ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Miniaci, Raffaele; Panteghini, Paolo; Rivolta, Giulia

Working Paper

The Estimation of Reaction Functions under Tax Competition

CESifo Working Paper, No. 6928

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Miniaci, Raffaele; Panteghini, Paolo; Rivolta, Giulia (2018) : The Estimation of Reaction Functions under Tax Competition, CESifo Working Paper, No. 6928, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/176947

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

CESifo WORKING PAPERS

6928
2018

February 2018

The Estimation of Reaction Functions under Tax Competition

Raffaele Miniaci, Paolo M. Panteghini, Giulia Rivolta

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6928*
Category 1: Public Finance

The Estimation of Reaction Functions under Tax Competition

Abstract

Tax competition has long been studied using panel models. According to this approach, each country's tax rate is assumed to depend on a weighted average of the tax rates applied in the rest of the world, where weights are exogenous. As a consequence, the estimated reaction functions of countries throughout the world have the same sign. This means that all tax rates are either strategic complements or strategic substitutes. Moreover, the intensity of a country's reaction depends on certain exogenous weights, with a unique proportional factor common across all countries. Our article departs from this standard approach and proposes a VAR model as an alternative estimation strategy. Accordingly, weights are no longer determined exogenously but rather endogenously. As such, we compare and explore the implications of the panel versus the VAR model in terms of structural contemporaneous parameters and impulse response functions. We show that results obtained with a VAR model differ from those obtained from a panel approach. In particular, we find that strategic complementarity between certain countries (with a positive slope of reaction functions) may co-exist with strategic substitutability between other countries (negative slope). Given these results, we can say that a standard panel approach is relatively restrictive and therefore can lead to unreliable estimates, and fail to provide helpful policy recommendations.

JEL-Codes: C540, H250, E620.

Keywords: tax competition, VAR models, Bayesian methods.

Raffaele Miniaci
Department of Economics
University of Brescia / Italy
raffaele.miniaci@unibs.it

Paolo M. Panteghini
Department of Economics
University of Brescia / Italy
panteghi@eco.unibs.it

Giulia Rivolta
Department of Economics
University of Brescia / Italy
giulia.rivolta@unibs.it

We wish to thank Wiji Arulampalam, Efrem Castelnuovo, Federico Revelli and the participants to the 8th RCEA Macro-Money-Finance Workshop for useful suggestions and comments.

1 Introduction

There has been a wave of market openness since the mid and late 1980s. This phenomenon has both stimulated the mobility of capital and fostered multinational activity. According to many scholars, this market openness has led to international tax competition. Subsequently, a growing amount of literature has looked at tax competition from both a theoretical and empirical point of view.

In empirical work, tax competition is typically analysed by means of a panel model approach. Thus, certain coefficients are assumed to be common across equations. This condition is useful to exploit both the cross-sectional and the time-series dimension of the data. With this approach, the dependent variable in each equation is the tax rate of one jurisdiction (which is its strategic tool), while the other countries' tax rates are independent variables, used to estimate the effects of competition. Moreover, panel models implicitly assume that strategic interactions between countries are of the same nature. As will be seen, this assumption has a relevant drawback in that it disregards the possibility that tax competition operates asymmetrically.

In order to study tax competition, it is necessary to identify an exogenous variation of the fiscal variable. This is difficult because under tax competition, countries probably adjust their tax rates simultaneously. This endogeneity problem can be tackled by using instrumental variable estimates. Moreover, it is usually assumed that each country reacts to a weighted average of other countries' tax rates. This approach therefore lies on two fairly strong assumptions: firstly, the instruments adopted are appropriate, and secondly, weights are perfectly known and exogenous.

Gibbons and Overman (2012) discussed the issue of identification and demonstrate that spatial models (including the ones dealing with tax competition) are not identified because different structural models, i.e., models coming from different theories, turn out to have the same reduced-form representation. They argue that the solutions applied in the literature do not appear to be appropriate, as these instruments are weak and the weighting matrix does not represent real-world linkages. Therefore, it is not possible to identify exogenous variations in endogenous covariates.

Starting from this observation, we have approached this problem by applying vector autoregressive models (henceforth VAR), a tool typical of macroeconomic analysis, to a tax competition context. This class of models, developed in the 1980s, originated from the seminal work of Sims (1980), and has been widely used for both forecasting and policy analysis. Reduced-form VARs are simply a statistical representation of the endogenous variables, whose dynamics is assumed to be described by their own lags and lags of endogenous variables in the system. In order to study tax policy, therefore it is necessary to move to a structural representation of the VAR model by retrieving contemporaneous relations. For this purpose, we have adopted a recursive identification scheme, which is based on zero restrictions on some contemporaneous correlations. Given this specification, when we compute the impulse response functions to an exogenous shock in one country, we allow the competing countries to react simultaneously. To do so, we control the ordering of the equations in the VAR system. It is worth noting that this structure does not impose simultaneous reactions, although this simply allows for this possibility. In other terms, if reactions are simultaneous, we can say that countries play Nash. If however, countries react only with a certain delay, we can conclude that they play Stackelberg.

VAR models may suffer from an overfitting problem whenever there are many parameters[1] and time series are relatively short so that degrees of freedom are not enough to get precise estimates. For this reason, Bayesian techniques have recently been applied to solve the overfitting problem through the development of large Bayesian VARs (Banbura et al., 2010). This approach looks promising for our context, as tax rate data have an annual frequency and hence the number of time series observations is rather small. Here, we will apply both VAR and Bayesian VAR (henceforth BVAR) models to estimate tax competition. Our aim is to obtain a more reliable estimation of tax reaction functions and so have a better understanding of tax competition.

In our empirical exercise, we have first applied the identification approach described above to both panel and VAR models. This allows us to understand the intrinsic implications into a panel setup and also enables us to compare it with the more flexible framework represented by the VAR model. In more details, we have started from a theoretical comparison of the two frameworks and moved on to evaluate the results from panel, VAR and BVAR models. In order to keep the framework as simple as possible, we have focused on four European countries, namely France, Germany, Italy and the UK. The dependent variable is a measure of the average corporate tax rate, computed as a ratio between the corporate tax revenues and GDP. The tax rate of each country has been regressed over a set of exogenous variables including FDI inflows, government consumption growth, unemployment rate and US stock prices. The rationale for these regressors is discussed in the empirical section. Since we aim to compare outcomes from different models, we have adopted the same specification for all models, i.e., each equation of the panel contains all exogenous regressors of all countries, so as to mimic a VAR model. Finally, we have considered a more traditional panel model where contemporaneous correlations are identified instrumenting endogenous regressors. We have then compared the estimated structural parameters with those coming from a panel identified with a Cholesky decomposition of the errors' variance covariance matrix.

We have found some interesting results. Firstly, from an economic point of view, the slope of reaction functions is positive for certain countries and negative for others. This result shows that tax competition between countries is heterogeneous. Secondly, results are consistent between two countries: in other words, if the tax rate of country A is a strategic complement for the tax rate of country B, the converse is also true. Thirdly, the results obtained with a panel model are different from those from a VAR approach. This implies that the (standard) restrictions imposed by panel models may not be supported by data and can lead to biased results. Fourthly, all significant responses are simultaneous to the shock: this supports the hypothesis that countries play Nash. However, these results are tentative and may be flawed due to the fact that we have used a small sample of countries. Therefore, we have not paid too much importance to their specific economic interpretation but rather we are interested in exploring how they are affected by different modeling approaches.

From a methodological point of view, our results can be summarized as follows: (i) the role of constraints over autoregressive parameters and lagged tax rates in our exercise is limited, even if they may play a more relevant role in richer models where the country-specific dynamics is more heterogeneous; (ii) the identification strategy significantly shapes the structural parameters and the Cholesky decomposition is what allows us to find heterogeneous responses; (iii) exogenous regressors are relevant to correctly iden-

[1] The number of parameters increases with the number of endogenous variables and the lag order of the system.

tify structural relations and so are the assumptions over their coefficients; (iv) results are also affected by the inclusion of FDI inflows as endogenous variables.

Overall, our exercise has showed how results can be influenced by the various frameworks, unveiling the potential drawback of not correctly modeling countries' heterogeneity. From this point of view, not only are constraints on the contemporaneous relations relevant, but also the assumptions over all remaining coefficients may unduly affect results, even if panel models allow for country-specific intercepts. Furthermore, we have also showed how VAR can be applied to study tax competition and we have proposed to apply an identification strategy consistent with the theory of tax competition.

Given the crucial role of the identification strategy aimed at modeling responses to exogenous shocks, in the last part of the paper we have evaluated the performance of panel and VAR models identified with the Cholesky decomposition of the errors' variance-covariance matrix when the true data generating process is a VAR model. We have used a Monte Carlo exercise where two datasets are simulated assuming a VAR structure, which differ for reduced-form coefficients. In particular, these coefficients are assumed to be either the same or different across equations. We have then evaluated the ability of the panel and VAR model to estimate parameters and responses to exogenous shocks. We have found that the VAR always outperforms the panel model in estimating structural contemporaneous relationships. It also leads to estimates of impulse response functions that are at least as good as those from the panel model in the least favorable scenario. However, these differences are less pronounced the smaller the number of observations becomes. Furthermore, the panel model tends to give less biased estimates of the reduced-form coefficients when the time span is limited, because it exploits the cross-sectional dimension of the dataset and it estimats fewer parameters. As a consequence, although the structural contemporaneous coefficients are less biased when estimated with a VAR, impulse response functions from a panel model may still be reliable if datasets have few observations or if the researcher has a strong belief that the structure of the economy is similar across the sample countries.

The paper is organized as follows. Section 2 reviews the theoretical and empirical literature dealing with tax competition. Section 3 explores and compares the theoretical structure of panel and VAR models. Section 4 takes our model to the data and presents the results. Section 5 presents the Monte Carlo exercise and section 6 concludes.

2 Literature Review

2.1 Theoretical Literature

There is a large amount of theoretical literature that studies tax competition, since the pioneering articles by Zodrow and Mieszkowski (1986) and Wilson (1986). Accordingly, tax rates are strategic tools, which means that if one country sets its own tax rate, other countries are expected to react in order to attract capital and/or profit.[2]

[2]Since the beginning of 1990s the literature has focused on asymmetric tax competition. In particular, Bucovetsky (1991) showed that small countries have lower equilibrium tax rates than larger ones. The reason for this is that the benefit from capital is greater in large countries than in smaller ones. Wilson (1991) examined this topic in a more general setting and Kanbur and Keen (1993) studied a model with commodity taxation and transportation costs. Both articles supported Bucovetsky's (1991) findings.

It is worth noting that if tax rates are strategic tools, there may be either strategic complementarity or strategic substitutability. However, according to most research, tax rates are considered strategic complements. This means that an increase in a country's tax rate leads to an increase in competing countries' tax rates and vice-versa. That is, if tax rates are strategic complements (substitutes), the reaction functions have a positive (negative) slope.

The theoretical literature on tax competition has proposed several possible explanations for strategic substitutability. For instance, Mintz and Tulkens (1986) pointed out that tax rates may be strategic substitutes if private consumption and public goods are complements. To understand this reasoning, let us assume that country j increases its tax rate τ_j. As a consequence, its competing country i is likely to enjoy an inflow of capital and thus a tax base broadening. If private consumption and public goods are complements, country i may find it optimal to cut its tax rate τ_i in order to further stimulate private consumption together with public goods provision. To sum up, a cut in τ_j may lead to an increase in τ_i and vice-versa.[3]

Furthermore, tax rates can also be set either simultaneously or sequentially. In the former case, a non-cooperative Nash equilibrium holds. Accordingly, Wildasin (1991) showed that the slope of reaction functions may depend on whether countries use tax rates or public spending as a strategic tool. In particular, if tax rates (public expenditures) are strategic tools, reaction functions have a positive (negative) slope. Moreover, if one country uses the tax rate and the other country uses public expenditure as the relevant strategic tool, the reaction function of the former has a positive slope, while the converse is true for the latter one.

More recently, Vrijburg and de Mooij (2016) have proven that the slope of reaction functions may also depend on a country's objective function. If the competing countries maximize tax revenues, the reaction functions slope is always positive. If however, countries maximize welfare, a negative slope (reflecting the fact that tax rates are strategic substitutes) can be obtained with realistic parameter values.

