e livello di istruzione per la maggior parte dei paesi del mondo nel 1990 e 2000, accessibili da: <http://econ.worldbank.org/programs/migration#Data>

- **Dossier Statistico Immigrazione Caritas/Migrantes**

Annuario socio-statistico sull'immigrazione in Italia, curato dalla Caritas e dalla Fondazione Migrantes. Raccoglie informazioni da diverse fonti statistiche, aggiornate e presentate annualmente sul contesto internazionale, i flussi migratori e gli *stock* in Italia, il mondo del lavoro, i contesti regionali, i rifugiati e il sistema di accoglienza.

Tutte le informazioni sono disponibili su: <http://www.dossierimmigrazione.it>

- **ISMU - Rapporto annuale sulle migrazioni**

La Fondazione ISMU cura un rapporto annuale sull'immigrazione in Italia, che analizza caratteristiche ed evoluzione del fenomeno, con la collaborazione di diversi studiosi.

Una lista dei rapporti è accessibile da: <http://www.ismu.org/index.php?page=82>

La Fondazione ISMU fornisce anche un utile servizio di raccolta e sistematizzazione di dati statistici sull'immigrazione nazionale e internazionale, con *link* alle pagine web di diverse istituzioni, accessibile da:

<http://www.ismu.org/index.php?page=233>

- **ISTAT - Demo-Geodemo**

L'ISTAT mette a disposizione i dati ufficiali più recenti sulla popolazione italiana e straniera residente nei Comuni italiani derivanti dalle indagini effettuate presso gli Uffici di Anagrafe. Sono possibili interrogazioni personalizzate (per

anno, territorio, cittadinanza, ecc.). Sono inoltre disponibili dati storici sui permessi di soggiorno e sui cittadini non comunitari regolarmente presenti in Italia.

Tutti i dati sono accessibili da: *<http://demo.istat.it/>*.

BIBLIOGRAFIA

- ACCETTURO A. - BUGAMELLI M. - LAMORGESE A., «Welcome to the Machine: Firms' Reaction to Low-Skilled Immigration», Banca d'Italia, *Temi di Discussione*, no. 846, 2012.
- ACCETTURO A. - MANARESI F. - MOCETTI S. - OLIVIERI E., «Don't Stand so Close to Me: The Urban Impact of Immigration», Banca d'Italia, *Temi di Discussione*, no. 866, 2012.
- ALTONJI J.G. - CARD D., «The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives», in ABOARD J.M. - FREEMAN R.B. (eds.), *Immigration, Trade and Labor*, Chicago, IL, University of Chicago Press, 1991.
- AMBROSINI M., *Utiles invasori: l'inserimento degli immigrati nel mercato del lavoro italiano*, Milano, F. Angeli, 1999.
- AMUEDO-DORANTES C. - DE LA RICA S., «Complements or Substitutes? Task Specialization by Gender and Nativity in Spain», *Labour Economics*, vol. 18(5), 2011, pages 697-707.
- AYDEMIR A. - BORJAS G.J., «A Comparative Analysis of the Labor Market Impact of International Migration: Canada, Mexico, and the United States», *Journal of the European Economic Association*, no. 5, 2007, pages 663-708.
- , -, «Attenuation Bias in Measuring the Wage Impact of Immigration», *Journal of Labor Economics*, vol. 29(1), 2011, pages 69-112.
- BARONE G. - MOCETTI S., «With a Little Help from Abroad: The Effect of Low-Skilled Immigration on the Female Labour Supply», *Labour Economics*, no. 18, 2011, pages 664-675.
- BARTEL A.P., «Where Do the New United-States Immigrants Live», *Journal of Labor Economics*, vol. 7(4), 1989, pages 371-391.
- BIANCHI M. - BUONANNO P. - PINOTTI P., «Do Immigrants Cause Crime?», *Journal of the European Economic Association*, forthcoming, 2012.
- BLANGIARDO G.C., «The Centre Sampling Technique in Surveys on Foreign Migrants. The Balance of a Multi-Year Experience», United Nations Statistical Commission and Eurostat, *Working Paper*, 2008, pages 12-29.
- BOERI T., «Immigration to the Land of Redistribution», *Economica*, vol. 77(308), 2010, pages 651-687.
- BORJAS G.J., «Immigrants, Minorities, and Labor Market Competition», *Industrial and Labor Relations Review*, no. 40, 1987, pages 382-392.
- , -, «The Economic Analysis of Immigration», in ASHENFELTER O. - CARD D. (eds.), *Handbook of Labor Economics*, vol. 3A, North-Holland, 1999.
- , -, «The Labor Demand Curve Is Downward Sloping: Reexamining The Impact Of Immigration On The Labor Market», *The Quarterly Journal of Economics*, vol. 118(4), 2003, pages 1335-1374.

- .-, «Native Internal Migration and the Labor Market Impact of Immigration», *Journal of Human Resources*, vol. 41(2), 2006, pages 221-258.
- .-, «The Analytics of the Wage Effect of Immigration», *NBER Working Paper*, no. 14796, 2009.
- BORJAS G.J. - FREEMAN R.B. - KATZ L.F., «Searching for the Effect of Immigration on the Labor Market», *American Economic Review Papers and Proceedings*, no. 86, 1996, pages 246-251.
- .-,-.-, «How Much Do Immigration and Trade Affect Labor Market Outcomes?», *Brookings Papers on Economic Activity*, no. 1, 1997, pages 1-67.
- BORJAS G.J. - GROGGER J. - HANSON G.H., «Imperfect Substitution between Immigrants and Natives: A Reappraisal», *NBER Working Paper*, no. 13887, 2008.
- .-,-.-, «Immigration and the Economic Status of African-American Men», *Economica*, no. 77, 2010, pages 255-282.
- .-,-.-, «Substitution between Immigrants, Natives, and Skill Groups», *NBER Working Paper*, no. 17461, 2011.
- BRUECKER H. - FACHIN S. - VENTURINI A., «Do Foreigners Replace Native Immigrants?», *Economic Modelling*, no. 28, 2010, pages 1078-1089.
- CARD D., «The Impact of the Mariel Boatlift on the Miami Labor-Market», *Industrial and Labor Relations Review*, vol. 43(2), 1990, pages 245-257.
- .-, «Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration», *Journal of Labor Economics*, no. 19(1), 2001, pages 22-64.
- .-, «How Immigration Affects US Cities», in INMAN R.P. (ed.) *Making Cities Work: Prospects and Policies for Urban America*, Princeton University Press, 2009a, pages 158-200.
- .-, «Immigration and Inequality», *American Economic Review*, no. 99, 2009b, pages 1-21.
- CARD D. - DINARDO J., «Do Immigrant Inflows Lead to Native Outflows?», *American Economic Review*, vol. 90(2), 2000, pages 360-367.
- CARD D. - DUSTMANN C. - PRESTON I., «Immigration, Wages, and Compositional Amenities», *Journal of the European Economic Association*, vol. 10(1), 2012, pages 78-119.
- CARD D. - LEMIEUX T., «Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis», *Quarterly Journal of Economics*, vol. 116(2), 2001, pages 705-746.
- CARD D. - LEWIS E.-G., «The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts», in BORJAS G.J. (ed.), *Mexican Immigration to the United States*, National Bureau of Economic Research, 2007, pages 193-228.
- CARRINGTON W.J. - DE LIMA P.J.F., «The Impact of 1970s Repatriates from Africa on the Portuguese Labor Market», *Industrial and Labor Relations Review*, no. 49(2), 1996, pages 330-347.

