Tax structure and economic growth: A panel cointegrated VAR analysis

Silvestro Di Sanzo[a,*], Mariano Bella[a,†], Giovanni Graziano[a,‡]
[a]Confcommercio Research Department, Piazza G.G. Belli 2, 00153 Rome, Italy

Abstract

The literature studying the growth effects of tax structure is based on a linear single-equation framework. Using annual data from 1970 to 2012 for a panel of twenty OECD countries, the paper re-evaluates the results of previous research relying on linear and non-linear panel cointegrated VAR models. The asymmetric impact of tax changes on growth is estimated adapting the threshold cointegration methodology proposed by Hansen (2002) to the panel framework. We find that property taxes are the least harmful for the growth while income and consumption taxes are negative for the growth and have comparable effects. Moreover, these findings are robust and significant when the tax burden is above the threshold value of 30%. Policy conclusions are also discussed.

Keywords: Tax structure, Long-run growth, Fiscal policy, Non-linear Panel VAR, Error correction model, OECD.

JEL Classification: C3, C5, E6, H3.

1 Introduction

The relationship between fiscal policies, such as government expenditure and total tax revenue, and growth has been at forefront of the research in theoretical and empirical macroeconomics in the last few decades.

While there are persuasive theoretical arguments for a key role of the fiscal policies on the economic growth (see Barro, 1990; King and Rebelo, 1990; Jones *et al.*, 1993) the empirical evidence with many empirical studies have analyzed the link between these variables yielding inconclusive results (see Arnold, 2008).

[*]Corresponding author. s.disanzo@confcommercio.it,
[†]m.bella@confcommercio.it
[‡]g.graziano@confcommercio.it

These ones range from the findings of a weak and non-robust, or only mild, significant relationship between fiscal policies and growth (see Easterly and Rebelo, 1993; Mendoza *et al.*, 1997; Agell *et al.*, 2006) to the evidence of a robust and significant negative relationship (see Barro, 1989, 1991; and Folster and Henrekson, 2001).

Recently, the economic research on this subject has focused on the link between long-run growth and tax structure under the government's revenue neutrality condition (see Arnold, 2008, for a survey)[1]. This topic has taken on added importance at present time, since in the wake of the recent financial and economic crisis, many countries must face the challenge of restoring public finances without jeopardising economic growth.

This means that given a number of fiscal instruments, i.e. taxes, and a given target of total revenue it is strictly preferable to choose the tax structure that minimize the effect on long-run growth.

Along this intuition, for example, Arnold *et al.* (2011) group taxes into four major categories: corporate income taxes, personal income taxes, consumption taxes and property taxes. The authors employ an error correction growth equation with short-term dynamics modeled explicitly to evaluate the impact of revenue-neutral tax structure changes on the long-run level of GDP per capita. Their parameter estimates, which are based on the Pooled Mean Group (PMG) estimator proposed by Pesaran *et al.* (1999), suggest the following rank of tax instruments in terms of effects on GDP per capita: property taxes are the least harmful for the growth, followed by consumption taxes, personal income taxes and corporate income taxes.

These ideas and these findings had a significant impact on the policy debate in Europe, and are at root of the recent OECD (2008, 2010) and European Commission (2013) policy recommendations that suggest to implement growth-oriented tax structure reforms in the direction of a tax shift towards consumption and property taxes.

Recently, some authors cast doubt on the robustness of the empirical results obtained by Arnold *et al.* (2011).

Xing (2012) carries out several robustness checks as different choices of sample and variables, time effects or period dummies and finds that the only robust result appears to be that shifts in tax revenues towards

[1]This condition implies that any change in a tax structure indicator is compensated by an equivalent change in another indicator to keep the total tax revenue as a share of GDP unchanged.

It is worth noting that there are also empirical studies that in the analysis of the growth effects of the tax structure do not adopt the hypothesis of tax revenue neutrality (see Kneller *et al.*,1999; Lee and Gordon, 2005). In this case, the influence of the tax structure on the economic growth becomes less evident.

property taxes are associated with higher level of GDP per capita in the long-run. Moreover, the author points out that the Arnold *et al.* (2011) results may be biased since the homogeneity coefficient hypothesis imposed in PMG estimations seem to be not valid when performing a Wald test of this restriction on long-run coefficients.

Arachi *et al.* (2015) study the relationship between growth and tax structure using tax ratios and implicit tax rates as indicators and they found that the link between tax structure and long-run GDP per capita is not robust to the adoption of different assumption on observable and unobservable heterogeneity across countries.

In our view, there are two relevant shortcomings that plague the existing literature in this topic.

First, it is based on a single-equation framework, then is assumed that the right-side variables are exogenously determined. In empirical applications, this hypothesis could be violated as tax revenues increase in expansion and decline in recessions (business cycle effects), thus the variables included in the growth model interact and feedback on each other when there are changes in the economy. Overlooking these feedbacks could lead to biased estimates and therefore inference results may be misleading.