Certain work have used a Stackelberg model, in which countries set their tax rates sequentially. In this case, a first mover (for instance, a dominant country such as the USA) sets its own tax rate and then the second player moves. Using this approach, Gordon (1992) showed that a leader country's optimal tax rate may be positive (whereas it would be nil in a non-cooperative Nash game). Altshuler and Goodspeed (2002) showed that in a Stackelberg model, the slope of countries' reaction functions may be ambiguous. In particular, it may depend on whether a country prefers to maintain public spending or private consumption.

2.2 Empirical Literature

The empirical literature has moved away from these theoretical models, as taking theoretical relationships to the data is not straightforward. Following Leibrecht and Hochgatterer (2012) we can classify empirical

[3] A similar result can be obtained if countries use public spending (instead of taxes) as their relevant strategic tool (see e.g., Wildasin, 1988). To clarify this point, assume that country j reduces public spending and hence cuts τ_j. Due to this change, country i will face a capital outflow and, hence, a decrease in its tax base. If country i is willing to keep the provision of public goods unchanged, it must increase its own tax rate. Again, a cut in τ_j may lead to an increase in τ_i . See Wilson (1999), Wilson and Wildasin (2004), Fuest et al. (2005) and Keen and Konrad (2013) for a review of the tax competition literature.

works into two different categories. The first category consists of papers that aim to estimate strategic interactions by controlling for market openness. The basic idea is that the more open the world economy, the lower the equilibrium tax rates are.

The second category consists of research that uses a direct approach. In this case, tax reaction functions are estimated by regressing each country's rate on a weighted average of other countries' tax rates as well as on some control variables.[4] The typical control variables are political and economic variables such as voter preferences, political majority, GDP, government expenditure, the structure of population, the tax rate on personal income, and measures of capital openness and trade. In these models, tax competition is studied by considering the coefficient on the average tax rate; if it is positive, tax rates are said to be strategic complements, while if it is negative they are strategic substitutes. The technical details of this approach will be explored in the next section. In terms of the results, most of this literature has found that tax rates are strategic complements (see e.g., Devereux et al., 2008, Redoano, 2014, and Egger and Raff, 2015), even though under certain conditions they can be considered to be strategic substitutes (e.g., Chirinko and Wilson, 2018).

Continuing, in this strand of literature, most works have assumed that tax competition can be described by a Nash game, which implies that tax rates are set simultaneously. The alternative is that tax rates are set sequentially following a Stackelberg game. These two theoretical assumptions have different empirical implications: if competitors play a Nash game, then other countries' current tax rates should be included as independent variables. On the other hand, under a Stackelberg game, the lagged values of other countries' tax rates should be used. Altshuler and Goodspeed (2015) are among the few authors who considered a Stackelberg game. They developed a three-country model where one country is the first mover and the other two countries play Nash after observing the first player's move. Based on this framework, they also tested the leadership role of the USA, the United Kingdom and Germany by including the lagged value of their tax rate in the estimating equation. Their empirical evidence supports the idea that only the USA has played as a leader. Another work supporting the idea that fiscal interactions occur with a lag is Hory (2017). In particular, the paper found that European countries mimick the previous year behaviour of their neighbours and the most likely explanation for this is yardstick competition.

As indicated by Revelli (2005), a tax reaction function also depends on tax base mobility. For this reason, it can be useful to investigate the relationship between the tax rate and the tax base. As such, a complete analysis of tax competition cannot solely rely on the estimation of reaction functions, but it should also consider capital mobility. Furthermore, most research has recognised the role of FDI flows in the tax competition process, though only few authors take this relationship explicitly into account.[5] Two papers modeling tax rates and tax bases are Egger and Raff (2015) and Ghinamo et al. (2010). Egger and Raff (2015) studied strategic interactions both for tax rates and in tax bases. Using a sample of both European and non-European countries (over the 1982-2005 period), they found that countries respond to the other countries' statutory tax rate cuts by reducing their own tax rates (although they are expected to raise depreciation allowances). Ghinamo et al. (2010) considered the simultaneous determination of the corporate tax rate and the FDI inflows and found that this model supports the hypothesis of strategic

[4] In some cases, the distinction between these two different categories is less neat. For instance, Ghinamo et al. (2010) use both market openness and other countries' tax rates to explain one country's tax rate setting.

[5] Typically, FDI flows are at best used to build exogenous weights.

complementarity.

Our approach lies in the latter (direct-approach) category. Namely, we have used other countries' tax rates as strategic variables and allowed for both simultaneous and lagged reactions. This means that countries are free to play either Nash or Stackelberg. To do so, we have identified strategic interactions through the Cholesky decomposition of the variance-covariance matrix of the residuals, given an appropriate ordering of the variables. If the reaction to a tax rate shock is immediate, we can say that results support the Nash-game hypothesis. On the other hand, delayed reactions are consistent with a Stackelberg strategy.

Finally, we have also taken into account the tax base determination: to do so, we have used a VAR model where FDI inflows are endogenised. Thus, the system will describe the evolution of both of these variables.

3 Methodological investigation

In this section, we compare the typical panel model used for analysing tax competition with a vector autoregression model. Our aim is to unveil the implications of these two econometric models in terms of economic theory trying to map the restrictions imposed through the panel structure in the VAR model.

3.1 The Panel Model

As pointed out, tax competition is usually analysed through a panel model where the dependent variables are the tax policy variable of the n countries in the sample. Here, we consider the most flexible specification, i.e., a dynamic panel. For the sake of simplicity, we consider only contemporaneous and lagged values of the dependent variables. We also assume that our system is composed of n equations, one for each country, where the dependent variables are the tax rates of the countries in the sample. Hence, the equation of country i takes the following form:

$$\tau_{it} = \alpha\tau_{i,t-1} + \bar{\beta}\bar{\tau}_{-i,t} + \bar{\alpha}\bar{\tau}_{-i,t-1} + g^{'}X_{it} + \phi_i + f_t + u_{it} \tag{1}$$

Here, $\tau_{i,t-1}$ is the lagged tax rate, $\bar{\tau}_{-i,t}$ and $\bar{\tau}_{-i,t-1}$ are the contemporaneous and lagged average tax rate of the other countries, X_{it} is a matrix of exogenous regressors, f_t is a time fixed-effect and $\phi_i + u_{it}$ is the error term. More precisely, the tax competition variable $\bar{\tau}_{-i,t}$ is a weighted average of the tax policy variable across countries other than i and it is formally computed as:

$$\bar{\tau}_{-i,t} = \sum_{j \neq i} w_{ij}\tau_{jt},$$

where w_{ij} are the pair-specific weights assigned to the countries in the sample. The assumption behind this structure is that country i reacts to average changes in the tax rates of the other countries. However, this assumption needs to cope with an estimation issue, as the contemporaneous tax rates are endogenous. This problem is usually overcome by instrumenting every contemporaneous tax rate with a combination of the corresponding country's exogenous characteristics. This approach has twofold implications. Firstly,

the choice of instruments is an identifying assumption that allows to retrieve structural contemporaneous coefficients. Secondly, the combination of tax rates is built by using exogenous weights. In addition, the existing literature imposes that the reaction to a change in the weighted average of contemporaneous tax variables of other countries is the same for each country, that is parameter $\bar{\beta}$ is constant across equations.

Taken together, these two assumptions shape the elements of the matrix of structural contemporaneous coefficients of the system, restricting the sum of elements on each row to be equal. More formally, in a system of n dependent variables, the matrix of contemporaneous coefficients is:

$$B_{PANEL} = \begin{bmatrix} 1 & \beta_{1,2} & \cdots & \beta_{1,n} \\ \beta_{2,1} & 1 & \cdots & \beta_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n,1} & \beta_{n,2} & \cdots & 1 \end{bmatrix}$$

Here, rows list coefficients by equation, and columns by variable. The restrictions implied by the panel models are as follows:

$$\begin{aligned} \beta_{1,2} + \beta_{1,3} + \ldots + \beta_{1,n} &= \bar{\beta} \\ \beta_{2,1} + \beta_{2,3} + \ldots + \beta_{2,n} &= \bar{\beta} \\ &\vdots \\ \beta_{n,1} + \beta_{n,2} + \ldots + \beta_{n,n-1} &= \bar{\beta} \end{aligned}$$

Returning to model 1, some authors (such as Chirinko and Wilson, 2018) also include the lagged tax competition variable, $\bar{\tau}_{-i,t-1}$. This choice is justified by the fact that imperfect capital mobility can delay tax reactions. For this reason, it is important to allow for such a lagged response in the tax rate. Also in this case, country i is assumed to react to a weighted average of other countries' lagged tax rates where for simplicity, we assume that the weights are the same as the ones used for computing $\bar{\tau}_{-i,t}$. A different although complementary source of time-dependency is the lagged dependent variable $\tau_{i,t-1}$. Very few authors[6] have estimated dynamic panels but this variable is relevant to capture gradual changes in the tax rate that are not due to the lagged tax competition variable.

The matrix X_{it} contains country-specific control variables, while the vector f_t is a time fixed-effect that can be used to capture the effects of common aggregate exogenous shocks. In most studies, this coefficient is common across countries (see e.g., Heinemann et al., 2010, and, Overesch and Rincke, 2011). Chirinko and Wilson (2018) relaxed this assumption and used the term $\zeta_i f_t$ to capture unobserved aggregate shocks, whose impact can vary by country. They showed that the introduction of f_t leads to find negative-sloped reaction functions: the underlying reason is that such shocks create common incentives that lead states to act synchronously.

Finally, the error of country i is assumed to consist of a time-invariant factor, ϕ_i, that can be considered as a fixed or random effect, and a time-varying one, u_{it}. Errors are assumed to be independent and identically distributed with the same variance across equations. Error independence is useful to simplify the estimation as, if the cross-covariances are zero, the OLS estimator is consistent. If however errors are

[6] See Chirinko and Wilson (2018), Heinemann et al. (2010).

correlated with each other, the GLS estimator should be used. Furthermore, this simplifying assumption is justified by the time trend f_t which is supposed to capture the covariance among equations.

The weights used to compute averages are determined exogenously. The relevant literature computes them by using either geographical distances, GDPs, population, FDIs, or some trade measure, and are normalized to sum to unity. Therefore, in certain cases, they are assumed to be constant over time; while in other cases they are allowed to be time-varying. The assumption that the sensitivity to other countries' tax rates can vary over time is reasonable, since economic relationships are not constant. Dimensional and institutional factors can be proxied by economic variables such as GDP. On the other hand, other measures such as FDIs or trade, aim to capture the intensity of relationships between two countries. Irrespective of the choice of weights, the appropriateness of such exogenous measures is however questionable.

In order to better explore the role of exogenous weights, let us consider the simultaneous reaction of country i to country k contemporaneous tax rate. Using model 1, it is equal to the following derivative:

$$\frac{\partial \tau_{it}}{\partial \tau_{k,t}} = \bar{\beta} w_{ik}$$

Notice that by construction, $\bar{\beta}$ is the same for all countries. This implies that the ratio between the reaction of country i to a tax change of country k and the reaction of country j to the same change is:

$$\frac{\partial \tau_{it}/\partial \tau_{kt}}{\partial \tau_{jt}/\partial \tau_{kt}} = \frac{\bar{\beta} w_{ik}}{\bar{\beta} w_{jk}} = \frac{w_{ik}}{w_{jk}}$$

This means that the relative intensity of reactions of the two countries i and j to changes in the tax rate of a third country k is completely determined by exogenous weights.