- CINGANO F. - ROSOLIA A., «Non sono concorrenti, ma complementari», *Libertà civili*, n. 2, 2010, pp. 45-56.
- COLOMBO A. - SCIORTINO G. (a cura di), *Assimilati ed esclusi*, Bologna, Il Mulino, 2002.
- CORTES P., «The Effect of Low-skilled Immigration on US Prices: Evidence from CPI Data», *Journal of Political Economy*, vol. 116(3), 2008, pages 381-422.
- CORTES P. - TESSADA J., «Low-Skilled Immigration and the Labor Supply of Highly Skilled Women», *American Economic Journal: Applied Economics*, vol. 3(3), 2011, pages 88-123.
- D'AMURI F. - OTTAVIANO G.I.P. - PERI G., «The Labor Market Impact of Immigration in Western Germany in the 1990s», *European Economic Review*, no. 54, 2010, pages 550-570.
- D'AMURI F. - PERI G., «Immigration and Occupations in Europe», Centro Studi Luca d'Agliano Development Studies, *Working Papers*, no. 302, 2010.
- DEVILLANOVA C., «I costi dell'immigrazione per la finanza pubblica», in Fondazione ISMU, *Quattordicesimo rapporto sulle migrazioni 2008*, Franco Angeli, Milano, 2009.
- .-., «Immigrazione e finanza pubblica», in Fondazione ISMU, *Quindicesimo rapporto sulle migrazioni 2009*, Franco Angeli, Milano, 2010.
- .-., «Immigrazione e finanza pubblica», in Fondazione ISMU, *Sedicesimo rapporto sulle migrazioni 2010*, Franco Angeli, Milano, 2011.
- DOCQUIER F. - RAPOPORT H., «Globalization, Brain Drain and Development», *Journal of Economic Literature*, forthcoming, 2012.
- DUSTMANN C. - FABBRI F. - PRESTON I., «The Impact of Immigration on the UK Labour Market», *Economic Journal*, vol. 115(507), 2005, pages F324-F341.
- DUSTMANN C. - FRATTINI T. - PRESTON I., «The Effect of Immigration along the Distribution of Wages», *Review of Economic Studies*, forthcoming, 2012, doi:10.1093/restud/rds019.
- DUSTMANN C. - GLITZ A., «How Do Industries and Firms Respond to Changes in Local Labor Supply?», *CReAM Discussion Paper*, no. 18/11, 2011.
- DUSTMANN C. - GLITZ A. - FRATTINI T., «The Labour Market Impact of Immigration», *Oxford Review of Economic Policy*, vol. 24(3), 2008, pages 477-494.
- DUSTMANN C. - PRESTON I., «Comment: Estimating the Effect of Immigration on Wages», *Journal of the European Economic Association*, vol. 10(1), 2012, pages 216-223.
- EINAUDI L., *Le politiche dell'immigrazione in Italia dall'Unità a oggi*, Editori Laterza, 2007.
- ELSNER B., «Emigration and Wages: The EU Enlargement Experiment», *IZA Discussion Papers*, no. 6111, 2011.
- FACCHINI G. - MAYDA A., «From Individual Attitudes to Immigration Policy: Theory and Evidence», *Economic Policy*, no. 56, 2008, pages 651-713.

- FACCHINI G. - MAYDA A., «Does the Welfare State Affect Individual Attitudes Towards Immigrants? Evidence Across Countries», *Review of Economics and Statistics*, no. 91, 2009, pages 291-314.
- .-.-. «Individual Attitudes Towards Skilled Migration: An Empirical Analysis Across Countries», *The World Economy*, forthcoming, 2011.
- FASANI F., «The Quest for “La Dolce Vita”? Undocumented Migration in Italy», in TRIANDAFYLLOU A. (ed.), *Irregular Migration in Europe: Myths and Realities*, Ashgate, 2010.
- FARRÉ L. - GONZALEZ L. - ORTEGA F., «Immigration, Family Responsibilities and the Labor Supply of Skilled Native Women», *The B.E. Journal of Economic Analysis & Policy (Contributions)*, vol. 11(1), 2011, page 34.
- FILER R.K., «The Effect of Immigrant Arrivals on Migratory Patterns of Native Workers», in BORJAS G.J. - FREEMAN R.B. (a cura di), *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, University of Chicago Press, 1992, pages 245-269.
- FRATTINI T., *Immigration and Prices in the UK*, mimeo, Università degli Studi di Milano, 2008.
- FREY W.H., «Immigration and Internal Migration Flight - A California Case Study», *Population and Environment*, vol. 16(4), 1995, pages 353-375.
- .-. «Immigration, Domestic Migration, and Demographic Balkanization in America: New Evidence for the 1990s», *Population and Development Review*, vol. 22(4), 1996, pages 741-763.
- FRIEDBERG R.M., «The Impact of Mass Migration on the Israeli Labor Market», *Quarterly Journal of Economics*, vol. 116(4), 2001, pages 1373-1408.
- FRIEDBERG R.M. - HUNT J., «The Impact of Immigration on Host Country Wages, Employment and Growth», *Journal of Economic Perspectives*, vol. 9(2), 1995, pages 23-44.
- GANDAL N. - HANSON G.H. - SLAUGHTER M.J., «Technology, Trade and Adjustment to Immigration in Israel», *European Economic Review*, vol. 48(2), 2004, pages 403-428.
- GASTON N. - NELSON D., «Immigration and Labour-market Outcomes in the United States: A Political-economy Puzzle», *Oxford Review of Economic Policy*, vol. 16(3), 2000, pages 104-114.
- GAVOSTO A. - VENTURINI A. - VILLOSIO C., «Do Immigrants Compete with Natives?», *Labour*, vol. 13(3), 1999, pages 603-621.
- GONZALEZ L. - ORTEGA F., «How Do Very Open Economies Absorb Large Immigration Flows? Evidence from Spanish Regions», *Labour Economics*, vol. 18(1), 2011, page 57-70.
- GROSSMAN J.B., «The Substitutability of Natives and Immigrants in Production», *Review of Economics and Statistics*, no. 54, 1982, pages 596-603.

- HATTON T.J., «A Dual Policy Paradox: Why have Trade and Immigration Policies always Differed in Labor Scarce Economies?», in HATTON T.J. - O'ROURKE K.H. - TAYLOR A.M. (eds.), *The New Comparative Economic History: Essays in Honor of Jeffrey G. Williamson*, Cambridge Mass, MIT Press, 2007.
- HATTON T.J. - TANI M., «Immigration and Inter-Regional Mobility in the UK, 1982-2000», *Economic Journal*, vol. 115(507), 2005, pages F342 - F358.
- HUNT J., «The Impact of the 1962 Repatriates from Algeria on the French Labor-market», *Industrial and Labor Relations Review*, vol. 45(3), 1992, pages 556-572.
- JAEGER D., *Skill Differences and the Effect of Immigrants on the Wages of Natives*, unpublished, William and Mary College, 2007.
- KATZ L. - MURPHY K., «Changes in Relative Wages 1963-1987: Supply and Demand Factors», *Quarterly Journal of Economics*, vol. 107(1), 1992, pages 35-78.
- KERR S.P. - KERR W.R., «Economic Impacts of Immigration: A Survey», *Finnish Economic Papers*, vol. 24(1), 2011, pages 1-32.
- LEAMER E.E. - LEVINSOHN J., «International Trade Theory: The Evidence», in GROSSMAN, G. - ROGOFF K. (eds.), *Handbook of International Economics*, vol. 3, North-Holland, 1995, pages 1339-1394.
- LEWIS E.G., «Immigration, Skill Mix, and Capital-Skill Complementarity», *Quarterly Journal of Economics*, vol. 126(2), 2011, pages 1029-1069.
- LONGHI S. - NIJKAMP P. - POOT J., «A Meta-Analytic Assessment of the Effects of Immigration on Wages», *Journal of Economic Surveys*, vol. 19(3), 2005, pages 451-477.
- , «Meta-Analysis of Empirical Evidence on the Labor Market Impacts of Immigration», *Region et Developement*, no. 27, 2008, pages 161-190.
- MAYDA A., «Why Are People More Pro-Trade than Pro-Migration?», *Economics Letters*, vol. 101(3), 2008, pages 160-163.
- MANACORDA M. - MANNING A. - WADSWORTH J., «The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain», *Journal of the European Economic Association*, vol. 10(1), 2012, pages 120-151.
- MILANOVIC B., *Worlds Apart: Measuring Global and International Inequality*, Princeton University Press, 2005.
- MOCETTI S. - PORELLO C., «How Does Immigration Affect Native Internal Mobility? New Evidence from Italy», *Regional Science and Urban Economics*, vol. 40(6), 2010, pages 427-439.
- MUNSHI K., «Networks in the Modern Economy: Mexican Migrants in the US Labor Market», *Quarterly Journal of Economics*, vol. 118(2), 2003, pages 549-599.
- OKKERSE L., «How to Measure the Labour Market Effects of Immigration: A Review», *Journal of Economic Surveys*, vol. 22(1), 2008, pages 1-30.
- OLIVIERI E., *Immigration and Prices*, mimeo, 2010.