Second, it uses linear models in the analysis, thus it is implicitly assumed that the effects of taxes on growth are symmetric. As suggested by some authors, this assumption may be too restrictive since the effects of taxes on growth are asymmetric.[2]

For example, Perotti (1999) shows that in times of fiscal stress shocks to government revenues have very different effects on private consumption than in normal times.

Recently, Jaimovich and Rebelo (2015) propose a theoretical model where the effects of taxation on growth are highly non-linear and show that marginal increases in tax rates have small growth impact when tax rates are low or moderate and large growth impact when tax rates are high. These theoretical results suggest that any satisfactory empirical growth model must take into account the asymmetric effects of tax policies on growth, because missing non-linearities may yield seriously biased estimates.

Taking into account these ideas and the econometric drawbacks in the existing literature, our paper proceeds in the following ways.

First, we study the link between long-run growth and tax structure using a linear and non-linear panel vector error correction model (PVECM) (see Engle and Granger, 1987). The analysis based on a

[2] For a more comprehensive reading, we remind to some empirical studies regarding fiscal policy asymmetries (Auerbach *et al.*, 2012, and Baum *et al.*, 2012)

PVECM is attractive since it allows us to avoid the simultaneity bias problems. Indeed, it explicitly considers the feedbacks across the variables included in the study allowing to estimate the effect of tax structure on growth as well as the reverse effect of growth on tax structure.

Moreover, in the PVECM it is also possible to verify if it is reasonable, as supposed by the previous research, to employ a single-equation approach in the empirical analysis. To test the single-equation hypothesis, we use the weak exogeneity tests proposed by Johansen (1988, 1991, 1995).

Second, we extend to the panel framework the threshold regression methodology proposed by Hansen (2002) and employ it to evaluate the asymmetric impact of taxation on growth. In particular, we are interested to investigate if there exists a total tax threshold (or "tipping point") value above which tax changes have a larger effect on long-run GDP per capita. Our study focuses in a sample of 20 OECD countries and covers the period 1970-2012.

The paper proceeds as follows. Section 2 presents the model specifications and the estimation methodology. Section 3 presents the empirical results. Section 4 concludes.

2 Model specifications

2.1 Linear model

We consider the following heterogenous p-variate $k_i th$ order panel VAR model:

$$y_{it} = \mu_i + \sum_{m=1}^{k_i} A_{im} y_{i,t-m} + \varepsilon_{it}, \quad (i = 1, ..., N;\ t = 1, ..., T) \qquad (1)$$

where $y_{i,t-m}, (m = 0, ..., k_i)$, the fixed-effect μ_i and the error term ε_{it} are $p \times 1$ vectors and A_{im},$(m = 1, ..., k_i)$, is an $p \times p$ coefficient matrix. The subscripts i denotes the countries and the subscripts t denotes the time. The errors ε_{it} are independent and identically distributed $\varepsilon_{it} \sim N_p(0, \Omega_i)$. Moreover, we adopt the standard panel data assumption of independence, namely $E(\varepsilon_{it}, \varepsilon_{jt}) = 0$ for all $i \neq j$, thus the processes ε_{it} are independent cross sectionally (see Pedroni, 2000)[3].

The error correction representation for the model (1) is the following (see Engle and Granger, 1987; Johansen, 1995):

[3] In the empirical analysis this hypothesis may be stretch. To control for common disturbances in the panel, we employed the strategy proposed by Pedroni (2000, 2004, 2007).

$$\Delta y_{it} = \mu_i + \Pi_i y_{i,t-1} + \sum_{m=1}^{k_i-1} \Gamma_{im} \Delta y_{i,t-m} + \varepsilon_{it}, \tag{2}$$

where Π_i is a $p \times p$ matrix. If the matrix Π_i is of reduced rank, then it can be expressed as $\Pi_i = \alpha_i \beta_i^{'}$ where α_i and β_i are full-column rank matrices of order $p \times r$, and r is the cointegrating rank. β_i is the matrix of cointegrating vectors (or long-run parameters) and the matrix α_i represents the speed of adjustment to the long-run equilibrium.

The multivariate cointegration method proposed by Johansen (1988, 1995) is based on the error correction model (2) and was developed for time-series data.

The author propose a Likelihood-based methodology that allows to estimate Π_i and make hypothesis testing on the cointegrating rank, the long-run coefficients β_i and the adjustment parameters α_i. Under the hypothesis of cross sectional independence, a panel estimator of the parameters of model (2) can be obtained using the Mean Group (MG) estimator proposed by Pesaran and Smith (1995) that allows both long-run and short-run coefficients to be country-specific. More specifically, one can estimate for country i, using the Johansen method, the parameters of the model (2) and then to average the coefficients over N[4]. This strategy provides consistent estimates of mean effects.