Similarly, the relative strength in the reaction of country i to the tax rates of country j and country k is:

$$\frac{\partial \tau_{it}/\partial \tau_{kt}}{\partial \tau_{it}/\partial \tau_{jt}} = \frac{\bar{\beta} w_{ik}}{\bar{\beta} w_{ij}} = \frac{w_{ik}}{w_{ij}}$$

Let us next focus on dynamic reactions. The dynamic reaction of country i to country k lagged tax rate is equal to:

$$\frac{\partial \tau_{it}}{\partial \tau_{k,t-1}} = \alpha \bar{\beta} w_{ik} + \bar{\delta} w_{ik} = \left(\alpha \bar{\beta} + \bar{\alpha}\right) w_{ik}$$

Here α, $\bar{\delta}$ and $\bar{\beta}$ are the same for all jurisdictions. As before, the relative strength in the reactions of country i with respect of country j to a change in country k tax rate and of country i to changes in the tax rates of country j and country k, is simply given by the ratio of exogenous weights, i.e.,

$$\frac{\partial \tau_{it}/\partial \tau_{k,t-1}}{\partial \tau_{jt}/\partial \tau_{k,t-1}} = \frac{\left(\alpha \bar{\beta} + \bar{\alpha}\right) w_{ik}}{\left(\alpha \bar{\beta} + \bar{\alpha}\right) w_{jk}} = \frac{w_{ik}}{w_{jk}}$$

and,

$$\frac{\partial \tau_{it}/\partial \tau_{k,t-1}}{\partial \tau_{it}/\partial \tau_{j,t-1}} = \frac{\left(\alpha \bar{\beta} + \bar{\alpha}\right) w_{ik}}{\left(\alpha \bar{\beta} + \bar{\alpha}\right) w_{ij}} = \frac{w_{ik}}{w_{ij}}$$

To sum up, in this panel model, the relative strength in responses is constant over time and solely determined by the relative weights. It is important to point out that not only are weights exogenous, but the theory of tax competition is also silent on how to compute them. This setup implies that the heterogeneity in response of country i to country j's tax rate is only governed by the exogenous weight. Some articles assume that heterogeneity can stem from the coefficient $\bar{\alpha}$ by interacting the term with dummy variables representing country-specific characteristics like capital openness or the membership to the European Union (see e.g., Redoano, 2014, and Davies and Voget, 2011). However, this solution only partially addresses the issue as these models assume that the tax competition coefficients are homogeneous within groups of countries.

To sum up, the most relevant assumptions implied by the panel structure are that: (i) countries react to a weighted average of other countries' tax rate, (ii) the coefficient capturing the reaction to the average of tax competition variables is the same for each country, and (iii) weights are exogenously given. Since these modeling constraints are not implied by tax competition theory, this does cast doubts on the appropriateness of the panel model to empirically evaluate tax competition.

From this analysis it is possible to conclude that panel models do not seem to be methodologically well suited to analyse tax competition for several reasons. However, the choice of a panel model in the empirical analysis is justified by data characteristics. As a matter of fact, data on tax competition are rather scarce as time series are short and observations are at annual frequencies. This calls for the need of shrinking the number of parameters to estimate and exploiting the cross-sectional dimension to gain some degree of freedom. Standard panel data models, through the imposition of cross-country parameter constraints, shrink the number of parameters and so allow the study of competition even with short time series.

3.2 The VAR model

Let us now focus on a VAR specification. We start with a reduced-form VAR with n dependent variables, one for each country. In order to be consistent with the previous panel model, we consider 1 lag of the dependent variables. Thus, each equation of the system has the following structure:

$$\tau_{it} = c_i + a_{i1}\tau_{1,t-1} + a_{i2}\tau_{2,t-1} + \ldots + a_{in}\tau_{n,t-1} + d_i^{'}X_t + \epsilon_{it} \tag{2}$$

Here, X_t is a $(m \times 1)$ matrix of exogenous variables, which are the same for each country so that we can drop the subscript i. This specification includes the case in which only a subset of exogenous variables enters the equation of each country. In this case, the system would be a SUR rather than a VAR.

In matrix notation, we can rewrite model 2 as:

$$\tau_t = c + A_1\tau_{t-1} + DX_t + \epsilon_t \tag{3}$$

Here, τ_t and τ_{t-1} are $(n \times 1)$ vectors of endogenous variables, A_1 and D are conformable matrices of parameters and ϵ_t is the vector of reduced-form errors. We use the standard assumption of normally distributed errors with zero mean and variance-covariance matrix Σ.

Note that Σ is left unrestricted: this means that every ϵ_{it} has a different variance and is connected with

other errors through the covariances. Reduced-form errors are simply statistical shocks, i.e., a convolution of structural unobserved shocks and dependent variables' contemporaneous relationships. From an economic point of view, it is interesting to study the reaction of dependent variables to fundamental economic shocks, as they are uncorrelated with each other and represent non-systematic economic developments.

Furthermore, the identification of structural shocks allows us to make the contemporaneous relationships between the dependent variables explicit. Indeed, extracting a matrix B_{VAR} from the reduced-form coefficients and residuals gives:[7]

$$\begin{aligned} \tau_t &= B^{-1}\psi + B^{-1}\alpha\tau_{t-1} + B^{-1}\gamma X_t + B^{-1}\varepsilon_t \\ B\tau_t &= \psi + \alpha\tau_{t-1} + \gamma X_t + \varepsilon_t \end{aligned} \tag{4}$$

System 4 is a structural-form VAR model. Here, B is a $(n \times n)$ matrix with ones on the main diagonal and general coefficients β_{ij} elsewhere and ε_t is $N(0, \Omega)$, where Ω is a diagonal matrix so that the structural errors are uncorrelated with each other, as is usual in the panel model. B is a non-singular matrix containing the contemporaneous relationships between the dependent variables. Thus, each equation of the structural VAR model has the following structure:

$$\tau_{it} = \psi_i + \alpha_{ii}\tau_{i,t-1} + \sum_{j \neq i} \beta_{ij}\tau_{jt} + \sum_{j \neq i} \alpha_{ij}\tau_{j,t-1} + \gamma_i^{'} X_{it} + \varepsilon_{it} \tag{5}$$

In more details, it is possible to move from model 2 to model 5 by retrieving the matrix B from the following relationship:

$$\epsilon_t = B^{-1}\varepsilon_t \tag{6}$$

Accordingly, the variance of ϵ_t can be written as:

$$\Sigma = \mathbb{E}\left(B^{-1}\varepsilon_t\varepsilon_t^{'}B^{-1'}\right) = B^{-1}\Omega B^{-1'} \tag{7}$$

Equation 7 defines a system of $n(n+1)/2$ equations with n^2 unknown elements. However, since $n^2 > n(n+1)/2$ for $n > 1$, there is an identification problem. More specifically, the estimation of the reduced-form VAR allows us to obtain c, A, D and Σ for a total of $\frac{3}{2}n^2 + \frac{3}{2}n + nm$ parameters. On the other hand, in order to retrieve the structural form, we would need to find B, ψ, α, γ, and Ω, i.e., $2n^2 + n + nm$ parameters. Overall, there are $\frac{1}{2}n^2 - \frac{1}{2}n$ unidentified parameters in the structural VAR model. In order to solve this problem, the literature has proposed several methodologies aimed at supplying the necessary extra identifying restrictions.

Given the identification of the structural VAR, we can now compare it with the panel model. The VAR model in 5 to collapse to the panel one in 1 by imposing the following restrictions:

1. $\alpha_{ii} = \alpha$;
2. $\sum_{j \neq i} \beta_{ij}\tau_{jt} = \bar{\beta}\sum_{j \neq i} w_{ij}\tau_{jt}$;
3. $\sum_{j \neq i} \alpha_{ij}\tau_{j,t-1} = \bar{\alpha}\sum_{j \neq i} w_{ij}\tau_{j,t-1}$;

[7] For simplicity, hereafter we will drop the subscript "VAR" unless it is strictly necessary.

4. $\gamma_i = g$.

Equalities 1 and 4 imply that the α_{ii} and γ_i are constant through equations. These constraints are relaxed in the VAR, given that both ratios are equation-specific. This means that each country is characterized by a different degree of partial adjustment in tax rates and by different sensitivity to variations in exogenous variables.

The second and third equalities regard the terms that capture the systematic reaction of tax rates to contemporaneous and lagged tax rates of other countries. In the VAR system, these terms depends on, respectively, structural coefficients β_{ij} and autoregressive coefficients α_{ii}. On the other hand, the panel model implies that the heterogeneity in the systematic responses is solely determined by exogenous weights w, as the coefficients $\bar{\beta}$ and $\bar{\alpha}$ are constant throughout all equations. This feature can be better analysed by considering the relative responses to systematic changes in other countries' tax rates. In the VAR model, the marginal reaction of τ_{it} to a change in τ_{jt} is:

$$\frac{\partial \tau_{it}}{\partial \tau_{k,t}} = \beta_{ik}$$

Therefore, the relative strength in the reaction of country i with respect of country j, to country k tax rate is:

$$\frac{\partial \tau_{it}/\partial \tau_{kt}}{\partial \tau_{jt}/\partial \tau_{kt}} = \frac{\beta_{ik}}{\beta_{jk}}$$

Similarly, the relative strength in the reaction of country i to the tax rates of countries j and k is:

$$\frac{\partial \tau_{it}/\partial \tau_{kt}}{\partial \tau_{it}/\partial \tau_{jt}} = \frac{\beta_{ik}}{\beta_{ij}}$$

The dynamic reaction of country i to country k lagged tax rate is:

$$\frac{\partial \tau_{it}}{\partial \tau_{k,t-1}} = \alpha_{ii}\beta_{ik} + \alpha_{ik}$$

The relative strength in the reactions of country i with respect of country j, to country k tax rate and of country i to the tax rates of country j and country k is:

$$\frac{\partial \tau_{it}/\partial \tau_{k,t-1}}{\partial \tau_{jt}/\partial \tau_{k,t-1}} = \frac{\alpha_{ii}\beta_{ik} + \alpha_{ik}}{\alpha_{jj}\beta_{jk} + \alpha_{jk}}$$

$$\frac{\partial \tau_{it}/\partial \tau_{k,t-1}}{\partial \tau_{it}/\partial \tau_{j,t-1}} = \frac{\alpha_{ii}\beta_{ik} + \alpha_{ik}}{\alpha_{ii}\beta_{ij} + \alpha_{ij}}$$

These derivatives show that in the VAR framework, the systematic relations between tax rates changes depend on internally estimated autoregressive coefficients α and structural coefficients β, whereas the panel model implies that these relations depends on exogenous weights w.

Given these results, let us now compare the matrices B_{VAR} and B_{PANEL} containing the structural

coefficients β_{ij} and $\bar{\beta}$. Overall, the VAR model is a more flexible framework as all reduced form coefficients are estimated and the structural form can be retrieved given supplementary identifying assumptions. In this sense, it is not necessary to use external weights or assume that some coefficients are the same across countries. Moreover, the VAR allows us potentially to extend the set of endogenous variables to any variable for which one can expect a feedback to and from tax rates. This is a crucial aspect so we can correctly identify causality relationships.

3.3 The Moving Average Representation

So far, we have compared the panel and the VAR models on the basis of structural coefficients. Since in a VAR framework the effects of economic shocks can be studied through impulse response functions, we will compare these setups by means of structural moving average representation, which describes how an endogenous variable reacts to exogenous shocks.

In order to analyse the moving average representation of the VAR system, it is useful to consider the reduced-form VAR in equation 3 and rewrite it using the lag operator:

$$A(L)\tau_t = c + DX_t + \epsilon_t, \tag{8}$$

where $A(L) = (I_n - A_1 L)$. It is then possible to derive the Wold representation by multiplying both sides of equation 8 by $A(L)^{-1}$, to obtain:

$$\tau_t = \mu + \Xi(L)\epsilon_t, \tag{9}$$

where:

$$\begin{aligned}
\mu &= \Xi(L)c + \Xi(L)DX_t \\
\Xi(L) &= (I_n - A_1 L)^{-1} = \sum_{k=0}^{\infty} \Xi_k L^k \\
\Xi_0 &= I_n \\
\Xi_k &= A_1^k
\end{aligned}$$

The structural moving average representation is obtained substituting $\epsilon_t = B^{-1}\varepsilon_t$ into equation 9:

$$\tau_t = \mu + \Xi(L)B^{-1}\varepsilon_t = \mu + \Theta(L)\varepsilon_t \tag{10}$$

where:

$$\begin{aligned}
\Theta(L) &= \sum_{k=0}^{\infty} \Theta_k L^k = \Xi(L)B^{-1} \\
&= B^{-1} + \Xi_1 B^{-1} L + \dots \\
\Theta_0 &= B^{-1}
\end{aligned}$$

$$\Theta_k = \Xi_k B^{-1} = A_1^k B^{-1},\ k = 0, 1, \ldots$$

Here, Θ_0 captures the initial impacts of structural shocks and determines the contemporaneous correlations between endogenous variables, which in our framework means that Θ_0 describes simultaneous reactions to changes in other countries' tax rates. The elements of the matrix Θ_k are the impulse responses (or dynamic multipliers) of endogenous variables to structural errors ε_t. Computationally, they are a convolution of the structural parameters in B and the estimated reduced-form coefficients in A. In more details, each equation of system 10 has the following structure:

$$\tau_{it} = \mu_i + \theta_{i1}^0 \varepsilon_{1t} + \ldots + \theta_{in}^0 \varepsilon_{nt} + \theta_{i1}^1 \varepsilon_{1,t-1} + \ldots \theta_{in}^1 \varepsilon_{n,t-1} + \ldots + \theta_{i1}^1 \varepsilon_{1,t-2} + \ldots \tag{11}$$

The impulse response functions are the plots of θ_{ij}^t against t and they describe how a shock in t affects the level of τ_i at time $t + s$.