- OTTAVIANO G.I.P. - PERI G., «Rethinking The Effect Of Immigration On Wages», *Journal of the European Economic Association*, vol. 10(1), 2012, pages 152-197.
- PERI G. - SPARBER C., «Task Specialization, Immigration and Wages», *American Economic Journal: Applied Economics*, vol. 1(3), 2009, pages 135-169.
- .-,-.-, «Highly-Educated Immigrants and Native Occupational Choice», *Industrial Relations*, vol. 50(3), 2011a, pages 385-411.
- .-,-.-, «Assessing Inherent Model Bias: An Application to Native Displacement in Response to Immigration», *Journal of Urban Economics*, vol. 69(1), 2011b, pages 82-91.
- PELLIZZARI M., «The Use of Welfare by Migrants in Italy», *International Journal of Manpower*, forthcoming, 2012.
- PISCHKE J.S. - VELLING J., «Employment Effects of Immigration to Germany: An Analysis Based on Local Labor Markets», *Review of Economics and Statistics*, vol. 79(4), 1997, pages 594-604.
- RAPHAEL S. - SMOLENSKY E., «Immigration and Poverty in the United States», *The American Economic Review*, vol. 99(2), *Papers and Proceedings of the One Hundred Twenty-First Meeting of the American Economic Association*, 2009, pages 41-44.
- ROMITI A., «Immigrants-Natives Complementarities in Production: Evidence from Italy», *CeRP Working Paper*, no. 105/11, 2011.
- ROMITI A. - ROSSI M., «Should We Retire Earlier in Order to Look After Our Parents? The role of Immigrants», *CeRP Working Paper*, no. 124/11, 2011.
- ROWTHORN R., «The Fiscal Impact of Immigration on the Advanced Economies», *Oxford Review of Economic Policy*, no. 24, 2008, pages 560-580.
- QUISPE-AGNOLI M. - ZAVODNY M., «The Effect of Immigration on Output Mix, Capital, and Productivity», *FRB Atlanta Economic Review*, Q1, 2002, pages 17-27.
- STAFFOLANI S. - VALENTINI E., «Does Immigration Raise Blue and White Collar Wages of Natives? The case of Italy», *Labour*, vol. 24(3), 2010, pages 295-310.
- VENTURINI A., «Do Immigrants Working Illegally Reduce the Natives' Legal Employment? Evidence from Italy», *Journal of Population Economics*, vol. 12(1), 1999, pages 135-154.
- VENTURINI A. - VILLOSIO C., «Labour Market Effects of Immigration into Italy: An Empirical Analysis», *International Labour Review*, vol. 145(1-2), 2006, pages 91-118.
- YANG D., «Migrant Remittances», *Journal of Economic Perspectives*, vol. 25, no. 3, 2011, pages 129-152.
- ZINCONE G. (a cura di), *Primo rapporto sull'integrazione degli immigrati in Italia*, Bologna, Il Mulino, 2000.
- .-, (a cura di), *Secondo rapporto sull'integrazione degli immigrati in Italia*, Bologna, Il Mulino, 2001.

RASSEGNA BIBLIOGRAFICA

Recensione del volume:
**The Globalization Paradox: Democracy and
the Future of the World Economy***

DANI RODRIK

a cura di

Giovanni Leghissa**

Università di Torino

Che Dani Rodrik ci metta in guardia di fronte ai pericoli della globalizzazione può suonar strano: l'ha sempre difesa strenuamente, e con argomenti piuttosto convincenti.¹ Anche in questo libro, del resto, non viene invocato alcun freno alla globalizzazione in quanto tale. Il titolo però ci ricorda che, di contro a una globalizzazione intelligente, ce ne può essere una stupida e pericolosa. È quest'ultima che deve essere invece combattuta, e con ogni mezzo. Ne va infatti non solo e non tanto della sopravvivenza del nostro sistema economico, ma soprattutto del nostro sistema democratico.

La tesi di Rodrik al riguardo è tanto semplice quanto meritevole di essere presa molto sul serio. Vengono messi in gioco tre elementi, tutti altrettanto importanti, di quella complessa rete istituzionale e organizzativa che costituisce il mondo della vita nell'età del capitalismo globale: gli stati nazionali, la politica democratica, l'iperglobalizzazione (il prefisso qui si riferisce a quel tipo di globalizzazione che non pone alcun freno alla circolazione transnazionale di prodotti finanziari). Ora, questi tre elementi, per loro natura, accettano interazioni tali da permettere che solo due su tre possano coesistere. Vivere in un mondo in cui convivano iperglobalizzazione, stati nazionali e politiche democratiche non è strutturalmente possibile. Rodrik ci invita dunque a sciogliere il seguente trilemma. O si sceglie di vivere in stati nazione democratici e sovrani, e allora si devono adottare misure

* Trad. it. *La globalizzazione intelligente*, Laterza, Roma-Bari, 2011.

** <giovannileghissa@yahoo.it>. Ricercatore confermato presso il Dipartimento di Filosofia e Scienze dell'Educazione dell'Università di Torino.

¹ Cfr. RODRIK D. (1997, 1999).

che contengano la volatile ubiquità e l'ingegnosa creatività dei mercati finanziari. Oppure si lasciano liberi i prodotti finanziari di rimbalzare tra Tokyo, Londra e New York alla velocità della luce, come fanno ora, ma si cerca di dare vita – e in tempi rapidi – a una qualche forma di *governance* mondiale, per evitare che gli spiriti animali di cui quei prodotti sono espressione finiscano per fagocitare il mondo. In questo caso si rinuncia alla forma di sovranità statale che, per semplificare, potremmo chiamare *post-westfalica*, e in compenso quelle forme di garanzia dei diritti individuali e collettivi che trovano ancoraggio nell'azione dei singoli stati vengono gestite globalmente. Terza e ultima ipotesi: non si pone alcun freno all'iperglobalizzazione, nel contempo si lasciano in pace gli stati nazionali, ciascuno libero di muoversi nell'arena globale a proprio piacimento (ovvero a seconda dello spazio di manovra di cui dispone), ma ci si rassegna a vivere in un mondo sempre meno democratico, dal momento che i disastri sociali causati progressivamente dall'iperglobalizzazione costringono via via gli stati nazionali a ridurre gli spazi di democrazia; a tale riduzione si giungerebbe, in tempi probabilmente molto rapidi, per gestire in qualche modo quella conflittualità sociale che si produrrebbe, in modo quasi automatico, non appena fosse compiuto lo smantellamento delle reti di protezione sociale, effetto indiretto ma inevitabile dell'invasività dei mercati finanziari a livello locale.

Non si fa fatica a capire, leggendo il libro, che l'Autore accorda la propria preferenza alla prima ipotesi. A chi gli potrebbe rimproverare che si tratta di una soluzione intrisa di keynesianismo (parola alquanto bistrattata, alla quale si associa di solito il sapore di una pietanza che ormai nessuno si sognerebbe di apprezzare perché legata a un gusto passato di moda, non più in grado di soddisfare palati fini e raffinati), Rodrik risponde per le rime, puntando essenzialmente su un unico argomento, per altro difficilmente confutabile: le politiche globali ispirate a Keynes, uscite insomma dagli accordi di Bretton Woods, hanno garantito pace, prosperità e democrazia per un lungo periodo di tempo, e non solo nei paesi ricchi del pianeta. Si trattava di politiche che, per definizione, non intendevano mettere il carro della globalizzazione davanti ai buoi dell'interesse nazionale: era quest'ultimo, semmai, a trainare quella. Ogniqualvolta subentrava un conflitto tra gli interessi di un'economia alle prese con i propri piani di sviluppo, le proprie specificità – o, al limite, le proprie lentezze – e gli imperativi dettati dai mercati, vi era sempre la possibilità di fare appello al supremo interesse nazionale. E di ottenere ragione. Con tali argomenti ci si riduce a fare l'apologia del protezionismo, nemico dei mercati? Niente affatto. Come mostra bene Rodrik citando come esempi probanti varie situazioni di conflitto che avevano caratterizzato il gioco

degli scambi gestiti secondo le regole di Bretton Woods, dietro gli imperativi dei cosiddetti “mercati” si nascondevano sempre gli interessi di un altro stato nazionale, di un altro sistema economico, magari più aggressivo a livello internazionale, ma pur sempre legato alla storia, alla cultura e alle sorti di una rete produttiva e commerciale incapsulata entro i confini di un territorio sovrano. In altre parole: più che tirar fuori lo spauracchio del protezionismo, è bene essere realisti – il che qui significa, anche, muniti di senso storico – e non negare l’evidente intreccio che da sempre lega assieme economie nazionali e sviluppo regolato del commercio internazionale. Quest’ultimo per definizione non si dà al di fuori di regole, e le regole vengono stipulate entro la cornice di accordi internazionali, siglati dai rappresentanti di singoli stati. I quali, nel siglare tali accordi, hanno sempre in mente gli interessi degli stati che rappresentano. Rodrik cita, per rendere perspicuo il proprio discorso e legarlo all’immediata attualità, il ben noto esempio dell’entrata della Cina nel sistema gestito dal WTO: l’11 dicembre del 2001 la Cina finalmente accetta di seguire le regole del commercio internazionale, ma ci si “dimentica” di far presente ai cinesi che il mantenimento di una moneta debole non risulta confacente alle suddette regole. Risultato: i cinesi non si sognano nemmeno per un istante di rivalutare il *renminbi*, e quindi di fatto è come se sovvenzionassero la propria economia, avvantaggiando in modo artificiale le proprie esportazioni, in barba ai principi basilari sanciti dal WTO. Cinesi cattivi? No, attenti a trarre vantaggi da un sistema economico globale che comunque non potrà mai prescindere dal ruolo giocato dai singoli stati nazionali.