2.2 Asymmetric model

In order to account for long-run asymmetric effects of tax policy on output, we introduce threshold type non-linearities in model (2). We consider the following threshold cointegration model[5]:

$$\Delta y_{it} = \mu_i + \alpha_i \beta_{i1}^{'} y_{i,t-1} I(q_{it} \leq \gamma_i) + \alpha_i \beta_{i2}^{'} y_{i,t-1} I(q_{it} > \gamma_i) + \sum_{m=1}^{k_i-1} \Gamma_{im} \Delta y_{i,t-m} + \varepsilon_{it}, \tag{3}$$

where $I(\gamma_i)$ is an indicator function taking value 1 if the values of the threshold series q_{it} is below a specific threshold value γ_i. We choose the total tax revenue as a share of GDP as our threshold variable q_{it}. In model (3) the observations are divided in two regimes and the threshold parameter γ_i determines whether an observation belongs to one regime or to the other one. The regimes are distinguished by differing cointegrating vectors, β_{i1} and β_{i2}. Thus, tax policies have a different impact on growth,

[4]The MG estimator may be sensitive to the presence of country-specific estimates with extreme values. To avoid this problem, we use the strategy proposed by Bond *et al.* (2010) that allows to obtain an outlier-robust MG estimator.

[5]For an extensive discussion of threshold regression models we refer to Tong (1990).

which depends on the regime of the economy. We estimate the model (3) adapting the threshold cointegration methodology proposed by Hansen (2002) to the panel framework. Basically the methodology involves three steps.

First, we estimate the model (3) for each value γ_i on specified subset Φ_i of the threshold series q_{it} and we retain the sum of squared residuals. In a second step, we select the value of $\widehat{\gamma}_i$ which minimizes the sum of squared residuals. In a third step, the panel estimates are obtained using the MG estimator proposed by Pesaran and Smith (1995).

We test if the threshold effect is statistically significant, namely the null hypothesis of a linear model versus the alternative of a nonlinear one. The null hypothesis of no threshold effect in (3) is defined by the linear constraint $H_0 : \beta_{i1} = \beta_{i2}$.

Under the null hypothesis H_0 the threshold parameter γ_i is not identified. As a result, the asymptotic distribution of conventional test statistics is not χ^2. This is a well-known problem in the literature on testing for regime switching type of non-linearities. This problem is usually handled by viewing the test statistic as a random function of the nuisance parameters and basing inference on a particular functional of the test statistic such as, for instance, its supremum over γ_i (see, Davies 1977 and 1987, Andrews and Ploberger 1994, Chan 1990, and Hansen 1996). Letting $LR_i(\gamma_i)$ the likelihood-ratio type test statistic obtained for each γ_i, we base our inferences on $SupLR_i = \sup_{\gamma_i \in \Phi_i} LR_i(\gamma_i)$. We use the bootstrap methods proposed by Hansen (2002) to approximate the sampling distribution of $SupLR_i$ under H_0[6].

The bootstrap method proposed by Hansen (2002) was developed in a time series context. To extend this methodology to the panel framework, we use a Fisher-type test which combines the p-values from N independent tests (see Maddala and Wu, 1999, for more details).

Let p_i be the p-value of the bootstrap linearity test for cross-section i, the Fisher test statistic is the following :

$$p_\lambda = -2\sum_{i=1}^{N} \ln(p_i). \tag{4}$$

As p_i $(1, 2, ..., N)$ are independent uniform $(0, 1)$ variables, then $-2\ln(p_i)$ has a χ^2 distribution with two degrees of freedom. Using the additive property of the χ^2 variables, we obtain that p_λ has a χ^2 distribution with $2N$ degrees of freedom .

[6] See Di Sanzo and Perez-Alonso (2011) for details about Monte Carlo and bootstrap simulations in threshold models.

3 Empirical results

3.1 Data and preliminary analysis

Our study concerns 20 OECD countries, namely, Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, United Kingdom and United States. We consider in the analysis the following non-fiscal variables suggested by the empirical literature: the GDP per capita, the population and the investment in physical capital[7]. The fiscal variables are the total tax revenue, the income taxes, the consumption taxes and the property taxes. The total tax revenue is expressed as a share of GDP while the other tax variables are expressed as a share of total tax revenue. Data comes from OECD database and covers the period from 1970 to 2012[8]. Table 1 lays out some indications about the total tax evolution and the tax composition in the sample. In the long-term the compound annual growth rate (CAGR) displays different patterns across countries (column 3). In most of these, it has been positive, while in some others (Ireland and Canada) it has been much steadier and, finally, in the United Kingdom and US it has followed a negative dynamics. Moreover, the total tax level (column 2) differs substantially across countries ranging from the 47.2% of Sweden to the 25.5% of the US. Finally, even if the countries have employed different types of taxes to generate revenues, the tax structure is biased towards income taxation (columns 4-6). Indeed, the values of income taxes range from the 72.8% of the Japan to the 51.4% of the Greece.