Following the same procedure, we can now analyse the moving average representation of the panel model in equation 1, which we show here for clarity:

$$\tau_{it} = \alpha \tau_{i,t-1} + \bar{\beta} \sum_{j \neq i} w_{ij} \tau_{jt} + \bar{\alpha} \sum_{j \neq i} w_{ij} \tau_{j,t-1} + g' X_{it} + \phi_i + f_t + u_{it} \tag{12}$$

This can be easily rewritten in form of equation 4 as follows:

$$W_\beta \tau_t = \phi + W_\alpha \tau_{t-1} + X_t g + u_t \tag{13}$$

Here, τ_t and τ_{t-1} are $(n \times 1)$ vectors of tax rates, u_t is a $(n \times 1)$ vector of errors, ϕ is a $(n \times 1)$ vector of individual and time effects, X_t is a $(n \times m)$ matrix of exogenous regressors, g is a $(m \times 1)$ vector of coefficients, and W_α and W_β are $(n \times n)$ matrices with autoregressive and contemporaneous coefficients that have the following structure:

$$W_\alpha = \begin{bmatrix} \alpha & \bar{\alpha} w_{12} & \cdots & \bar{\alpha} w_{1n} \\ \bar{\alpha} w_{21} & \alpha & \cdots & \bar{\alpha} w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\alpha} w_{n1} & \bar{\alpha} w_{n2} & \cdots & \alpha \end{bmatrix}, \quad W_\beta = \begin{bmatrix} 1 & -\bar{\beta} w_{12} & \cdots & -\bar{\beta} w_{1n} \\ -\bar{\beta} w_{21} & 1 & \cdots & -\bar{\beta} w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -\bar{\beta} w_{n1} & -\bar{\beta} w_{n2} & \cdots & 1 \end{bmatrix}$$

This can be reconciled with the reduced-form VAR equation as follows:

$$\begin{aligned} \tau_t &= W_\beta^{-1} \phi + W_\beta^{-1} W_\alpha \tau_{t-1} + W_\beta^{-1} X_t g + W_\beta^{-1} u_t \\ \tau_t &= z + W \tau_{t-1} + W_\beta^{-1} X_t g + v_t \end{aligned}$$

where: $z = W_\beta^{-1} \phi$, $W = W_\beta^{-1} W_\alpha$, and $v_t = W_\beta^{-1} u_t$. Using the lag operator we can re-write this system as follows:

$$\begin{aligned} W(L) \tau_t &= z + W_\beta^{-1} X_t g + v_t \\ W(L) &= (I_n - WL) \end{aligned}$$

From this equation it is possible to derive the structural moving average representation which would have the same form as equation 11. The only difference between the moving average representation of the panel model with respect to the one of the VAR model is in the W matrix, as in the case of the panel it depends on exogenous weights w_{ij}. Therefore, the exogenous weights not only determine the systematic response of tax rates to other countries' rates, but also the impulse response functions of the system.

It is now reasonable to ask how different impulse response functions of a VAR model are with respect to those from a panel model.

3.4 Panel VAR models

A hybrid framework between panel and VAR models is represented by panel VAR models. Canova and Ciccarelli (2013) provided an overview of panel VAR models, as well as their use, the estimation algorithms and possible identification strategies.

A panel VAR model describes the evolution of each endogenous variable with their own lags and lags of other endogenous variables in the system, in line with VAR models. However, it also exploits the cross-sectional dimension of data (by grouping variables into units). Therefore, panel VARs have three distinguishing features. Firstly, they allow for dynamic inter-dependencies because lags of all endogenous variables enter the equations of each unit. Secondly, they can capture static inter-dependencies by allowing the errors to be correlated across units. Thirdly, they can feature cross-sectional heterogeneity if the intercept, the coefficients and the variance of shocks are assumed to be unit-specific.

As such, panel VAR models appear to be well-suited to study the transmission of idiosyncratic shocks across different units. However, the flexibility of these models has a cost in terms of the very high number of parameters to estimate. Therefore, restricted models are often estimated in practical application. Restrictions may be imposed setting the dynamic and static inter-dependencies to zero or assuming cross-sectional homogeneity in certain parameters. Hence, starting from an unrestricted panel VAR model, it is possible to formulate a large number of restricted models.

Subsequently, a panel VAR can be seen as a restricted version of a large-scale VAR model. For the purpose of this work, the formal analysis of panel VARs would not add much to the comparison between simple panel and VAR models.

4 Empirical investigation

Let us now move our attention towards an empirical comparison of the models described above. In particular, we will compare the outcome from different econometric models, namely a panel, a VAR, and a Bayesian VAR model. We will then start with a panel model where all coefficients are constant through equations. We will then partially relax this assumption and introduce a country-specific autoregressive term. Finally, we will move to VAR and BVAR models, where no restrictions are imposed on reduced-form coefficients. This exercise will allow us to understand whether the adoption of alternative modeling strategies leads to different conclusions about the importance and nature of tax competition.

As pointed out, empirical studies have analysed tax competition including a weighted average of contemporaneous tax rates of other countries. Since these tax rates are endogenous, they are usually

instrumented with the subject country's exogenous characteristics. Here, we follow a different approach in that contemporaneous relations are not directly modeled but rather derived from the structural identification of reduced-form models. In other words, our models only consider lagged dependent variables, and susequently exploit contemporaneous residuals' correlations to compute the effects of a change in tax rates. Given our interest in reactions to unexpected tax changes, the first part of our analysis compares impulse response functions stemming from different frameworks.

More specifically, in all subsequent models, identification is achieved through a Cholesky decomposition of the variance-covariance matrix of residuals as described by equations 6 and 7, i.e., we impose certain short-run zero restrictions. This procedure models contemporaneous relations through a lower-triangular matrix, which implies that the order of variables is crucial to determine the simultaneous effects of a shock. In other terms, variables placed below the variable we shock are assumed to react simultaneously, while variables ordered first are allowed to react with a lag.[8] In our context, we let countries react simultaneously to changes in a country's tax rate so that the country that moves its tax rate is placed first in the system. This assumption is justified by the periodicity of our data, which is annual, so that it is plausible that even after an unexpected shock, countries have enough time to move their tax rate within the same period. However we do not impose that the contemporaneous responses are statistically significant but rather we simply allow for this possibility. Indeed, it may well be that significant responses occur with a lag rather than simultaneously. In this sense, our identification strategy is appropriate as it is agnostic to the timing of the response. For the comparability of the responses across different models, we normalize the size of the initial shock to one percentage point.

A further important issue is that of computing interval bands for impulse responses. Confidence intervals for impulse response functions are computed using a residual-based bootstrap with 3000 replications. Subsequently, the 5th and 95th percentiles are displayed. In particular, after re-sampling the residuals, the dependent variable is generated recursively using the estimated parameters, the exogenous regressors and initial values of the original series.

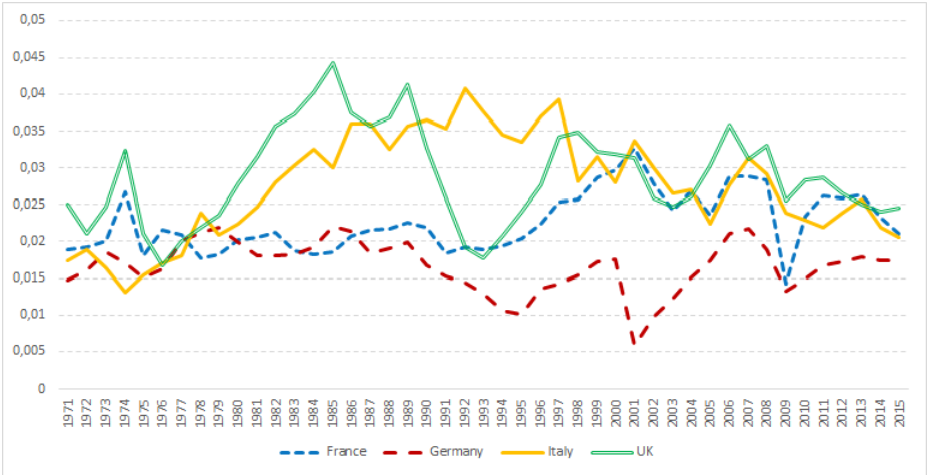
In order to link our work to the literature, the last step of our analysis focuses on the matrices of contemporaneous coefficients, as they are used to evaluate tax competition. In this respect, we consider a panel model identified with instrumental variables, whose objective is to represent the typical model used in the literature. From this model, we take the coefficient of contemporaneous relations and compare it with the matrix of structural coefficients of a panel model identified with the Cholesky decomposition. This comparison clarifies the differences, in terms of contemporaneous coefficients, between a panel identified with instrumental variables and a panel identified with zero contemporaneous restrictions.

4.1 Data

Our dataset consists of tax rates, foreign direct investment inflows, government consumption, unemployment and US stock prices. All series are annual and range from 1971 to 2015. To measure the tax rate, we use the average corporate tax rate defined as the ratio between corporate tax revenues and GDP as computed by the OECD Revenue Statistics. This measure provides us with fairly long time series of

[8] For this reason, the Cholesky identification is also called recursive identification as it is said to be based on the recursiveness assumption.

Figure 1: Average corporate tax rates



tax burden on corporate income, profits, and capital gains and has been used, among others, by e.g., Besley et al. (2001) and Altshuler and Goodspeed (2015).[9] Data concerning FDI inflows, supplied by the UNCTAD, are on a net basis, i.e., they are net changes in the amount invested in national assets and are expressed in US dollars. In order to account for each country's budget constraints, we use growth in government consumption expressed in real terms. Data come from the World Bank and series are in the national currency. Finally, unemployment and US stock prices are assumed to capture the national and world business cycle, respectively. In detail, unemployment data come from either the World Bank or the OECD, while for US stock prices, we use the returns on the S&P500 index.

In order to keep the model simple and tractable, we have decided to focus on four European countries with similar characteristics: France, German, Italy and the United Kingdom. We chose these countries for at least four reasons. Firstly, they are comparable in terms of both wealth (GDP in 2016 ranges from 1680 billion Euros to Italy to 3144 billion Euros in Germany)[10] and population. Secondly, they have a common cultural heritage.[11] Thirdly, these countries are relatively close or even bordering; so, transportation costs are not a significant driver of FDI flows. Fourthly, given their membership in the EU (and previously in the EEC, from 1973), these countries have long being experiencing free trade. Without tariffs therefore, we can focus on corporate taxation as a possible determinant of competition.

Figures from 1 to 5 display the variable used in this empirical exercise. Certain preliminary data analyses are necessary to correctly specify the model. In particular, we need to test whether tax rates are stationary and whether they are cointegrated.

To check the presence of a unit root we have used the ADF test on the individual variables with one or two lags and with constant, or constant and trend. All these specifications confirmed the stationarity

[9] Many other papers use statutory corporate tax rates or forward-looking effective tax rates (see e.g., Devereaux and Loretz, 2013, Leibrecht and Hochgatterer, 2012, and Baskaran and Lopes da Fonseca, 2014). Unfortunately, these measures are usually available for a shorter period. Moreover, forward-looking rates are usually based on some arbitrary assumptions. As pointed out by Altshuler and Goodspeed (2002), e.g., they do not consider certain aspects such as enforcement. Furthermore, they do not account for the heterogeneity due to double tax treaties and are neither firm- nor project-specific.

[10] These figures come from the Eurostat.

[11] We are aware that UK applies Common Law while the other three countries use civil law. However, we use country-specific dummies to control for heterogeneity.