Non è scontata l’insistenza con cui Rodrik, nel corso di tutta la trattazione, sottolinea il ruolo giocato dagli stati nazionali. Dall’uscita del famoso libro di Susan Strange² in avanti, non solo a livello accademico, ma soprattutto a livello mediatico, fino a diventare parte del senso comune, si è imposta la *vulgata* secondo la quale il termine globalizzazione andrebbe inteso come sinonimo di fine del peso giocato dallo stato nazione di fronte all’onnipotenza dei mercati – e ciò non solo in politica estera, ma persino in materia di politiche sociali, per definizione soggette alla sovranità di parlamenti e organi centrali di uno stato. Come richiamato sopra, sulla scorta di una letteratura ormai consolidata sull’argomento,³ Rodrik ci ricorda non solo che ai tempi di Bretton Woods l’espansione del commercio internazionale è convissuta con un sano tasso di conflittualità tra stati

² STRANGE S. (1996, 1998).

³ Rodrik cita ampiamente RUGGIE J.G. (1982), ormai divenuto un vero e proprio classico sull’argomento, da integrare proficuamente con HELLEINER E. (1994).

che perseguivano interessi nazionali diversi, ma anche che nell'età presente, in cui si tenta di venir fuori dalle secche del Doha Round, a tirare le fila del dibattito sulla liberalizzazione dei servizi e sull'opportunità o meno di sovvenzionare i piccoli agricoltori sono i rappresentanti degli stati, e non quelli delle grandi *corporations*. Alle orecchie di alcuni l'argomento suonerà specioso, o per lo meno insufficiente, dal momento che si suppone risulti chiaro a tutti che i rappresentanti degli stati agiscono in nome e per conto delle grandi *corporations* – sussunte collettivamente, a volte, sotto il titolo di “capitale globale”. In realtà, è tutt'altro che evidente l'assunto secondo cui a tirare le fila delle politiche nazionali sarebbe il capitale globale – un attore, quest'ultimo, dai contorni mai precisati e definiti sino in fondo, *et pour cause*: la funzione retorica di tale entità è infatti quella di giustificare l'operatività di una dimensione mitica in seno al discorso chiamato a descrivere il gioco dei mercati. Rodrik si guarda bene dall'introdurre simili entità mitiche ed è molto attento a far presenti i limiti di un modello esplicativo che attribuisce un potere quasi assoluto all'attore che opera nei mercati. Quest'ultimo acquista potere, se del caso, sempre e solo in presenza di attori istituzionali che glielo concedono. Ed è in questo senso che va inteso il richiamo storico di Rodrik alle politiche internazionali nate da Bretton Woods: non si tratta, con un gesto passatista, di aspirare a farle rivivere in modo puro e semplice; si tratta piuttosto di prendere coscienza del carattere non inevitabile, non necessario, non prescritto da alcun destino ineluttabile, delle politiche che attualmente favoriscono l'incontrollata proliferazione ed espansione, al di fuori di qualsiasi controllo statale, di quella rete societaria che, a livello globale, produce capitale finanziario e minaccia sia di distruggere le reti della produzione materiale, sia la funzione sociale di quest'ultima, intesa quale perno attorno cui far ruotare il benessere collettivo.

Merita ricordare che Rodrik, su questo punto, si trova ormai in buona compagnia. Quanti tra i sociologi hanno voluto riprendere le fila del discorso iniziato a suo tempo da Polanyi e hanno sottolineato l'*embeddedness* dell'economico nel sociale,⁴ hanno poi ampiamente mostrato come i capitali non circolino mai nella forma di puri spiriti, spinti da un conato che ne predeterminerebbe la moltiplicazione (in base alla legge secondo cui il capitale affluisce immancabilmente dove il suo rendimento è più efficiente); essi si muovono, piuttosto, in consonanza tanto con le motivazioni individuali, incluse quelle non articolabili nella forma dell'autointeresse, quanto con i rapporti di forza che regolano in vario modo accessibilità a risorse e informazioni e predeterminano quindi la posizione da cui

⁴ GRANOVETTER M. (1985, 1991).

ciascun attore parte allorché decide di partecipare ai giochi dello scambio. Precisamente tale intreccio – che lega l’azione economicamente rilevante, la presenza e l’operatività di cornici motivazionali e cognitive che potremmo in senso lato definire anche come “culturali”, la cogenza dei rapporti di potere che attraversano ogni rete di attori – è ciò che serve a spiegare il funzionamento tanto dei mercati in generale⁵ quanto di quelli finanziari in particolare.⁶ Partire da questo punto di vista è la premessa indispensabile – *in primis* a livello metodologico – per poter poi impostare un discorso minimamente coerente sulle politiche miranti a frenare – o a favorire – la libera circolazione dei flussi di danaro prodotti entro il mercato finanziario. Il punto è cruciale non appena si consideri il modo in cui si articola la difesa dell’iperglobalizzazione. Quest’ultima viene per lo più presentata come la logica e inevitabile conseguenza, anzi come il compimento, della liberalizzazione su scala globale di beni e servizi. Questa mossa dovrebbe suscitare immediate perplessità, stante l’enorme differenza tra la liberalizzazione del commercio di merci come automobili o cotone e la circolazione non regolata di prodotti finanziari la cui massa supera ormai di molte volte il PIL mondiale.⁷ In realtà è una mossa decisiva, che permette non solo di occultare la natura peculiare della liberalizzazione dei prodotti finanziari, ma anche di trarre vantaggio dal successo in precedenza ottenuto dagli argomenti a favore della liberalizzazione del commercio su scala globale. Questi ultimi si erano imposti non attraverso l’esplicitazione – e quindi la discussione pubblica – delle scelte politiche che avevano condotto i maggiori paesi del mondo industrializzato a promuovere la liberalizzazione del commercio mondiale, bensì, al contrario, attraverso il rimando al carattere inevitabile e necessario, ovvero non dipendente da scelte istituzionali, del processo di globalizzazione. Sulla stessa linea, le decisioni che hanno reso possibile la libera circolazione dei flussi finanziari si sono articolate lungo i sentieri ben noti delle teorie che deducevano i vantaggi della liberalizzazione dalla capacità innata dei mercati di autoregolarsi – spinti a ciò non dall’ormai obsoleta mano invisibile, bensì dalla conformità del comportamento di ciascun attore ai dettami della *Rational Choice Theory*. A nulla è mai valso il riferimento alla bassa empiria, ovvero al fatto che i

⁵ BECKERT J., DIAZ-BONE R. e GANßMANN H. (2007).

⁶ KNORR-CETINA K. e PRED A. (2005).

⁷ Un dato abbastanza recente: nel 2007 il valore degli attivi finanziari era di 241 trilioni di dollari, mentre il PIL mondiale ammontava a 54 trilioni (per contro, nel 1980 il valore degli attivi finanziari e il PIL mondiale erano più o meno pari). Su ciò, cfr. PALMA J.G. (2009), che offre anche un’acuta analisi sociologica di come si sia giunti a un tale immenso accumulo di ricchezza nelle mani dei detentori di capitale finanziario.

mercati non si regolano da soli, oppure al fatto che la mancanza totale di regole produce disastri, sia su scala nazionale che internazionale – in riferimento ai primi, la mente corre subito al fallimento di Long-Term Capital Management, l'*hedge fund* gestito da John Meriwether, il cui fallimento, nel 1998, costituisce se vogliamo la madre di tutti i fallimenti finanziari a venire;⁸ quanto ai disastri internazionali, non c'è che l'imbarazzo della scelta: si può pensare alla crisi monetaria del 1992/93, alla crisi che investì le Tigri asiatiche a partire dal 1997, oppure alla crisi argentina dei primi anni 2000, ampiamente citate e discusse da Rodrik nel suo libro. Ormai forti di un consenso generalizzato, capace di contenere le spinte critiche che provengono sia dall'accademia che dalla società, i sostenitori dell'iper-globalizzazione hanno sempre accolto con un'alzata di spalle ogni riferimento ai dati empirici appena ricordati, e hanno continuato ad avvalersi del richiamo, evocato sopra, alla natura destinale del moto globale che conduce inevitabilmente all'iper-globalizzazione. In consonanza sia con i rappresentanti della scuola regolazionista,⁹ sia con altri autori che più di recente hanno riconsiderato il problema,¹⁰ Rodrik in questo libro sottolinea tuttavia la necessità di porre un freno al flusso globale del capitale finanziario, pena il venir meno di ogni capacità, da parte dei cittadini del pianeta, di prendere decisioni autonome e responsabili in materia di distribuzione delle risorse e di gestione del bene comune. Qualche passo in questa direzione, almeno negli Stati Uniti, è stato fatto, anche se non tutti ritengono che il Dodd-Frank Act – e la cosiddetta Volcker Rule in esso contenuta – serva poi molto al fine di contenere la creatività socialmente irresponsabile dei *manager* della finanza.¹¹ Tuttavia, nella situazione in cui siamo, da qualche parte bisogna pur cominciare. L'Europa, dal canto suo, pare per il momento voler restar fedele all'architettura istituzionale voluta a suo tempo da quanti, come per esempio Jacques Delors, si fecero accaniti propugnatori della deregolamentazione del mercato della finanza. E fin che manca una politica comune a livello globale, concordata da attori sovrani, soluzioni come per esempio l'introduzione della Tobin Tax (che Rodrik non sembra scartare, almeno in linea di principio) restano assai lontane, immerse nel mondo dei pii desideri.