To identify the degree of integration of the variables involved in our study, we have implemented the panel unit root tests proposed by Im *et al.* (IPS) (2003) and Hadri (2000). Table 2 summarizes the unit root tests results. For all variables the IPS test fails to reject the null hypothesis of non-stationarity and the Hadri test rejects the null of stationarity.

To check if the results on the degree of integration are robust to the presence of structural breaks in the data generating processes of the series, we also have performed the panel unit root test with structural breaks proposed by Im *et al.* (2005). The test results are reported in Table 3 and provide strong evidence in favor of non-stationarity.

[7] We also checked if the empirical results depend from the choice of non-fiscal growth determinants. Using different sets of non-fiscal variables (such as human capital, trade openness indicators etc.) we found similar estimates, thus we settle on this parsimonious specification of the model.

[8] Country and estimation period is dictated by data availability only and more precisely, for Australia, Greece, Japan, Netherlands and Switzerland the full sample period is 1970-2011, while for Portugal the sample reduced to 1989-2012.

3.2 Estimation results

We study the growth effects of tax composition under the assumption of the revenue neutrality. Accordingly, following Arnold *et al.* (2011), we take the total tax revenue as a control variable and drop one tax instrument at time in the regression. Therefore, the changes of the tax structure indicators included in the regression are offset by changes in the tax instrument omitted to keep the overall tax burden unchanged. In other words, the estimated coefficients of the fiscal variables measure the growth effects of a shift from the tax instrument omitted to the others included in the regression. For example, we omit the variable consumption taxes from the regression when we analyze the impact of the income taxes and property taxes on the growth. In this case, it is supposed that the changes in income and property taxes are compensated by changes in consumption in opposite direction to keep the overall tax burden unchanged.

Unit root tests have provided evidence of breaks in the series thus, to verify whether there are long-run relationships between variables, we tested for cointegration using the Johansen *et al.* (JMN) (2000) method. JMN (2000) developed a time series cointegration test in the presence of structural breaks at known points in time. Their approach is a slight generalization of the likelihood-based cointegration analysis proposed by Johansen (1988, 1995) and it is robust under the presence of structural breaks. We extend this methodology to the panel framework, using a Fisher-type test.

The results for all models considered are reported in Table 4. The cointegration test suggests the presence of one relation of cointegration among the variables.

Table 5 summarizes the estimates of the cointegrating vector for the liner model. Column (1) reports the effects of a hypothetical shift towards income taxes and property taxes financed by a reduction of consumption taxes. Therefore, we have omitted the consumption tax variable from the regression as indicated at the bottom of the table. The estimated coefficient on income taxes suggests negative and statistically significant effects of these taxes on the the long-run level of GDP per capita. On the other hand, the shift towards property taxes has a positive effects on the long-run level of GDP per capita. Column (2) considers a similar exercise where is evaluated the impact on the long-run GDP per capita of a hypothetical shift from income taxes to consumption and property taxes. As in column (1), the property taxes have a positive effect on long-run GDP per capita level. Consumption taxes have a negative effect, but the GDP per capita reduction is stronger of that obtained with the income taxes in column (1). In column (3) we consider a shift from property taxes towards income and consumption taxes. The coefficient estimates confirm the results obtained in columns (1)-(2): income and consumption taxes lead to a reduction of the long-run level of the GDP per capita and consumption taxes have the most negative effects on economic performance. Finally, it is important to note that for all models we find that the coefficient on total tax revenue is negative and statistically significant.

To resume, in contrast to the existing literature (Arnold *et al.*, 2011; Xing, 2012), we find that consumption taxes are associated with lower long-run level of GDP per capita than taxes on income, while, in according with it, our results suggest that shifts in total tax revenue towards property taxes may be associated with a higher long-run level of GDP per capita.

3.3 Endogeneity test

As we pointed out in the introduction, the results obtained by previous research may be biased since were obtained using a single-equation model. In this framework is likely to have problems of endogeneity bias since the fiscal variables are correlated with the business cycle. Moreover, when the variables are not weakly exogenous the single-equation approach will lead to long-run estimator inefficiency (Juselius, 2006).

Despite the fact that the previous literature have identified this important econometric problem, not much has been done to properly account for it. Arnold *et al.* (2011) control for endogeneity purging the variables of all possible correlations with business cycle, by regressing them on the output gap in a first stage and using the residuals of these regressions instead of the tax variables in the growth regressions. This strategy is frequently used in empirical analysis, but is imperfect, because some endogeneity may remain. Arachi *et al.* (2015) use in the analysis time series not sufficiently long, then they cannot implement the strategy proposed by Chudik and Pesaran (2015) to control for endogeneity. This strategy implies to include in the model lags of all cross-sectional averages of the dependent and of the control variables, as well as a sufficient number of cross-sectional averages of one or more additional covariates. Accordingly, to verify the reasonability of the single-equation hypothesis, we have tested for weak exogeneity using the procedure described in Johansen (1988, 1995). Basically, a variable is said to be long-run weakly exogenous when the cointegration vector do not have any influence on its dynamics, thus the null hypothesis of the weak exogeneity test is $\alpha = 0$. Table 6 shows the test results for separate and joint restrictions on $\alpha's$, giving evidence that weak exogeneity cannot be assumed. Thus, empirical results obtained by previous literature in a single-equation framework must be interpreted with caution since the estimates obtained may be biased and the inference may be misleading.