Figure 2: FDI inflows

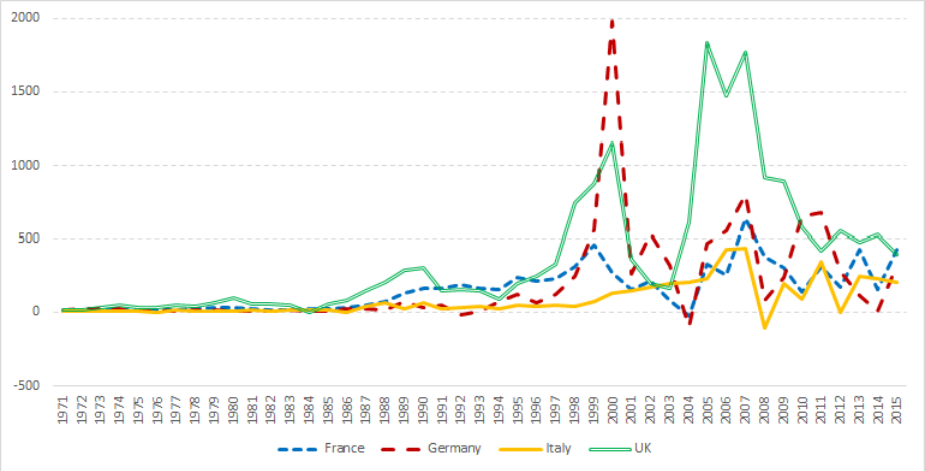


Figure 3: Real government consumption YoY

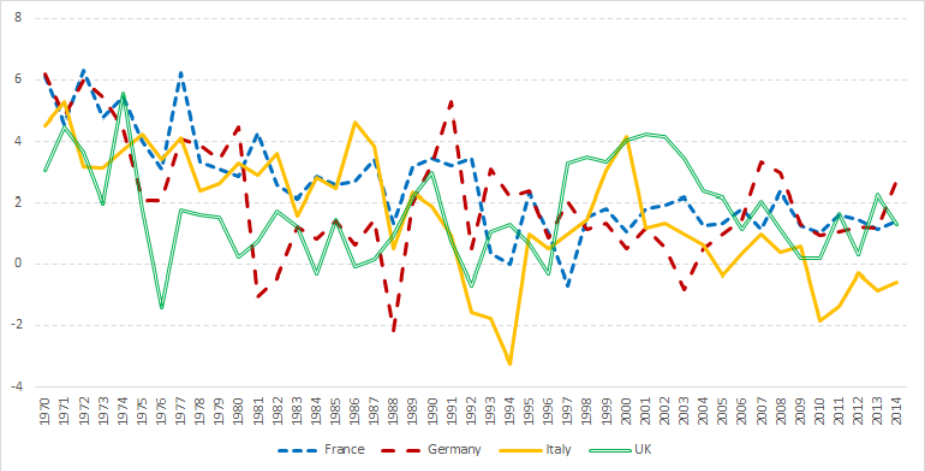


Figure 4: Unemployment rate

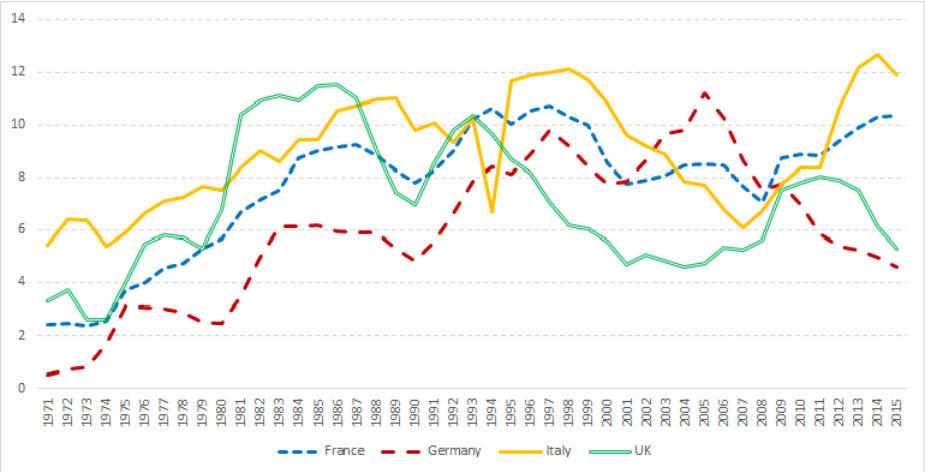
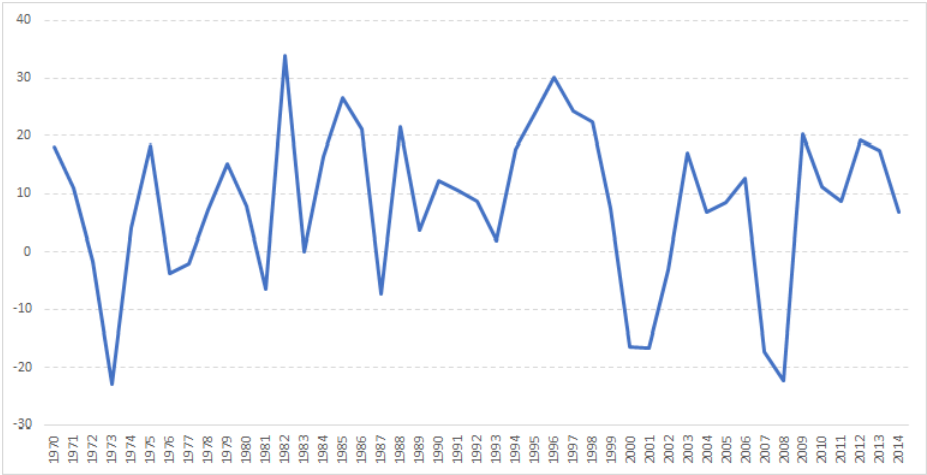


Figure 5: Return of the S&P500 index



of all series with the exception of Italy. We have also performed panel unit root tests, and in particular we used the Levin et al. (2002), the Im et al. (2003) and Fisher-type tests. The null hypothesis of these tests is that all panels contain a unit root while they differ for the alternative hypothesis, which are that all panels are stationary for the Levin et al. (2002) test, some panels are stationary for the Im et al. (2003) test and that at least one panel is stationary for the Fisher-type test. The results from these tests are mixed and reflect the fact that one of the tax rates is non-stationary.

We have checked cointegration both in a time series and in a panel setting. As for the time series framework, we have used the Johansen test that estimates a VECM model and runs two tests on the matrix of coefficients of the equilibrium term. These tests are the trace test and the maximum eigenvalue test, and they show that there is no cointegration if we consider an unrestricted constant and one lag of the dependent variables.

In the panel setting, we have evaluated the presence of cointegration with the tests developed by Westerlund (2007). The null hypothesis of these tests states that there is no cointegration, while the alternative hypothesis is either that the panel is cointegrated as a whole or that at least one unit is cointegrated. These tests specify an ECM for each variable and evaluate whether the coefficients of the lagged dependent variable in levels are statistically different from zero as they represent the speed of adjustment to the equilibrium relationship, i.e., the relevance of the ECM term. We have also added a constant term and the values of other countries' tax rates as exogenous regressors to the model. We have then included either one or two lags of the dependent variables as regressors, and either none or one lead of exogenous regressors. For all of these models, the four tests accept the null hypothesis of no cointegration with rather high p-values.

Therefore, we will now proceed by using the first difference of the tax rates as dependent variables.

4.2 Panel models

In this section, we estimate several different panel models starting from the specification that constrains most regression parameters'. We then progressively relax these assumptions to move closer to a VAR model. Our first goal is to analyse the impulse response functions implied by a panel model identified

with the Cholesky decomposition of the errors' variance-covariance matrix, as in the VAR model, and we normalize the initial shock to 1 percentage point. The system is estimated using least squares, which appears to be the best approach as we deal with a low number of countries ($N = 4$). In particular, we use four specifications which differ for assumptions on the autoregressive coefficient and the coefficient on the lagged average tax rate. Here, the dependent variable is the first difference of tax rates and in order to keep notation simple, we refer to it as τ.

To begin with, we consider a simplified version of equation 1, where we drop the time fixed-effects and contemporaneous response of other countries' tax rates:

$$\tau_{it} = \alpha\tau_{i,t-1} + \bar{\alpha}\bar{\tau}_{-i,t-1} + g^{'}X_{i,t-1} + \bar{g}^{'}X_{-i,t-1} + \phi_i + u_{it} \tag{14}$$

Here, each tax rate is regressed over its own first lag $\tau_{i,t-1}$, the average of the lagged tax rates of the other countries $\bar{\tau}_{i,t-1}$, a country-specific dummy variable ϕ_i and a set of control variables in the matrices $X_{i,t-1}$ and $X_{-i,t-1}$, i.e., FDI inflows, the growth rate of government consumption, unemployment and the return of the S&P500 index. With regard to these exogenous variables, we divide them into two matrices: $X_{i,t-1}$ contains the own variables and $X_{-i,t-1}$ contains the variables of the other countries. The coefficients of these two groups are restricted to be equal across countries. To simplify the model, the average tax rate is computed using the arithmetic mean of the other countries' tax rates. Note that each equation contains the values of FDIs, public spending and unemployment of all the countries in the sample and these variables are lagged by one period to avoid endogeneity issues. This contrasts to the usual procedure in the related literature as, in a panel setup, each countries' dependent variable is regressed over its own contemporaneous control variables and endogeneity is evaluated with a Hausman test. However, we insert lagged regressors of all countries in each equation as this allows us to obtain a specification comparable with a VAR model. Furthermore, it is also economically meaningful to assume that each country takes into account other countries' variables when setting its own tax rate. Finally, the coefficients α, $\bar{\delta}$ and g are restricted to be equal across countries.

Let us now consider the case in which the autoregressive coefficients are country-specific. The panel model becomes:

$$\tau_{it} = \alpha_i\tau_{i,t-1} + \bar{\alpha}\bar{\tau}_{-i,t-1} + g^{'}X_{i,t-1} + \bar{g}^{'}X_{-i,t-1} + \phi_i + u_{it} \tag{15}$$

We therefore relax the assumption that tax rates react to other countries' average tax rates by including individual country tax rates. The first version of this model assumes that all coefficients except for ϕ_i are common across equations:

$$\tau_{it} = \alpha\tau_{i,t-1} + \sum_{j\neq i}\alpha_j\tau_{j,t-1} + g^{'}X_{i,t-1} + \bar{g}^{'}X_{-i,t-1} + \phi_i + u_{it} \tag{16}$$

The second version of this model allows for country-specific autoregressive terms.

$$\tau_{it} = \alpha_i\tau_{i,t-1} + \sum_{j\neq i}\alpha_j\tau_{j,t-1} + g^{'}X_{i,t-1} + \bar{g}^{'}X_{-i,t-1} + \phi_i + u_{it} \tag{17}$$

The impulse response functions to shocks of the four tax rates are displayed in figures from 6 to 9. In each figure, four sets of responses are presented given that our identification strategy requires each tax rate to be ordered first when the effects of its unexpected change is evaluated. In all cases, we display the cumulated impulse response functions, so that the effects on tax rate levels can be evaluated. Table 1 summarizes the information from impulse response functions.

Overall, several impulse response functions are significant and, at a first sight, results appear to be similar across different models. However, some interesting insights can be inferred from this exercise. First of all, by comparing the impulse response functions from model 14 with those from model 15 and those from model 16 with those from model 17 we can evaluate the effects of using country-specific autoregressive terms instead of assuming that they are constant across countries. The differences in the two frameworks with respect to the significance of impulse response functions are very small, even though the dynamics of several responses is affected by the different assumption over the autoregressive term. However, this does not have any relevant effect over the significance of responses, whose differences across the two frameworks are negligible. Nevertheless, this result may be limited given that our model only considers the first autoregressive term, while in a model with richer dynamics the effect of restricting the autoregressive terms may be much larger. This insight concerning constraints on the autoregressive term is noteworthy, as it unveils that a constraint may relevantly affect the reaction of tax rates to exogenous shocks.

Additionally, a comparison of the impulse response functions derived from both model 14 with 16, and model 15 with those from model 17, allows us to evaluate the effects of considering the lagged tax rates of other countries as regressors rather then their weighted averages. Our results are more significant on condition that the coefficients of the lagged tax rates are assumed to be constant throughout rather than considering each tax rate individually. This may be due to the fact that, in the latter case, some degree of freedom is lost.

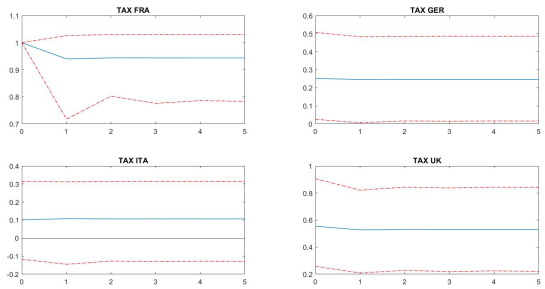
Significant responses show that UK reacts to changes in the French and German tax rates and vice-versa. In particular, the French and German tax rates are strategic complements for the UK rate. Furthermore, there is mild evidence of strategic substitutability between the tax rate of Germany and Italy and between those of Italy and the UK. This evidence is weak because these are only one-way relationships, i.e., the Italian tax rate reacts to changes in the German tax rate but the opposite is not true, while the UK tax rate reacts to changes in the Italian tax rate but not vice-versa. These results are summarized in table 1. Therefore, from a tax competition point of view, results are mixed and no common pattern can be identified in tax responses.