⁸ LOWENSTEIN R. (2000) e BOOKSTABER R.M. (2007).

⁹ Si veda per esempio AGLIETTA M. e RIGOT S. (2009).

¹⁰ GRAZ J.C. e NÖLKE A. (2008); HELLEINER E., PAGLIARI S. e ZIMMERMANN H. (2010); DOYRAN M.A. (2011).

¹¹ CHATTERJEE R.R. (2011).

È forse da qui che bisogna partire per prendere in considerazione l'ipotesi di una *governance* globale? Constatando le difficoltà che gli stati sovrani incontrano nel trovare accordi che permettano di dar vita a una qualche forma di regolamentazione dei mercati finanziari, pare quasi scontato rivolgersi all'ipotesi secondo cui al posto di un semplice accordo tra stati debba subentrare una vera e propria entità sovranazionale, in grado, questa sì, di imporre regole vincolanti a tutte le organizzazioni che agiscono nei mercati finanziari globali. Rodrik analizza a fondo questa ipotesi, giungendo però alla conclusione che una *governance* globale non possa costituire la soluzione del problema.

A chi pensare quando si evoca tale soluzione? Le prime istituzioni che possono venir ragionevolmente considerate sono gli organismi internazionali attualmente esistenti, ai quali si potrebbe conferire un potere maggiore – un potere, cioè, che permetta loro di sanzionare quegli attori che si discostano da una serie di regole in precedenza statuite. Ora, perché il discorso sulla *governance* globale abbia un senso, è doveroso auspicare che le istituzioni transnazionali chiamate a garantire l'applicazione delle regole siano effettivamente rappresentative, siano cioè in grado di svolgere quella funzione di mediazione politica che ora viene svolta dagli stati nazionali. Ma in virtù di quale ingegneria istituzionale si giungerebbe poi a rendere effettivi simili meccanismi di rappresentanza? Come si coniugherebbe l'efficienza di un apparato istituzionale che non si può che immaginare estremamente complesso con la trasparenza delle procedure di *accountability* che servirebbero a garantire che i rappresentanti della popolazione mondiale agiscano effettivamente nell'interesse di quest'ultima? Non si correrebbe il rischio di veder agire, all'interno delle sedi istituzionali transnazionali che stiamo qui immaginando, la stessa *élite* composta da burocrati ed esperti che ora popola la Banca Mondiale, il Fondo Monetario Internazionale e la miriade di altre istituzioni chiamate a implementare la liberalizzazione dei mercati su scala globale? Come mostrato a suo tempo da Stiglitz, è del tutto vano fare affidamento sulla speranza che questa *élite* possa mai prendere decisioni che risultino contrarie o non conformi agli interessi delle organizzazioni attive nel settore della finanza globale.¹² Tale *élite* viene quasi sempre reclutata all'interno della classe capitalista transnazionale,¹³ di cui condivide interessi e modelli mentali. Questi ultimi vanno ben tenuti presenti quando si affronta la questione del ruolo da attribuire a un gruppo di individui chiamati, in qualità di esperti, a prendere decisioni rilevanti per la collettività. Chiunque operi

¹² STIGLITZ J.E. (2002, 2006).

¹³ SKLAIR L. (2001); CARROLL W.K. (2010).

in seno a un'istituzione accetta determinate regole del gioco, restrizioni formali e informali, *patterns* cognitivi; questi, a loro volta, non possono discostarsi troppo da quelle rappresentazioni condivise la cui funzione consiste precisamente nel rendere possibile il buon funzionamento e la riproduzione dell'istituzione entro cui operano gli attori. Avendo in mente questa mutua compenetrazione tra modelli mentali condivisi e regole istituzionali, si dovrebbe allora supporre che i futuri rappresentanti della collettività globale possano venir reclutati tra quella schiera ormai sempre più folta di individui che hanno maturato la convinzione che, per il bene dell'umanità, sia necessario limitare il flusso globale dei mercati finanziari. Al limite seguaci di Polanyi, comunque lettori attenti di Amartya Sen e di quanti, come Pogge, stanno lavorando alla definizione di un nuovo concetto di giustizia globale, ma anche in grado di apprezzare la sostanziale differenza tra *Austrian Economics* e i dettami della Scuola di Chicago, i futuri artefici dell'auspicata *governance* globale dovrebbero insomma emergere dalla rete che unisce la confusa e variegata galassia di quanti oggi sono consci del fatto che l'iperglobalizzazione rischia tanto di distruggere l'ecosistema globale quanto di negare a una fetta enorme di individui la stessa possibilità di condurre una vita decente. Salta agli occhi, però, una difficoltà: quanto variegata ce la vogliamo immaginare questa galassia? Ipotizziamo che si estenda dal movimento *Occupy Wall Street* ai guerriglieri naxaliti che combattono nelle foreste del subcontinente indiano? Anche a voler mettere da parte la facile ironia, pare proprio non sia facilmente percorribile la strada che conduce alla formazione di una classe politica transnazionale a partire da una società civile consapevole, informata, critica – senza quest'ultima, come insegna Dahl, non vi è autentica democrazia.

Certo, voler essere realisti a tutti i costi e tarpare le ali allo spirito dell'utopia, in una fase storica come l'attuale, può risultare quasi vile, o almeno poco lungimirante. Tuttavia, prima di entusiasinarsi per uno scenario utopico in cui i guasti del mercato finanziario globale vengono tenuti a freno da una democrazia globale ancora tutta da inventare, è bene tener presente l'ultima variante del trilemma proposto da Rodrik. Magari proprio da qui lo slancio utopico trae rinnovata energia – oppure, più semplicemente, ci si mette alla ricerca di soluzioni realisticamente praticabili nell'immediato, al fine di scongiurare la catastrofe.

La catastrofe in questione scaturirebbe dall'ovvia e immediata reazione degli organismi nazionali e internazionali preposti a difendere l'ordine costituito di fronte all'immane ondata di proteste che verrebbe generata dalla dissoluzione di ogni forma di protezione sociale. Quest'ultima, a sua volta, sarebbe l'altrettanto ovvia conseguenza dell'iperglobalizzazione – un mercato finanziario senza limiti,

infatti, prima o poi metterebbe le mani, direttamente o indirettamente, sulla rete di servizi gestita dagli stati, determinandone l'inaccessibilità a fette molto larghe di popolazione. Si tratta di una eventualità che, in quanto *worst case*, probabilmente non viene esclusa da coloro che non arretrerebbero dinanzi a nessun ostacolo pur di realizzare compiutamente l'agenda neoliberale, vero cuore dell'iperglobalizzazione. Rodrik non si dilunga sulla questione da lui stesso sollevata, ma va ascritto comunque al suo libro il merito di aver concepito questa possibilità. A chi pensasse che qui si tratti di un puro esercizio accademico, vanno ricordate due cose. Primo, in questo momento l'iperglobalizzazione sta provocando un aumento della povertà che rischia di erodere i vantaggi che la globalizzazione intelligente aveva recato con sé nel recente passato. Tale aumento della povertà può ben costituire la premessa per un aumento del disordine sociale su vasta scala. Secondo, una volta che una fetta della popolazione, spinta dalla mancanza di risorse, si organizza per dar vita a forme di protesta anche violente, è probabile che la restante parte della popolazione accetti una diminuzione sostanziale delle libertà civili e politiche pur di veder regnare la pace sociale. Tanto più che quest'ultima verrebbe garantita, probabilmente, non solo e non tanto dall'introduzione di misure repressive eccezionali, quanto piuttosto da un mirato aumento dei meccanismi di controllo già attualmente operanti – il che renderebbe in qualche misura meno gravosa la rinuncia alla libertà. Se ci chiediamo come mai è ragionevole supporre che una maggioranza silenziosa, in caso di disordini sociali, potrebbe barattare la libertà per ottenere in cambio sicurezza, va ricordato che, come mostrato da un'ampia letteratura, la sicurezza da sempre costituisce non solo uno dei beni primari presenti nel paniere di scelte di cui dispone un attore, ma anche uno di quei beni che più spesso vengono preferiti.¹⁴ Ripetuti esperimenti in cui si gioca il gioco del dittatore insegnano che gli umani tendenzialmente sono inclini se non ad attribuire un alto valore alla giustizia distributiva, per lo meno a provare fastidio di fronte a smaccate ingiustizie. Tuttavia, alta – forse troppo alta – dovrebbe essere la ricompensa per indurre numeri molto consistenti di individui a preferire la solidarietà verso gli esclusi e quindi a mettere a repentaglio la propria vita per impegnarsi in lotte sociali dall'esito incerto al fine di garantire un accesso ai servizi primari non dipendente dal reddito dei cittadini.