3.4 Checking for a threshold

In this section we investigate if there is a total tax threshold (or "tipping point") value above which tax changes have a larger effect on the long-run level of GDP per capita.

The parameter estimates for the non-linear model are reported in Table 7. For all models considered, the linearity test, i.e. $H_0 : \beta_1 = \beta_2$, indicates that the assumption of linearity can be rejected, therefore it is not appropriated to employ a linear model to study the relationship between tax structure and economic growth.

Regardless of the model specification, we find a tax threshold value of around 30%, which is significant for all models considered. Moreover, the link between tax policies and the long-run level of GDP per capita is robust and significant only in the unfavorable regime, namely, when the total tax revenue as a share of GDP is above the threshold value of 30%. This is a very strong result and it is consistent with the theoretical predictions reported in Jaimovich and Rebelo (2015): fiscal policies based on tax shifts are effective only when a total tax threshold is exceeded, otherwise the effects on GDP per capita disappear and become not significant.

The magnitude of the estimated coefficients differs substantially between linear and non-linear model. With respect to the linear model, in the non-linear one the impact of tax structure on GDP per capita is smaller, while the effect of total tax revenue is larger. Moreover, it is important to note that in non-linear model the consumption and income taxes have comparable effects. Therefore, the evidence favouring consumption taxes over income taxes is weaker when we introduce non-linearities in the analysis. To conclude, since there are considerable differences in the results obtained with linear and nonlinear models, in empirical analysis one has to take caution to assume the hypothesis of a linear model when it is present a nonlinear relationship between the series under study, since it could lead to misleading conclusions about the relationship between tax policies and economic growth.

4 Conclusions

In this paper we used a linear and non-linear panel cointegrated VAR model to study the link between tax structure and long-run growth. We have focused on a panel of twenty OECD countries during the period 1970- 2012. To evaluate the asymmetric impact of tax changes on growth, we extend the threshold cointegration methodology proposed by Hansen (2002) to the panel framework. To test for linearity, a Fisher-type version of the Hansen (2002) time-series linearity test is proposed.

The tests indicate that the weak exogeneity and linearity hypotheses are not supported by data, thus the empirical results obtained by previous literature in a linear single-equation framework must be interpreted with caution since the estimates obtained may be biased and the inference may be misleading.

As in the recent literature results (Arnold *et al.*, 2011; Xing, 2012), we find that recurrent taxes on immovable property seem to be the least harmful for the growth, while, in contrast with it, we do not found compounding evidence favouring consumption taxes over income taxes. In particular, our results suggest income and consumption taxes as the most harmful for the growth and having comparable effects. Therefore, our results do not fully support the policy prescriptions proposed by many international organizations (such as OECD, International Monetary Funds and European Commission) aimed at shifting the tax burden from income taxes to consumption and property taxes.

Moreover, these results are robust and significant when a total tax threshold of 30% is exceeded, consistently with the theoretical predictions reported in Jaimovich and Rebelo (2015) that found a negative non-linear relationship between taxes and economic growth that are effective only when a total tax threshold is exceeded. Otherwise the effects of tax policies on GDP per capita disappear and become not statistically significant.

Our empirical results are explained in a theoretical terms, as follow (see Johansson *et al.*, 2008, and Arnold *et al.*, 2011, for more details).

- Most of the OECD countries give various tax preferences for owner-occupied housing (such as tax deductibility of interest on house loans and exemption from capital gains tax), which produce misallocation of capital from other investement to housing. Recurrent taxes on immovable property are the least harmful for the growth because increasing in this type of taxes will shift some investments from housing toward other type of productive investments and so, *ceteris paribus*, increase the growth.

- Consumption taxes appear to be harmful for the growth not less than income taxes, because it is almost universally agreed that these type of taxes are regressive. This because these taxes fall on the consumption of goods and services that make up a larger share of the budgets of poorer than richer households. Therefore, the increase in these taxes may lead to a higher level of inequality and thus to a lower long-run level of GDP per capita. Probably, this is one of the channels trough which operates the endogeneity between growth and taxes and explains the rejection of the weak exogeneity hypothesis.

Two important policy implications emerge from our analysis.

First, since property taxes appear to be the least harmful for the growth, tax reforms aimed to replace revenues from income and consumption taxes with property taxes would increase long-run GDP per capita.