As regards competition, all significant responses are contemporaneous to shock so that it is possible to conclude that our results support the hypothesis that countries play a Nash game.

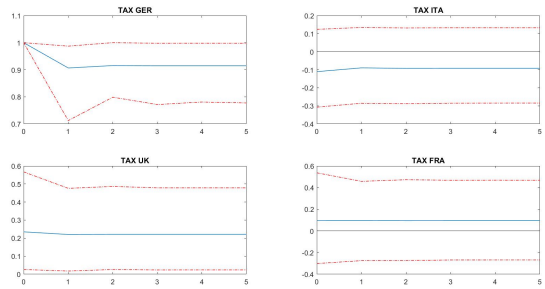
Finally, the size of responses, when significant, is always smaller than the size of the shock and never exceeds 0.5 percentage points on average, i.e., the countries react less than one-to-one to the shock.

Figure 6: IRFs of panel model 14

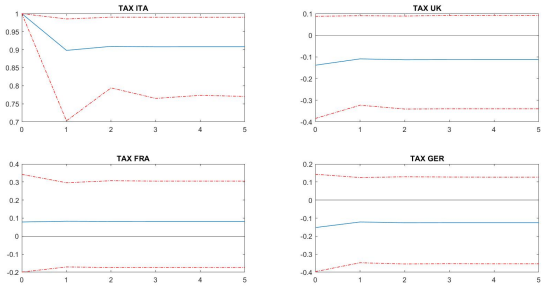
(a) Shock to the French tax rate



(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate

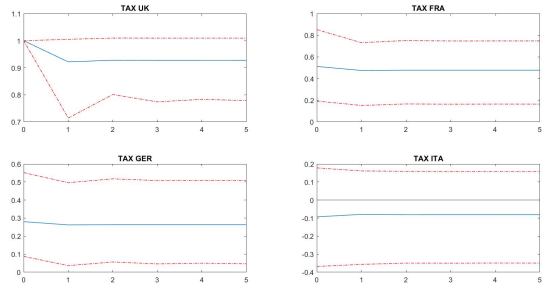
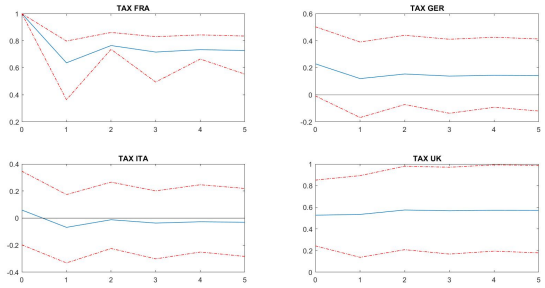
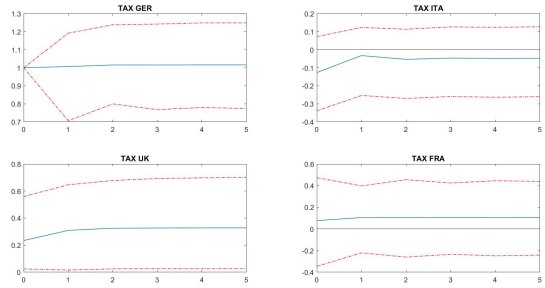


Figure 7: IRFs of panel model 15

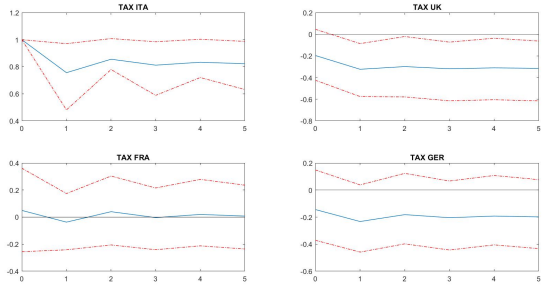
(a) Shock to the French tax rate



(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate

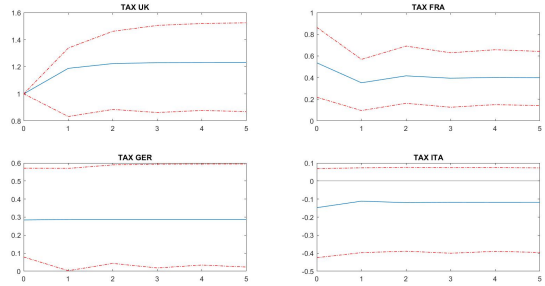
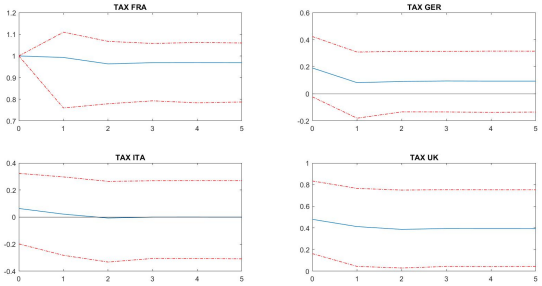
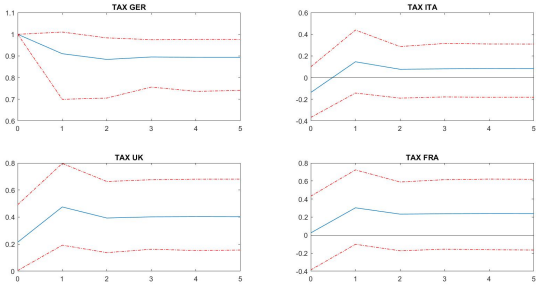


Figure 8: IRFs of panel model 16

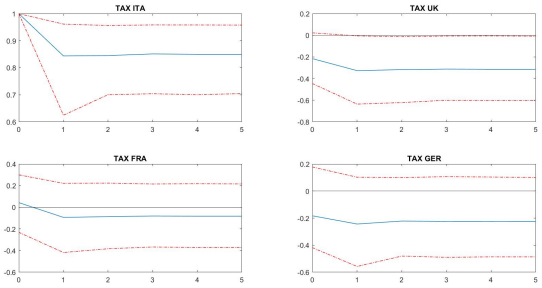
(a) Shock to the French tax rate



(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate

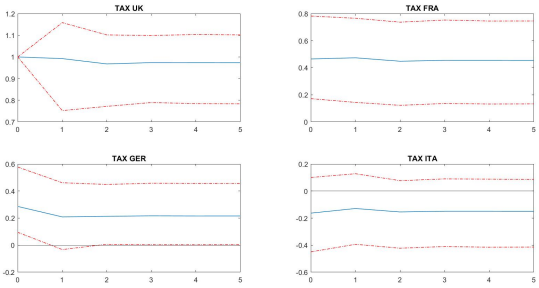
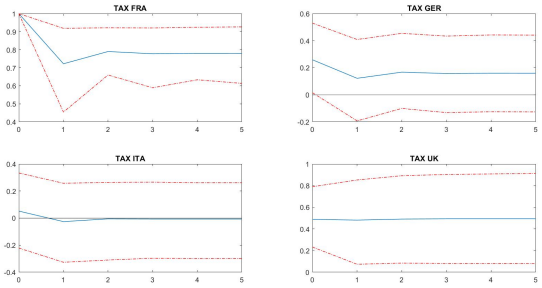
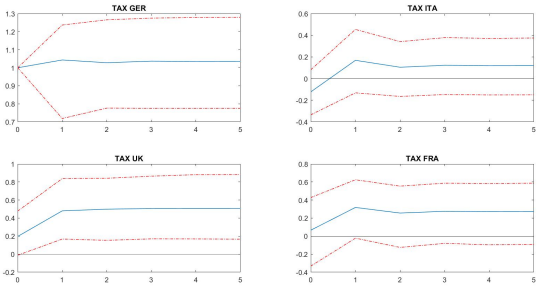


Figure 9: IRFs of panel model 17

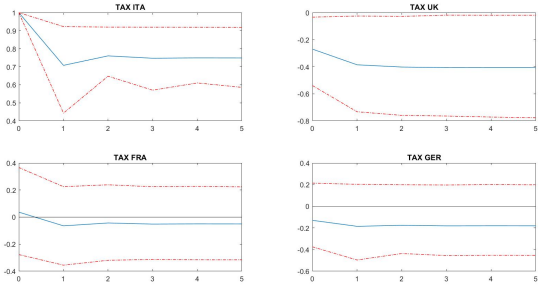
(a) Shock to the French tax rate



(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate

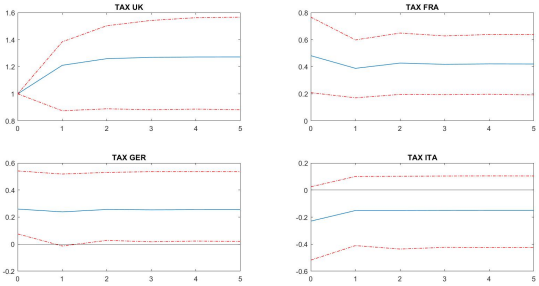


Table 1: Summary of information contained in IRFs from panel models

		Exogenous Shock			
		France	Germany	Italy	UK
Reactions	France	***	0	0	+
	Germany	0	***	0	+
	Italy	0	-	***	0
	UK	+	+	-	***

As pointed out, the literature focuses on structural coefficients from the estimation of panel models, rather than looking at impulse response functions. In order to establish a closer link with tax competition literature, we also analyse and compare the structural coefficients of model 14 with those from a model identified with instrumental variables:

$$\tau_{it} = \alpha\tau_{i,t-1} + \bar{\beta}\bar{\tau}_{-i,t} + \bar{\alpha}\bar{\tau}_{-i,t-1} + g^{'}X_{i,t-1} + \bar{g}^{'}X_{-i,t-1} + \phi_i + u_{it} \tag{18}$$

This is equation 1, where time fixed-effects have been dropped and the regressors of the other countries have been added. The difference compared to model 14 is that here we directly consider the average tax rate of other countries $\bar{\tau}_{-i,t-1}$ and estimate $\bar{\beta}$ instead of deriving it from the errors' covariance matrix. As in the existing literature, we face an endogeneity issue relating to contemporaneous tax rates and we address that by instrumenting $\bar{\tau}_{-i,t}$ with a weighted average of public expenditures, unemployment and US stock prices. The standard diagnostic tests for underidentification, weak identification and overidentification are satisfied. The estimated value of the parameter $\bar{\beta}$ is 0.9635, which is significant at a 5% level. A conventional analysis would conclude that there is evidence of tax competition and that tax rates are strategic complements. However, here we want to map this value into matrix B of contemporaneous relations of model 14. To do so, we need to divide the value of $\bar{\beta}$ by the number of tax rates included in the term $\bar{\tau}_{-i,t}$. The matrices of contemporaneous coefficients from models 14 and 18 are as follows:

$$B_{14} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.2514 & 1 & 0 & 0 \\ -0.1535 & 0.2068 & 1 & 0 \\ -0.5172 & -0.2103 & 0.1514 & 1 \end{bmatrix}$$

$$B_{18} = \begin{bmatrix} 1 & 0.3212 & 0.3212 & 0.3212 \\ 0.3212 & 1 & 0.3212 & 0.3212 \\ 0.3212 & 0.3212 & 1 & 0.3212 \\ 0.3212 & 0.3212 & 0.3212 & 1 \end{bmatrix}$$

As for matrix B_{14}, we considered the case in which France is ordered first and the implied impulse response functions are displayed in panel a of figure 6. On the other hand, the ordering of variables does not influence the values in matrix B_{18}. Both matrices of contemporaneous coefficients are shaped by the relative identification strategies although matrix B_{18} is also influenced by restrictions over coefficients.

Model 14 implies that some tax rates are strategic complements, while others are strategic substitutes. On the other hand, model 18 shows that all tax rates are strategic complements. This analysis confirms that the panel model unduly constrains the reactions of tax rates to exogenous shocks.

Overall, this section highlighted that the effects of constraining the coefficients of autoregressive terms and lagged tax rates of other countries had very marginal effects on the results. A much more relevant role in retrieving the true strategic relations was played by the identification strategy as contemporaneous structural parameters may differ considerably according to the strategy pursued.