Nel caso precedentemente esaminato si trattava di immaginare un aumento del numero di individui disposti a trasformare le spinte provenienti dalla società

¹⁴ Da ultimo, HIBOU B. (2011).

civile in creazione di stabili strutture democratiche globali – di immaginare cioè che gli attuali clienti del mercato globale inizino a trasformarsi in cittadini della democrazia cosmopolita, come auspicato, tra gli altri, da Barber.¹⁵ Nell'ultimo caso esaminato si tratta, con uno sforzo di immaginazione non minore, di ipotizzare che la lotta per la democrazia e la difesa di quei diritti sociali faticosamente acquistati in alcuni paesi del mondo diventi interesse primario dei cittadini degli stati nazionali più esposti all'invadenza dei mercati finanziari. Molto minore, però, sembra lo sforzo di immaginazione richiesto da Rodrik quando ci presenta l'ipotesi di una nuova Bretton Woods dei mercati finanziari. Certo, nulla ci garantisce che la classe politica attualmente al governo nei vari paesi del mondo sia in grado di comprendere la necessità di introdurre – e in tempi rapidi – misure simili alla Tobin Tax o al Glass-Steagall Act votato dal Parlamento statunitense nel 1933 al fine di separare banche di deposito e banche di investimento e abrogato nel 1999. Si tratterebbe di passi in avanti non secondari, in grado di indebolire se non altro il peso degli attori che operano nel mercato finanziario – esiste infatti un legame essenziale tra peso politico raggiunto dalle organizzazioni operanti nel mercato finanziario e la complessità delle loro strutture societarie. Né alcunché ci assicura che l'attuale classe politica internazionale trovi la forza per intervenire in modo restrittivo sulle operazioni *over the counter* – intervento altrettanto necessario, in quanto è l'immenso volume e l'intrinseca non trasparenza di queste operazioni (se fossero trasparenti non sarebbero redditizie!) a costituire la premessa per ogni crollo sistemico a venire.¹⁶ Tuttavia, pur senza indulgere in facili ottimismo, Rodrik ci ricorda che nulla impedisce all'attuale classe dirigente del pianeta di attuare misure come quelle appena prospettate – se solo lo volesse.

Se solo lo volesse! – verrebbe da dire con un sospiro. E con ciò siamo portati a una riflessione conclusiva, concernente il valore formativo della scienza economica. Cosa può formare la volontà, cosa può convincere la classe dirigente a intraprendere misure di tale efficacia da porre un freno alla libera e incontrollata circolazione di flussi di capitale su scala globale? La risposta è semplice ma non banale: un'altra mentalità, un insieme di modelli mentali condivisi alternativi rispetto a quelli attualmente dominanti. La formazione di questi ultimi richiede senza dubbio la diffusione di una scienza economica diversa da quella insegnata nelle maggiori università del mondo. Il libro di Rodrik è costellato da continui riferimenti a questo aspetto del problema. Vi sono pagine in cui l'Autore ironizza

¹⁵ BARBER B.R. (2008, 2010).

¹⁶ CROTTY J. (2008).

sulla presunta scientificità dei modelli costruiti da tanti suoi colleghi – e sulla tendenza degli economisti a non voler quasi mai discutere i fondamenti sui cui tale scientificità poggia dal punto di vista epistemologico. A tal proposito, secondo Rodrik, è urgente tornare a considerare la scienza economica come una scienza sociale, e non come una scienza naturale. Non più chiamata a sviluppare modelli le cui condizioni di partenza sono tanto irrealistiche quanto difficilmente verificabili (condizioni rese però stabili dalla magica formuletta *ceteris paribus*), una scienza economica ricondotta nell'alveo delle scienze sociali non solo non si renderebbe complice dei deliri della finanza globale (sono noti i danni che può causare un utilizzo meccanico e dogmatico delle equazioni di Black-Scholes-Merton), ma potrebbe tornare a dialogare con quei saperi che sono chiamati a formulare e rendere concettualmente chiare le nozioni di diritto, giustizia, equità, libertà, responsabilità, democrazia, eccetera. Come ben sapeva Knight, la sfera economica dipende da quella politica, in quanto è a quest'ultima sfera che compete l'elaborazione dei valori condivisi, senza i quali sarebbe impossibile argomentare a favore di questa o quella scelta economicamente rilevante per la collettività¹⁷ – si pensi solo al caso delle politiche fiscali, o alla politiche chiamate a regolamentare il mercato del lavoro. Non è impossibile trovare buoni studi di epistemologia o di sociologia della conoscenza che mettono in rilievo i contesti entro cui si è sviluppata la scienza economica contemporanea, fino a proporsi come scienza naturale.¹⁸ C'è da sperare che, accanto a questi, il libro di Rodrik, vista la crucialità del trilemma ivi enunciato, spinga la comunità scientifica degli economisti a interrogare con maggior profondità i presupposti epistemologici del proprio sapere al fine di cogliere come i primi siano sempre legati alla valenza politica del secondo.

¹⁷ KNIGHT F.H. (1940).

¹⁸ AMADAE S.M. (2003); CARTWRIGHT N. (2007); FOURCADE M. (2010).

BIBLIOGRAFIA

- AGLIETTA M. - RIGOT S., *Crise et rénovation de la finance*, Jacob, Paris, 2009.
- AMADAE S.M., *Rationalizing Capitalist Democracy. The Cold War Origins of Rational Choice Liberalism*, The University of Chicago Press, 2003.
- BARBER B.R., *Consumati: da cittadini a clienti*, Einaudi, Torino, 2010.
- BECKERT J. - DIAZ-BONE R. - GANßMANN H. (eds.), *Märkte als soziale Strukturen*, Campus, Frankfurt am Main, 2007.
- BOOKSTABER R.M., *A Demon of Our Own Design: Markets, Hedge Funds, and the Perils of Financial Innovations*, Wiley, Hoboken (NJ), 2007.
- CARROLL W.K., *The Making of a Transnational Capitalist Class: Corporate Power in the 21st Century*, Zed Books, London - New York, 2010.
- CARTWRIGHT N., *Hunting Causes and Using Them: Approaches in Philosophy and Economics*, Cambridge University Press, 2007.
- CHATTERJEE R.R., «Dictionaries Fail: The Volcker Rule's Reliance on Definitions Renders It Ineffective and a New Solution Is Needed to Regulate Adequately Proprietary Trading», Columbia Law School, *Working Paper*, 2011.
- CROTTY J., «Structural Causes of the Global Financial Crisis: A Critical Assessment of the "New Financial Architecture"», Economic Department, *Working Paper Series*, no. 16, University of Massachusetts - Amherst, 2008.
- DOYRAN M.A., *Financial Crisis Management and the Pursuit of Power*, Ashgate, Burlington (VT), 2011.
- FOURCADE M., *Economists and Society: Discipline and Profession in the United States, Britain and France, 1890s to 1990s*, Princeton University Press, 2010.
- GRANOVETTER M., «Azione economica e struttura sociale: il problema dell'embeddedness», in MAGATTI M. (a cura di), *Azione economica come azione sociale. Nuovi approcci in sociologia economica*, Angeli, Milano, 1991.
- GRAZ J.C. - NÖLKE A., *Transnational Private Governance and Its Limits*, Routledge, London, 2008.
- HELLEINER E., *States and the Reemergence of Global Finance: From Bretton Woods to the 1990s*, Cornell University Press, Ithaca - London, 1994.
- HELLEINER E. - PAGLIARI S. - ZIMMERMANN H. (eds.), *Global Finance in Crisis: The Politics of International Regulatory Change*, Routledge London, 2010.
- HIBOU B., *Anatomie politique de la domination*, La Découverte, Paris, 2011.
- KNIGHT F.H., «What is Truth in Economics?», *The Journal of Political Economy*, no. 48, 1940.
- KNORR C.K. - PRED A. (eds.), *The Sociology of Financial Markets*, Oxford University Press, 2005.

- LOWENSTEIN R., *When Genius Failed. The Raise and Fall of Long Term Capital Management*, Random House, New York, 2000.
- PALMA J.G., «The Revenge of the Market on the Rentiers. Why Neo-liberal Report on the End of History Turned out to Be Premature», *Cambridge Journal of Economics*, vol. 33, 2009.
- RODRIK D., *Has Globalisation Gone Too Far?*, Inst. for International Economics, Washington, 1997.
- .-, «Globalisation and Labour, or: If Globalisation is a Bowl of Cherries, Why Are There So Many Gum Faces Around the Table?», in BALDWIN R. - COHEN D. - SAPIR A. - VENABLES A. (eds.), *Market Integration, Regionalism, and the Global Economy*, CEPR, Cambridge, 1999.
- RUGGIE J.G., «International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order», *International Organizations*, no. 36, 1982.
- SKLAIR L., *The Transnational Capitalist Class*, Blackwell, Oxford, 2001.
- STIGLITZ J.E., *La globalizzazione e i suoi oppositori*, Einaudi, Torino, 2006.
- STRANGE S., *Chi governa l'economia mondiale? Crisi dello stato e dispersione del potere*, Il Mulino, Bologna, 1998.