Second, our results suggest that the negative effects of total taxes on growth become strongly and significant when the tax burden is above the threshold value of 30%. Thus, especially for countries with a high tax burden, a mix of policies directed to reduce the total tax revenue and to promote the tax structure least harmful for the growth are more appropriate to guarantee growth and fiscal sustainability.

Two important caveats apply to our analysis shared with other studies in this topic. First, our results are based on the Mean Group (MG) estimator proposed by Pesaran and Smith (1995), thus they indicate, for the countries included in the study, the average growth effects of the tax policies. Therefore, they cannot be employed to make predictions about the suitability of a generalized tax reform for a particular country in the sample.

Second, we have evaluated the tax policies in terms of their effects on the GDP per capita. However, to evaluate the goodness of a given reform, the governments have to take into account others factors, such as inequality concerns, simplicity and compliance costs. There exists a large literature on the effects of tax structure on efficiency, income distribution etc., which are discussed in Institute for Fiscal Studies (2010).

References

[1] Agell, J., H. Ohlsson, and P. Thoursie, 2006. Growth Effects of Government Expenditure and Taxation in Rich Countries: A Comment. *European Economic Review* 50, 211-218

[2] Andrews, D.W.K., and W. Ploberger, 1994. Optimal tests when a nuisance parameter is present only under the alternative," Econometrica, 62, 1383-1414.

[3] Arachi, G., V. Bucci, and A. Casarico, 2015. Tax structure and macroeconomic performance. *International Tax and Public Finance* 22, 635–662

[4] Arnold, Jens, 2008. Do tax structures affect aggregate economic growth? Empirical evidence from a panel of OECD countries. *OECD Working Paper no.* 643.

[5] Arnold, J., B. Brys, C. Heady, A. Johansson, C. Schwellnus, and L. Vartia, 2011. Tax policy for economic recovery and growth. *The Economic Journal* 121, 59–80.

[6] Auerbach, A. J., Gorodnichenko Y., 2012. Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy* 2012, 4(2): 1-27.

[7] Barro, R. J., 1989. Economic growth in a cross section of countries. *NBER Working Paper no.* 3120.

[8] Barro, R. J., 1990. Government spending in a simple model of endogenous growth. *Journal of Political Economy* 98, 103–117.

[9] Barro, R. J. 1991. Economic growth in a cross section of countries. *The Quarterly Journal of Economics* 106, 407-443.

[10] Baum A., Poplawski-Ribeiro M., Weber A., 2012. Fiscal Multipliers and the state of the economy. *IMF Working Paper no.* 286.

[11] Bond, S., A. Leblebicioglu, and F. Schiantarelli, 2010. Capital accumulation and growth: a new look at the empirical evidence. *Journal of Applied Econometrics* 25, 1073-1099.

[12.] Chan, K.S., 1990. Testing for threshold autoregression. *The Annals of Statistics* 18, 1886-1894.

[13] Chudik, A., and Pesaran, M. H., 2015. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics* 188, 393–420

[14] Davies, R.B., 1977. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 64, 247-254.

[15] Davies, R.B., 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 74, 33-43.

[16] Di Sanzo, S., and A. Pérez-Alonso, 2011. Unemployment and Hysteresis: A Nonlinear Unobserved Components Approach. *Studies*

in Nonlinear Dynamics & Econometrics 15, 24-50.
[17] Easterly, W., and S. Rebelo, 1993. Fiscal Policy and Economic Growth: An Empirical Investigation. *Journal of Monetary Economics* 32, 417-458.
[18] Engle, R. F., and C. W. J. Granger, 1987. Cointegration and error correction: Representation, estimation and testing. *Econometrica* 55, 251–276.
[19] European Commission (2013). Taxation trends in the European Union. *Publications Office of the European Union.*
[20] Fölster, S., and M. Henrekson, 2001. Growth effects of government expenditure and taxation in rich countries. *European Economic Review* 45, 1501-1520.
[21] Hadri, K., 2000. Testing for stationarity in heterogeneous panel data. *Econometrics Journal* 3, 148-161.
[22] Hansen, B.E.,1996. Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* 64, 413-30.
[23] Hansen, B.E., and S. Byeongseon, 2002. Testing for two-regime threshold cointegration in vector error correction models. *Journal of Econometrics* 110, 293-318.
[24] Im, K. S., M. H. Pesaran, and Y. Shin, 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115, 53–74
[25] Im, K. S., J. Lee, and M. Tieslau, 2005. Panel LM unit-root tests with level shifts. *Oxford Bulletin of Economics and Statistics* 67, 393–419.
[26] Institute for Fiscal Studies, 2010. Dimension of Tax Design: The Mirless Review. *Oxford: Oxford University Press.*
[27] Jaimovich, N., and S. Rebelo, 2015. Non-linear Effects of Taxation on Growth, forthcoming *Journal of Political Economy.*
[28] Johansen, S., 1988. Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control* 12, 231–54.
[29] Johansen, S., 1991. "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59, 1551–80.
[30] Johansen, S.,1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. *Oxford: Oxford University Press.*
[31] Johansen S., R. Mosconi, and B. Nelsen, 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. *Econometrics Journal* 3, 216-249.
[32] Jones, L., R. Manuelli, and P. Rossi, 1993. Optimal taxation in models of endogenous growth. *Journal of Political Economy* 101, 485–519.
[33] Juselius, K., 2006. The Cointegrated VAR Model: Methodology