4.3 VAR models

Let us now focus on a standard VAR model that can be conceived as an unconstrained version of the panel models above as all coefficients are allowed to be equation-specific. The reduced form of the system is described in equation 3, shown here for clarity:

$$\tau_t = c + A_1\tau_{t-1} + GX_{t-1} + \epsilon_t \tag{19}$$

As in the panel models, each country's tax rate is regressed over the first lag of all tax rates in the system. The first lag of FDI inflows, government spending growth, unemployment and contemporaneous value of the return on the S&P500 index are used as exogenous regressors.

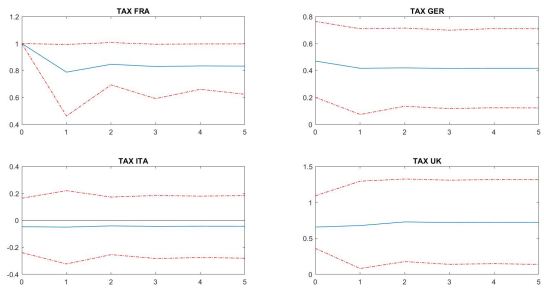
Figure 10 shows the sets of impulse response functions to one percentage point unexpected increase in each tax rate, together with 90 percent bootstrapped confidence intervals. Table 2 summarizes the results from the VAR model. As before, the UK tax rate increases after an unexpected change in the tax rate of France and Germany and the opposite is also true. Furthermore and unlike previous findings, the French and German tax rates also react positively to shocks in each others' tax rates. This implies that the tax rates of these countries are strategic complements. On the contrary, the Italian and UK tax rates are strategic substitutes as they show a marginal decrease after an unanticipated increase in each others' tax rates. Finally, the German and Italian tax rates do not react significantly to shocks in each others' tax rates.

As already pointed out, these responses are somewhat different from those from panel models, even from the less constrained one represented by equation 17, implying that the choice between a panel or a VAR setup is a relevant one. More specifically, the discrepancies between the results of the different panel models appear to be much smaller than the differences detected comparing the results from models 17 and 19. In particular, the VAR model leads to more significant results and the size of the significant responses is always greater than in the panel model. In some cases the average response is even twice as large as those stemming from the panel setup and around or above 0.5 percentage points. The panel models embed various assumptions over the coefficient of their autoregressive terms and the lagged tax rates of other variables. However, the less constrained panel model 17 and VAR model 19 differ for the constraints over the coefficient of the exogenous variables, which are assumed to be constant through equations in the panel models and equation-specific in the VAR model. Therefore, the differences between model 17 and 19 results can be attributed to the sensitivity of the tax rates to exogenous variables. This finding emphasizes the role of exogenous variables for correctly identifying structural parameters. Finally, as before, all significant

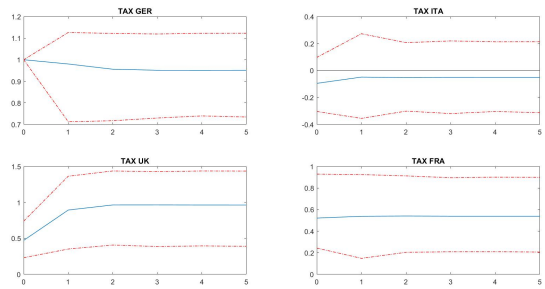
responses are simultaneous to the shock and so again we can conclude that competition across countries can be considered as a Nash game.

Figure 10: IRFs of VAR model for Tax Rates

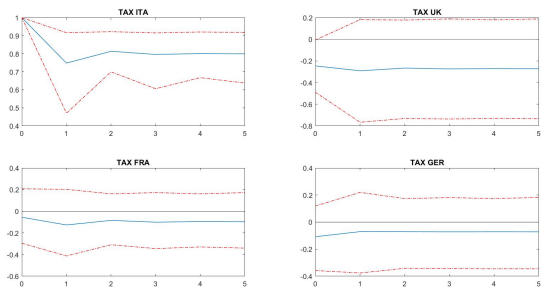
(a) Shock to the French tax rate



(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate

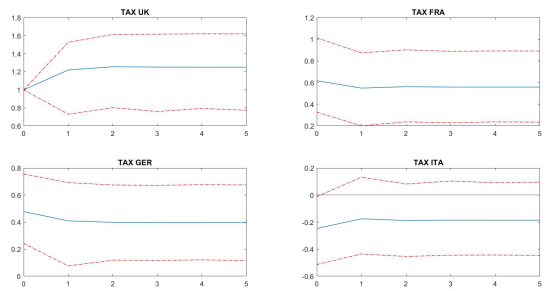


Table 2: Summary of information contained in IRFs from VAR models

		Exogenous Shock			
		France	Germany	Italy	UK
Reactions	France	***	+	0	+
	Germany	+	***	0	+
	Italy	0	0	***	-
	UK	+	+	-	***

In addition, VAR models allow us to more exhaustively capture the implications of exogenous changes in tax rates. Indeed, as pointed out by Revelli (2005), a tax reaction function arises because the tax base is mobile and reacts to tax rate differentials. Therefore, a complete analysis of tax competition cannot solely rely on the estimation of tax reaction function, but should also consider capital as an endogenous variable. In order to consider an encompassing model, we generalized the previous one so that FDI inflows are endogenised giving the following system:

$$\begin{bmatrix} x_t \\ \tau_t \end{bmatrix} = c + A_1 \begin{bmatrix} x_{t-1} \\ \tau_{t-1} \end{bmatrix} + \epsilon_t \tag{20}$$

Here x_t is a $(n \times 1)$ vector of FDI inflows of the countries considered and, as before, τ_t is a $(n \times 1)$

vector of tax rates so that the whole system has 2 equations for each country.

In the literature, very few studies have considered a tax base and tax rate equation for each entity and those that do so, estimate the systems with an instrumental variables approach, e.g., Brett, Pinske (2000).

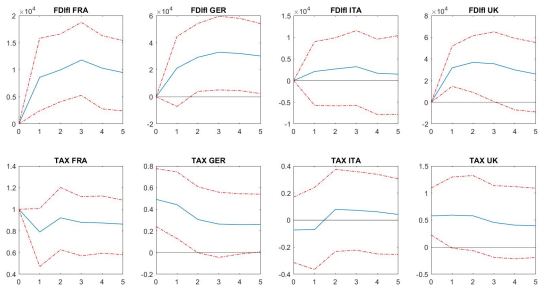
The advantages of the VAR approach over a panel setup also apply in this case. More specifically, the identification of an exogenous shock to tax rates follows the same rationale as before but besides we also assume that FDI inflows react to tax rate shock with one lag and their equations are ordered first in the system.

Figure 11 shows our results. As can be seen, the results are slightly different from the previous ones. As a matter of fact, here we are directly modeling the contemporaneous relationships between tax rates and FDI inflows. More specifically, the results regarding the tax rates are qualitatively the same as those coming from model 19. There is some difference in the size of certain impulse response functions and overall confidence bands are slightly tighter when FDI are considered as endogenous. This seems at odds with the increased number of parameters which may lead to an overfitting problem. However, the effects of tax rate shocks in model 20 are shorter-lived compared to model 19, as some impulse response functions become indistinguishable from zero after a few periods. However, these discrepancies are marginal and do not affect the general interpretation of the strategic relations, even if they highlight the fact that including FDI inflows among the endogenous variables can influence tax reaction functions.

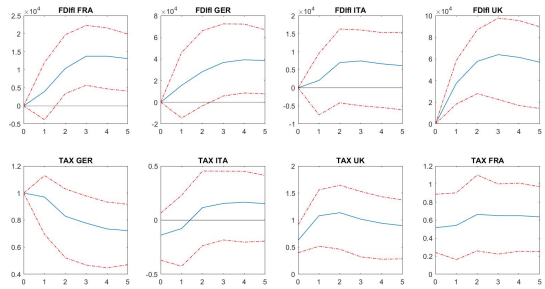
It is worth noting that FDI inflows responses are often non significant. The significant responses show that FDI inflows increase after a rise in the tax rate and the magnitude is quite considerable given the average size of inflows so that our results cast doubts on the effectiveness of tax competition as a policy to attract FDIs. Of course, this result is preliminary and further investigation is desirable to confirm it. Indeed, there are two possible explanations for this finding. Firstly, FDI may be shaped not only by changes in the tax rate of one country, but also by changes in the tax rates of other countries. Secondly, there may be other factors which we do not explicitly consider in this model: e.g., tax rate changes in a fifth country or other developments in some other economy. From this point of view it would be desirable to extend the model to include other relevant economies. For now, our objective has been to show what can be done with VAR models. Furthermore, a model extension would lead to a sharp increase in the number of parameters to be estimated.

Figure 11: IRFs of VAR model for Tax Rates and FDI Inflows

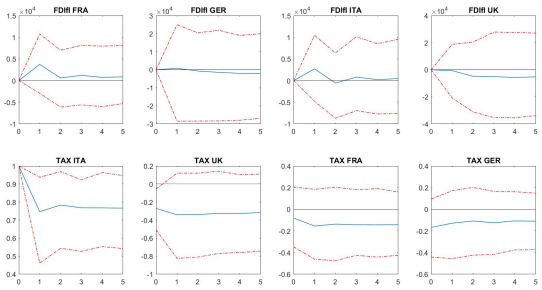
(a) Shock to the French tax rate



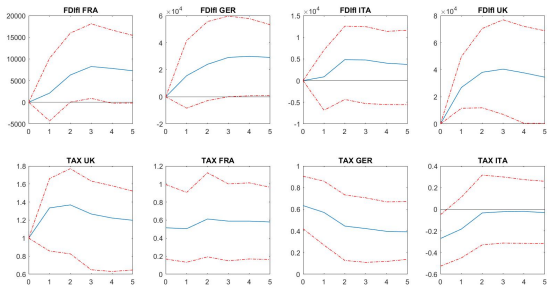
(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate



Overall, this exercise has allowed us to understand that, not only is the identification strategy relevant, but also, how the constraints of model's parameters can distort inference regarding structural relationships. This distortion appears to be relevant for some impulse response functions and marginal for others. As such, it is not possible to exactly predict the effect of assuming that certain parameters are constant throughout equations. However, we do argue that our application demonstrates that such a distortion may be large enough to consider adopting a framework that does not impose unrealistic constraints or that, at least, these constraints are strongly supported by both theoretical and empirical evidence.

4.4 Bayesian VAR models

The last step of our analysis compares our previous findings with large Bayesian VAR models, introduced by Banbura et al. (2010) as a tool to handle systems of many variables. Indeed, a drawback of VAR models is that they usually have many parameters to estimate. When time series have only a few observations, it becomes very easy to fall into the overfitting problem, i.e., that there are not enough degrees of freedom left to obtain significant estimates. Thus, Bayesian inference allows to avoid the overfitting problem by using informative priors that shrink an overparametrized model towards a more parsimonious one represented by the prior distributions. This approach is essentially based on Bayesian shrinkage which means increasing the tightness of the prior as more variables are added into the model. In particular, Banbura et al. (2010) proposed to choose the shrinkage parameter to target a desired in-sample fit. We however adopt the Giannone et al. (2015)'s setup in which the model is given a hierarchical structure, hyperparameters are treated as unknown and the appropriate degree of shrinkage is automatically. Bayesian inference is also

applied to hyperparameters by assigning them a flat hyperprior so that maximizing their posterior simply amounts to maximizing the marginal likelihood with respect to them, i.e., an empirical Bayes approach.

More specifically, the prior on the coefficients and on the variance-covariance matrix is a Normal - Inverse-Wishart distribution, namely:

$$\begin{aligned} \Sigma_\varepsilon &\sim IW\left(\psi I_n, n+2\right) \\ vec\left(A\right)|\Sigma_\varepsilon &\sim N\left(a, \Sigma_\varepsilon \otimes \Omega\right) \end{aligned}$$

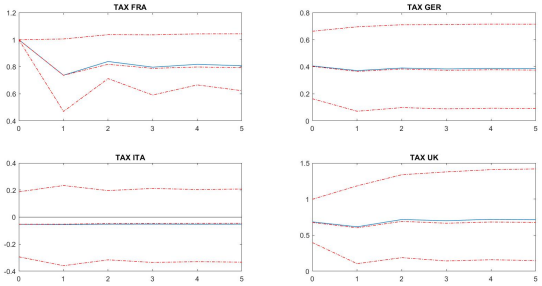
Here, the parameters a and Ω are functions of a small set of hyperparameters γ, and the scale parameter ψ contains the residual variances of an AR(1) process fitted on each VAR equation, as prescribed by the Minnesota prior. Bayesian shrinkage is achieved through the combination of Minnesota, sum-of-coefficients and dummy-initial-observation priors for the VAR coefficients. The Minnesota prior assumes that the limiting form of each VAR equation is a random walk with drift, while the sum-of-coefficients prior and the dummy-initial-observation prior are necessary to account for unit root and cointegration.