Recensione del volume:
The Return to Keynes*

BATEMAN B. - HIRAI T. - MARCUZZO M.C. (eds.)

a cura di

Annalisa Rosselli**

Università di Roma Tor Vergata

Concepito all'indomani del collasso finanziario del 2008 che ha trascinato con sé l'economia reale di quasi tutto il mondo, il titolo attribuito a questo eccellente volume suona oggi già datato e eccessivamente trionfalistico. È più che comprensibile che i curatori l'abbiano scelto al momento di consegnare il lavoro alle stampe: i mesi successivi al fallimento della Lehman Brothers hanno visto infatti quasi tutti i governi varare pacchetti di misure di stimolo all'economia più o meno ampi, fatti di sgravi fiscali, incentivi al consumo, salvataggio di banche insolventi, diminuzione dei tassi di interesse tramite ampliamento dell'offerta di moneta – tradizionali strumenti keynesiani in caso di stallo della domanda del settore privato, la cui efficacia però era stata rifiutata sul piano politico e teorico nei decenni precedenti. Nel momento dell'emergenza lo Stato aveva dunque cessato di essere il problema ed era tornato ad essere la soluzione. Tuttavia si è trattato di un breve ritorno di fiamma. Oggi il titolo dovrebbe rinunciare all'articolo determinativo e diventare semplicemente "Return to Keynes", cioè un'esortazione, un ordine o una preghiera di abbandonare il modello basato sulla fiducia nella libertà dei mercati, incontrollati movimenti di capitale e bilanci pubblici in pareggio che ha già almeno due disastri sulle spalle – quello del 1929 e del 2008 con le conseguenti lunghe recessioni – ma che non appare significativamente scalfato dal suo predominio intellettuale.

* Raccolta di saggi pubblicata da The Belknap Press of Harvard University Press, Cambridge Mass. e London UK, 2010.

** <annalisa.rosselli@uniroma2.it>, Dipartimento di Economia, Diritto e Istituzioni.

Per gli autori dei quattordici capitoli di questo libro però non si è trattato di un ritorno a Keynes. La macroeconomia keynesiana o il pensiero di Keynes sono infatti il tema di una parte molto significativa della loro produzione scientifica che quasi sempre abbraccia molti anni e molte pubblicazioni. In realtà i saggi contenuti nel volume erano stati preparati per un convegno che si è tenuto in Giappone nel 2007 con la partecipazione di alcuni dei massimi esperti di Keynes, ma, come accade sempre in questi casi, hanno visto la luce nella loro forma definitiva e con poche modifiche solo alcuni anni più tardi. Il fatto di essere stati scritti prima della scoppio della crisi però non ne diminuisce il valore. Anzi, in un certo senso lo aumenta perché risaltano maggiormente le loro qualità di prodotto di un approfondito lavoro di ricerca e di riflessione, non dettato da esigenze contingenti da *instant book*, ma guidato dalla volontà di valutare, alla luce di settant'anni di storia di fatti e di idee, la validità del contributo keynesiano e di fare il punto sui problemi interpretativi che l'opera di Keynes ha ancora lasciato aperti. Forse l'unico rimprovero che si può muovere ad alcuni di questi saggi, particolarmente a quelli concentrati sull'aspetto interpretativo, potrebbe essere quello di rivolgersi a lettori che conoscono Keynes già molto bene attraverso la lettura diretta dei suoi lavori e che costituiscono un pubblico ormai non molto vasto. Già nel 1962 Harry Johnson aveva osservato che la *Teoria Generale* di Keynes aveva raggiunto da tempo lo stato di opera classica, cioè "un libro di cui tutti hanno sentito parlare e nessuno ha letto".

Il volume è diviso in quattro parti. La prima è dedicata alla ricostruzione delle vicende nelle tre principali economie mondiali – Stati Uniti, Giappone e Europa – delle politiche di controllo della domanda che vanno sotto il nome di politiche keynesiane e che per molto tempo sono state identificate con il modello IS-LM a cui, alla fine degli anni '50, si è aggiunta la curva di Phillips con la sua proposta di semplice alternativa tra disoccupazione e inflazione. Nessuno degli autori dei tre saggi contenuti in questa prima parte – Bateman, Ono e Trautwein – pur essendo eccellenti conoscitori del pensiero di Keynes, è interessato a rinverdire un'antica polemica sulla corrispondenza di quel modello con le idee del presunto ispiratore. Il loro lavoro parte dall'assunto felicemente esposto da chi il modello IS-LM l'ha creato, John Hicks, in un saggio in cui valutava criticamente il proprio contributo alla costruzione della vulgata keynesiana: «Con teoria keynesiana intendendo solo la teoria formalizzata nei libri di testo. Ci hanno ricordato [...] che in Keynes c'è molto di più e sono totalmente d'accordo. Ma non possiamo mettere in dubbio che c'è anche questo. Quando perdiamo Keynes di vista, lo troviamo che è salito già un altro bel tratto verso la cima della montagna; ma questo è un

posto dove ha piantato per un po' le tende» (Hicks 1979, p. 73, traduzione mia).

È a questo modello, che a quell'epoca prevaleva nell'analisi macroeconomica accademica e nell'approccio dei *policy-maker*, che si riferiva il presidente USA Nixon quando diceva nel 1971 "Siamo tutti keynesiani ora". Ma come Bateman ci illustra con grande chiarezza espositiva nel suo saggio dedicato agli Stati Uniti, in realtà il predominio del keynesismo negli Stati Uniti è stato brevissimo, poco meno di un decennio negli anni '60 sotto le presidenze democratiche di Kennedy e Johnson. Dal punto di vista della politica economica, contrariamente a quanto spesso si dice, non c'è Keynes dietro il New Deal Rooseveltiano, anche se Keynes aveva tentato di convincere Roosevelt fin dalla fine del 1933 a una politica di aumento della spesa pubblica finanziata con l'emissione di titoli. Nonostante le notorie capacità argomentative di Keynes, sarebbe stato del resto ben difficile che un presidente americano si facesse guidare da un economista inglese. Infatti il primo governo che adotta esplicitamente le politiche keynesiane di sostegno alla domanda per convinzione teorica e non per semplice buon senso e urgenza politica, è quello canadese nel 1939. Anche negli anni '50, dominati dalle preoccupazioni ideologiche della Guerra Fredda, gli USA hanno guardato con oscillante diffidenza all'intervento dello Stato nell'economia. È solo a livello accademico che l'influenza della *Teoria Generale* è stata dominante nel dopoguerra e attraverso alcuni dei migliori economisti del tempo è arrivata ai programmi di sostegno dell'occupazione e di spesa sociale dell'era kennediana. La storia successiva è ben nota: il surriscaldamento dell'economia statunitense in seguito alle spese della guerra in Vietnam porta alla fine di Bretton Woods e lascia spazio agli anni della stagflazione. L'impossibilità di applicare le ricette keynesiane per affrontare contemporaneamente disoccupazione e inflazione segna la fine del keynesismo, che è rapida e totale, sia sul piano politico che su quello scientifico. Il monetarismo che prende il suo posto si afferma come la prima versione della dottrina che la politica fiscale è controproducente e la politica monetaria può essere dannosa a meno che non si attenga strettamente a una regola anticipabile dal pubblico. La regola nel corso degli anni cambia, ma nel dibattito tra il partito a favore delle regole contro quello a favore delle politiche discrezionali (che risale almeno alle guerre napoleoniche), è il primo che prende il netto sopravvento a partire fin dagli anni '70. A partire da allora, l'apparato teorico a suo sostegno si evolve ma non abbandona l'innovazione di mettere al suo centro il comportamento dell'agente economico razionale e le sue reazioni agli annunci delle misure dei governi.

Su questo punto i saggi di questo volume sono interessanti per due aspetti che sono messi in dovuto rilievo. Il primo è l'influenza che la teoria economica ha ef-

fettivamente esercitato sulle decisioni dei governi. Nel caso degli Stati Uniti, terminato il decennio dorato clintoniano con lo scoppio della bolla dot.com, abbiamo un ritorno massiccio a politiche aggressivamente keynesiane ben prima dell'11 settembre o del recente collasso finanziario. Dell'anti-keynesismo resta solo l'apparato ideologico e la sua fede che gli agenti economici abbiano aspettative razionali per il loro futuro e i mercati debbano essere lasciati liberi di operare se non si vuole peggiorare la situazione. In Europa invece trionfano le teorie originate negli USA sulla necessità dell'equilibrio di bilancio e sull'inflazione come unico obiettivo per la politica monetaria gestita da un'autorità indipendente e credibile. Spetta a loro giocare un ruolo determinante nella creazione dell'Unione Monetaria a cui pure alcuni paesi avevano aderito nella speranza di recuperare un controllo sulla politica economica che nello SME era di fatto decisa a Bonn.