and Applications. *Oxford: Oxford University Press.*
[34] Lee, Y., and R. H. Gordon, 2005. Tax structure and economic growth. *Journal of Public Economics* 89, 1027–1043.
[35] King, R., and S. Rebelo, 1990. Public Policy and Economic Growth: Developing Neoclassical Implications. *Journal of Political Economy* 98, S126–S150.
[36] Kneller, R., M. F. Bleaney, and N. Gemmell, 1999. Fiscal policy and growth: evidence from OECD countries. *Journal of Public Economics* 74,171–190
[37] Maddala, G. S., and S. Wu, 1999. A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics* 61, 631–652.
[38] Mendoza, E., G. Milesi-Ferretti, and P. Asea, 1997. On the effectiveness of tax policy in altering long-run growth: Harberger's superneutrality conjecture. *Journal of Public Economics* 66, 99–126.
[39] OECD. (2008). Tax and economic growth. *Economics Department Working Paper No.* 620.
[40] OECD. (2010). Going for growth 2010. *Paris: OECD Publishing.*
[41] Pedroni, P., 2000. Fully-modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity. *Advances in Econometrics* 15, 93–130.
[42] Pedroni, P., 2004. Panel cointegration; asymptotic and finite sample properties of pooled time series tests, with an application to the PPP hypothesis. *Econometric Theory* 20, 575–625.
[43] Pedroni, P., 2007. Social capital, barriers to production and capital shares: implications for the importance of parameter heterogeneity from a nonstationary panel approach. *Journal of Applied Econometrics* 22, 429–451.
[44] Perotti, R., 1999. Fiscal Policy in Good Times and Bad. *Quarterly Journal of Economics* 114, 1399-1436.
[45] Pesaran, M. H, and R.P. Smith, 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics* 68, 79-113.
[46] Pesaran, M.H., Y. Shin, and R. Smith, 1999. Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association* 94, 621–634.
[47] Tong, H., 1990. Nonlinear Time Series: a Dynamical System Approach. *Oxford: Oxford University Press.*
[48] Xing, J., 2012. Tax structure and growth: How robust is the empirical evidence? *Economics Letters* 117, 379–382.

1 Tables

In this section there are the tables in the paper.

Table 1: Revenue shares of total tax and its composition.

Country	*CAGR* GDP p.c.	Total Tax	*CAGR* Total Tax	Income Taxes	Property Taxes	Consumption Taxes
Australia	1.5	27.2	0.6	61.8	8.9	29.4
Austria	2.0	40.6	0.6	65.2	2.1	30.8
Belgium	1.7	42.6	0.7	69.1	4.4	26.2
Canada	1.4	32.7	0.0	59.8	10.3	28.2
Denmark	1.5	46.1	0.5	59.6	4.5	34.0
Finland	2.2	41.5	0.8	64.7	2.4	32.7
France	1.6	41.5	0.7	61.6	6.3	28.6
Germany	1.8	36.0	0.4	69.1	3.0	27.7
Greece	1.3	27.3	1.2	51.4	5.7	41.5
Ireland	2.9	31.1	0.0	52.0	6.3	41.6
Italy	1.6	36.4	1.3	66.2	4.3	27.6
Japan	2.3	25.8	1.0	72.8	9.9	17.1
Netherlands	1.6	40.9	0.2	68.0	3.9	27.4
New Zealand	1.1	32.3	0.6	59.7	6.8	30.7
Norway	2.3	41.6	0.5	62.7	2.4	34.7
Portugal	2.0	30.4	1.4	53.7	3.3	41.7
Spain	1.7	28.9	1.7	65.8	6.1	27.2
Sweden	1.7	47.2	0.4	71.2	2.5	26.1
Switzerland	0.9	26.0	0.9	67.1	8.4	22.3
United Kingdom	1.8	34.9	-0.1	57.3	11.7	30.6
United States	1.6	25.5	-0.1	70.0	11.9	18.1

Note: the table gives for columns 2, 4, 5 and 6 the country averages from 1970 to 2012 while for the first and third column gives the Compound Annual Growth Rate (CAGR) of GDP per capita and Total Tax during the sample period. Total Tax is expressed as percentage of GDP, while Income, Property and Goods and Services Taxes are expressed as percentage of Total Taxes.

Table 2: Panel unit root tests

Variable	IPS test	Hadry test
GDP per capita	0.636	8.101***
Population	-0.480	10.332***
Investment	-0.209	9.784***
Total Tax	1.323	11.002***
Income Taxes	-0.541	14.722***
Consumption Taxes	0.250	12.575***
Property Taxes	-0.802	10.204***

Note: The test statistics are asymptotically normally distributed.
*** significant at 1%. The null hypothesis of IPS test is: unit root.
The null hypothesis of Hadry test is: stationarity.

Table 3: Panel unit root test with structural breaks (Im *et al.* (2005))

Variable	one break	two breaks
GDP per capita	0.193	0.740
Population	-0.365	-0.780
Investment	-0.464	-0.662
Total Tax	1.011	1.311
Income Taxes	-0.390	-0.416
Consumption Taxes	0.555	0.693
Property Taxes	-0.320	-0.487

Note: The test statistic are asymptotically normally distributed.

Table 4: Cointegration test with structural breaks (JMN, 2000)

Cointegrating rank (omitted tax variable: Consumption Taxes)	Test statistic
r=0	286.882***
r≤ 1	39.861
r≤ 2	15.774
Cointegrating rank (omitted tax variable: Income Taxes)	Test statistic
r=0	185.221***
r≤ 1	28.464
r≤ 2	11.881
Cointegrating rank (omitted tax variable: Property Taxes)	Test statistic
r=0	158.477***
r≤ 1	29.022
r≤ 2	13.659

Note: The test has a chi-square distribution with 2N degree of freedom.
** significant at 1%.

Table 5: Estimation results of cointegrating vector (linear model)

GDP per capita	(1)	(2)	(3)
Population	-0.272**	-0.380**	-0.490***
	(36.332)	(36.524)	(38.887)
Investment	0.621**	0.551***	0.750**
	(35.913)	(38.082)	(35.725)
Total Tax	-0.513***	-0.340***	-0.421***
	(43.296)	(45.994)	(46.011)
Income Taxes	-1.852***		-0.791 ***
	(39.031)		(42.073)
Consumption Taxes		-2.464***	-1.754***
		(52.455)	(47.335)
Property Taxes	0.738**	1.292***	
	(36.771)	(45.537)	
Omitted tax variable	Consumption Taxes	Income Taxes	Property Taxes

Note: Likelihood ratio (LR) test statistic in parenthesis.
The LR statistic has a chi-square distribution with N degrees of freedom.
/ significant at 1% and 5%.

Table 6: Test statistics for weak exogeneity (Johansen, 1995)

Restricted α	(1)	(2)	(3)
Total Tax	35.810**	32.229**	38.031***
Income Taxes	36.286**		41.123***
Consumption Taxes		45.443***	42.545***
Property Taxes	34.885**	33.237**	
Joint	113.322***	104.731***	115.449***
Omitted tax variable	Consumption Taxes	Income Taxes	Property Taxes

Note: The individual test is asymptotically chi-squared distributed with N degrees of freedom.
The joint test is asymptotically chi-square distributed with 3N degrees of freedom.
/ significant at 1% and 5%.

Table 7: Estimation results of cointegrating vector (non-linear model)

GDP per capita	(1)		(2)		(3)	
Regime	$q_{t-1} \leq \gamma$	$q_{t-1} > \gamma$	$q_{t-1} \leq \gamma$	$q_{t-1} > \gamma$	$q_{t-1} \leq \gamma$	$q_{t-1} > \gamma$
Total Tax	-0.074	-0.823***	-0.177	-0.521***	-0.056	-0.713**
	(16.425)	(40.643)	(22.018)	(42.752)	(18.710)	(36.986)
Income Taxes	-0.179	-1.035***			-0.263	-0.513***
	(12.324)	(41.109)			(20.229)	(45.079)
Consumption Taxes			-0.184*	-1.331***	-0.379*	-0.679***
			(28.925)	(55.630)	(29.521)	(49.661)
Property Taxes	0.236	0.498**	0.157	0.748***		
	(15.803)	(35.514)	(21.895)	(42.509)		
Population	-0.353**		-0.445***		-0.585***	
	(35.770)		(38.452)		(37.992)	
Investment	0.702***		0.623***		0.831***	
	(42.003)		(45.381)		(39.766)	
Treshold Estimate	30.223***		30.552**		30.223***	
	(64.821)		(66.081)		(63.772)	
Linearity Test Statistic	71.31***		59.88**		69.91***	
Omitted tax variable	Consumption Taxes		Income Taxes		Property Taxes	

Note: Likelihood ratio (LR) test statistic in parenthesis. The LR statistic has a chi-square distribution with N degrees of freedom. The linearity test statistic has a a chi-square distribution with 2N degrees of freedom.
***/**/* significant at 1%, 5% and 10% respectively.