Il secondo aspetto della relazione tra teoria e realtà messo in evidenza nella prima parte del volume è la cesura che si è manifestata tra teoria ed evidenza empirica. Come illustra ampiamente Trautwein nel suo saggio sulla politica macroeconomica europea, l'evidenza per la Bundesbank, che per tutti gli anni '80 e '90 ha costituito un vero modello di credibilità per la sua indipendenza e politica monetaria di stretto controllo dell'inflazione, non offre sostegno alla teoria prevalente. In base a questa, la Germania sarebbe dovuta passare in quegli anni da un equilibrio con un tasso di inflazione più elevato a uno con saggio più basso o nullo senza allontanarsi troppo dal tasso naturale di output e disoccupazione, muovendosi lungo la curva verticale di Phillips. Invece i costi della disinflazione in termini di output e occupazione non sono stati minori per la Germania e rispetto agli altri paesi che non avevano agganciato la moneta al marco. Ma l'evidenza empirica non sembra aprire una breccia in un apparato teorico che ha solo vacillato dopo gli shock ben maggiori dello scoppio di due bolle (ma i mercati finanziari non erano "efficienti" e gli agenti non avevano aspettative razionali?) che ci hanno portato nell'attuale drammatica situazione.

La seconda parte del volume è dedicata alla Sintesi neoclassica, nella sua forma originaria di Hicks e in quella più recente della "New Neoclassical Synthesis". Alla prima è indubbiamente legato il nome di James Tobin, che non ha mai rinnegato nella sua lunga carriera né i principi ispiratori di Keynes né gli insegnamenti ricevuti in gioventù a Harvard, prendendo le distanze da post-keynesiani e neo-keynesiani nella stessa misura. La battaglia di Tobin contro ogni tentativo di negare esplicitamente o surrettiziamente la disoccupazione involontaria e gli effetti reali della politica monetaria è ricostruita nel saggio di Dimand. Alla Sintesi neoclassica nuova versione sono invece dedicati il saggio di Arena e quello di Bo-

ianovsky e Trautwein. Arena imposta il confronto tra vecchia e nuova Sintesi neoclassica sul concetto di razionalità strumentale – la capacità di individuare la miglior soluzione possibile dato un insieme di informazioni – e razionalità cognitiva, la capacità di effettivamente raggiungere la completezza dell’informazione. Keynes ha sempre accettato la prima e dubitato ampiamente della seconda che è invece elemento comune alla Sintesi neoclassica vecchia e nuova. Boianovsky e Trautwein discutono la nuova Sintesi neoclassica nella versione di Woodford, per esaminare quanto sia fondato il suo richiamo alle teorie del ciclo Wickselliane e pre-*Teoria Generale*.

La terza e la quarta parte del volume spostano l’attenzione su Keynes vero e proprio. Innanzitutto sulla sua persona. A chi pensa che sia impossibile che si possa dire qualcosa di nuovo su Keynes dopo tre ampie biografie – una scritta da Harrod suo contemporaneo e collaboratore, una da Moggridge che ha passato anni a leggerne le carte per curare la pubblicazione della sua opera e una da Skidelsky in tre massicci volumi – bisogna ricordare quale lavoratore instancabile della parola fosse Keynes. I trenta volumi della sua opera contengono solo una parte di quello che ha scritto nel campo dell’economia e quasi nulla in quello della filosofia, per non parlare delle centinaia di lettere, appunti e rapporti che ci ha lasciato. Dostaler, nel capitolo giustamente intitolato “Keynes and the war of words” cerca di darci un’idea del ruolo che la parola ha avuto nella complessità della vita di Keynes e nella molteplicità delle sue attività. È un bel saggio che ci fa ancora più rimpiangere la prematura scomparsa del suo autore. Gli altri capitoli invece riportano l’attenzione sull’opera di Keynes. Ovviamente all’interno di questa la *Teoria Generale* ha un ruolo dominante e Hirai – forse l’unico al mondo ad averne esaminato con pazienza certosina tutte le stesure fino alle ultime bozze – ne ricostruisce la genesi mentre Backhouse ricorda l’accoglienza dell’opera percepita come “troppo matematica” (e spiega perché). Ma il maggiore interesse di quest’ultima parte del volume è nell’esame di due scritti di Keynes, entrambi meno noti anche se per motivi differenti.

Il capitolo diciassettesimo della *Teoria Generale* è sempre risultato ostico per la sua oscurità e complessità, anche se il concetto centrale è relativamente semplice: dal momento che concedere un prestito è come vendere moneta *spot* e comprarla *future* e questa operazione si può fare su molti altri mercati e per molte altre merci, ne segue che come è possibile calcolare dalla differenza tra i due prezzi un tasso di interesse monetario, così dalla stessa differenza si può ricavare un tasso di interesse proprio di ogni merce. In equilibrio prezzi *spot* e *future* sono eguali e tutti i tassi sono uguali a quello monetario; ma cosa accade fuori dall’equilibrio

quando i tassi diventano diversi? Quale è quello dominante? Cosa ha di speciale il tasso di interesse sulla moneta? O, ancora in termini più generali, cosa rende diversa un'economia monetaria da un'economia senza moneta? È forse una delle domande più affascinanti che si possano fare a un economista. Keynes nel capitolo diciassettesimo offre una sua spiegazione che Sraffa non ha mai accettato, anche se non ha mai pubblicato il suo dissenso che trapela da poche lettere e appunti (che Kurz illustra dando ragione a Sraffa). De Cecco e Kregel appoggiano invece la posizione di Keynes, riportando alla luce come la conoscenza dei mercati finanziari abbia permeato la costruzione della *Teoria Generale* in una maniera molto più profonda e strutturale di quanto i passi fin troppo citati sul "concorso di bellezza" e il mercato finanziario trasformato in casinò lascino intendere.

Infine, l'altro scritto di Keynes su cui intervengono Carabelli e Cedrini nel capitolo conclusivo del volume, è il *memorandum* del 1945 che servirà di base alle negoziazioni per la concessione di un prestito degli Stati Uniti alla Gran Bretagna alla fine della Seconda Guerra mondiale. Non è molto conosciuto perché solo in tempi recentissimi si è tornati a parlare del contributo di Keynes al disegno di un modello di ordine mondiale. Nel *memorandum* infatti Keynes cerca di delineare un modello di aggiustamento internazionale in un contesto di forte indebitamento estero, nel tentativo di convincere il più potente alleato che l'interdipendenza delle economie, che la teoria economica incoraggia e la libertà di commercio crea, deve richiedere anche "condivisione delle responsabilità". L'alternativa è la soluzione del mercato che non può essere che l'austerità: *Starvation Corner*, come la definisce Keynes, con tutto il peso dell'aggiustamento sulle spalle più deboli del debitore. È fin troppo facile istituire un parallelo tra l'Inghilterra di allora e gli Stati Uniti di ora – entrambi gravati dalle conseguenze di aver giocato troppo a lungo il ruolo di Grande Consumatore o di "importatore di ultima istanza". Anche ora, ci ricordano gli autori del capitolo, la soluzione dovrebbe essere concordata a livello mondiale in modo equo e a vantaggio di tutti. Ma il *memorandum* non portò le negoziazioni alla soluzione che Keynes avrebbe voluto e i suoi disegni per un diverso ordine mondiale, a differenza delle sue teorie per la stabilizzazione del reddito, non hanno mai avuto nemmeno il beneficio di essere messe alla prova.

Keynes ha dato a una sua raccolta di saggi uscita nel 1931 il titolo di "Essays in Persuasion", ma nella prefazione aveva scritto che l'avrebbe voluta chiamare "Essays in Persuasion and Prophecy" (ed *Esortazioni e profezie* si è chiamata la traduzione italiana) perché purtroppo la *Profezia* aveva avuto maggiore successo che la *Persuasione* (Keynes 1931, p. xvii). Speriamo che in questo non sia stato profeta e che a lungo andare e dopo molti disastri la persuasione si faccia strada. Volumi come questo contribuiscono a mantenerne viva l'esigenza.

BIBLIOGRAFIA

HICKS J., *Causality in Economics*, Oxford, Blackwell, 1979.

KEYNES J.M., «Essays in Persuasion», London, Macmillan, in *CWK* 9, 1931.

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. - SZAPARY 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 settembre 2012

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:

Crescita economica ed economia sommersa. Un approccio di simulazione e verifica empirica per l'Italia
Roberto Dell'Anno

Il controllo dell'economia: bilancio e prospettive
Marisa Faggini

Sulla *political economy* della spesa pubblica nell'Italia liberale
Emma Galli, Roberto Ricciuti

Le scelte di localizzazione delle opere pubbliche: il fenomeno *Nimby*
Roberta Occhilupo, Giuliana Palumbo, Paolo Sestito

Financial Development and Economic Growth. A Theoretical and Empirical Overview
Filomena Pietrovito

ANNO CI - SERIE III
luglio/settembre 2012
Fascicolo VII-IX

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