The estimation algorithm combines a Metropolis step to draw the vector of hyperparameters with a Gibbs sampler to draw the models' parameters conditional on the former. From the conditional posterior distribution, we extract 20.000 draws, the first 10.000 are discarded and the latter 10.000 are used for inference. Further details on the prior specification and estimation procedure can be found in Giannone et al. (2015).

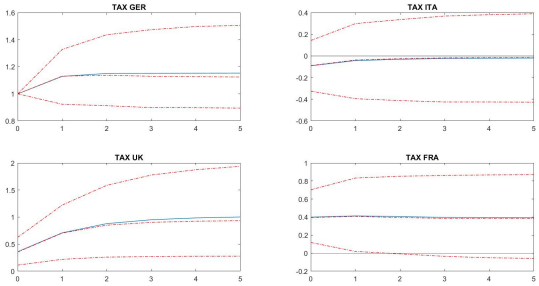
Using this tool, we estimate that the same model described in the previous section, and figures 12 and 13 display results that are very similar to those from the VAR model. This can be considered as proof of the validity of the Bayesian tool we used as it means that the priors are not constraining parameters to unrealistic values. More specifically, the qualitative interpretation of impulse response functions is unchanged and the summary in table 2 remains valid. Moreover, the magnitude and dynamics of the responses is slightly, but not significantly different. This is so, both when comparing the results from the BVAR with those from the VAR model, i.e., figures 10 and 11 with figures 12 and 13 respectively, as well as when comparing figure 12 with figure 13, i.e., the results from two different specifications of the VAR model. Furthermore, given the use of prior distribution to shrink the dimensionality of the system, one would expect to find tighter credibility intervals. In our application, this is true only for certain impulse responses and mainly on impact or on short horizons, and it does not alter our previous findings. However, this exercise shows that Bayesian inference can help to retrieve significant structural relations in large systems where classical inference cannot overcome the overfitting problem. From this point of view, the advantage of such an estimation procedure is grater, the larger the system of equations to estimate is.

Figure 12: IRFs of BVAR model for Tax Rates

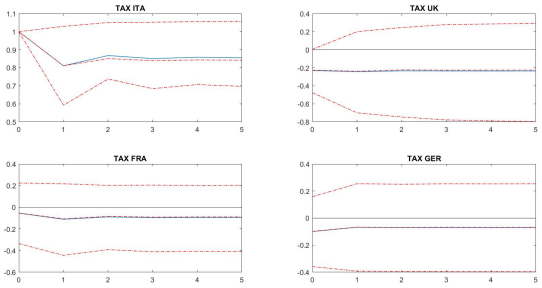
(a) Shock to the French tax rate



(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate

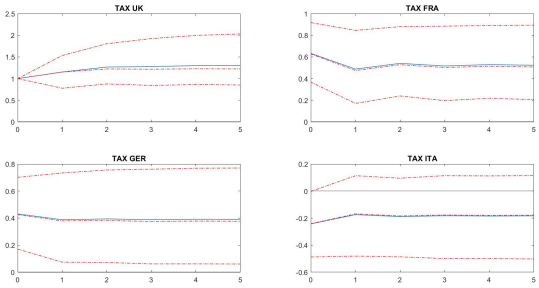
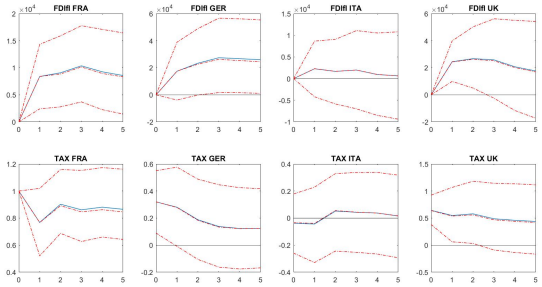
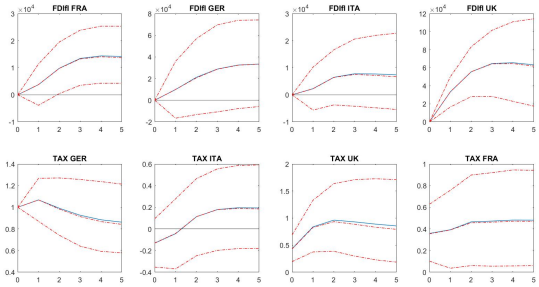


Figure 13: IRFs of BVAR model for Tax Rates and FDI Inflows

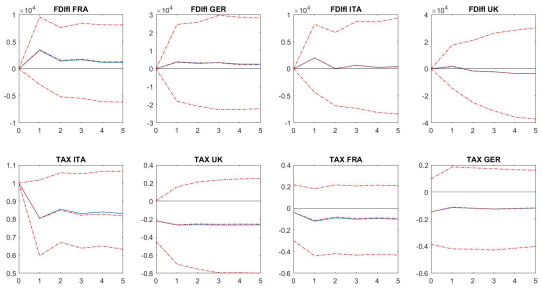
(a) Shock to the French tax rate



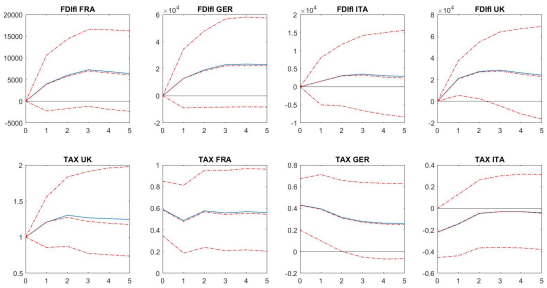
(b) Shock to the German tax rate



(c) Shock to the Italian tax rate



(d) Shock to the UK tax rate



5 Monte Carlo Experiment

In this section, we set up a Monte Carlo simulation exercise to evaluate the effects of estimating a panel model on data generated by a VAR process.

The previous empirical exercise has shown that the identification strategy plays a crucial role to study responses to exogenous shocks, while the effects of constraints over reduced-form coefficients are more muted and may depend on data characteristics. Given these insights, we will evaluate the performance of panel and VAR models identified with a Cholesky decomposition of the errors' variance-covariance matrix under two different dimensions. In particular, we will look at the estimated matrices of reduced-form coefficients and the impulse response functions.

In more details, we set up two different Monte Carlo exercises. In the first one, we assume that the coefficients are the same across equations, as implicit in the panel setup. The variance-covariance matrix of errors is randomly generated, with the only constraint of being positive semidefinite and symmetric. In the second one, we generate random coefficients so that they are different in each equation, and retain the same errors' variance-covariance matrix of the previous exercise. Subsequently, in both cases, we use these matrices to generate 100 series from a VAR process and use these artificial datasets to estimate a panel and VAR model. By considering the results, we will be able to evaluate what the error induced by estimating a panel model is, when the true data generating process is represented by a VAR model. Finally, the length of the simulated series will be in a first case 45, and in a second case 300. The reason for considering different length series is that VAR models can easily suffer from an overfitting problem, so that estimates will be influenced by low degrees of freedom. This may be the case when the time-series dimension is 45 observations, as in the previous sections. On the contrary, when the time series features 300 observations, we will be able to unveil the actual benefit of using a VAR rather than a panel model.

For the first exercise, we use the following matrix of reduced-form coefficients:

$$A_1 = \begin{bmatrix} 0.2 & 0.05 & 0.05 & 0.05 \\ 0.05 & 0.2 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.2 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.2 \end{bmatrix}$$

Here, the coefficients of every equation are listed by column and rows represent the lags of the variables. We therefore assume that each variable has an autoregressive coefficient of 0.2, while its sensitivity to the lags of other variables is 0.05.

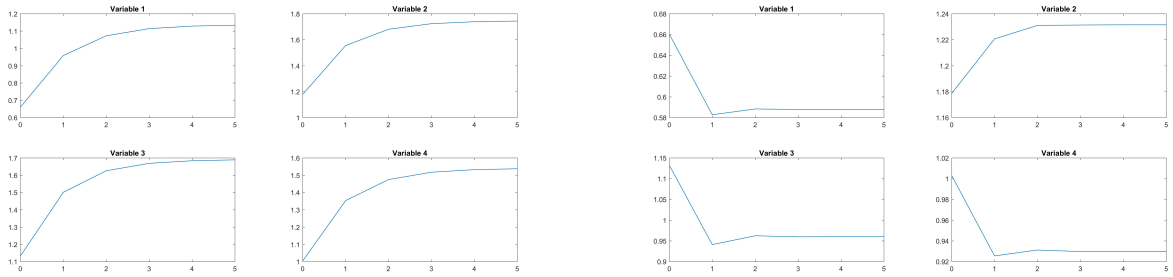
In regards constant terms, we assume the following:

$$c = \begin{bmatrix} 0.1 & -0.45 & -0.3 & 0.8 \end{bmatrix}$$

For simplicity, we define $D_1 \equiv \left[c', A_1'\right]'$.

The variance-covariance matrix of the errors is generated by randomly drawing a matrix S from a uniform distribution and then imposing $\Sigma = S \cdot S'$, so that Σ is positive semidefinite and symmetric. This allows us to obtain the following matrix:

Figure 14: IRFs implied by artificial matrices A_1 or A_2 and Σ



(a) IRFs with A_1 and Σ

(b) IRFs with A_2 and Σ

$$\Sigma = \begin{bmatrix} 0.436 & 0.778 & 0.747 & 0.662 \\ 0.778 & 1.901 & 1.522 & 1.189 \\ 0.747 & 1.522 & 1.649 & 1.086 \\ 0.662 & 1.189 & 1.086 & 1.034 \end{bmatrix}$$

In the second exercise, we randomly generate a matrix of contemporaneous coefficients by drawing from a uniform distribution between 0 and 1 and re-scaling values between -1 and 1 so that we obtain:

$$A_2 = \begin{bmatrix} -0.0328 & 0.0388 & -0.0794 & -0.0892 \\ -0.0069 & 0.0914 & -0.0312 & -0.0552 \\ 0.0035 & -0.0244 & -0.0836 & -0.045 \\ -0.0515 & -0.0635 & -0.0072 & 0.0969 \end{bmatrix}$$

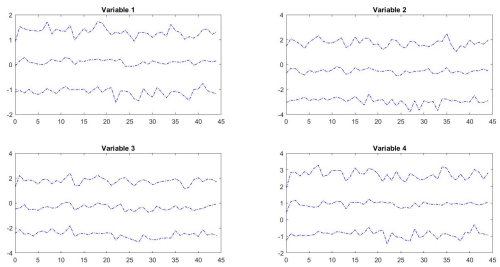
As for the errors' variance-covariance matrix and the constant terms, we retain Σ and c from the first Monte Carlo exercise. This allows us to show that the results will only be influenced by the different hypotheses over the reduced-form coefficients on the lagged endogenous variables. As before, we define $D_2 \equiv \left[c', A_2'\right]'$.

The impulse response functions implied by these matrices are displayed in figure 14. The responses are cumulated to be consistent with the empirical exercise.

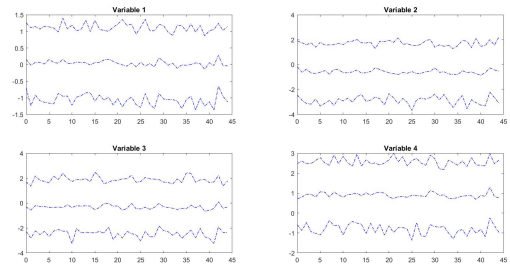
These matrices are used to simulate 100 series for each variable in our dataset from a simple VAR model with 1 lag and no exogenous regressors. In our empirical analysis, the time series have 45 observations so that we also start by simulating 45 observation. Then, for the reasons explained above, we also simulate a dataset with 300 observations for each series. The first observation is simulated from its unconditional distribution, while the following ones are simulated from their conditional distributions. Figure 15 shows the 5th, 50th and 95th percentiles of the distribution of the four artificial datasets.

We use these series to estimate a panel model with the same structure of equation 14, but no exogenous regressors and a VAR model with 1 lag and no exogenous regressors. In both cases, we identify the structural relations with a Cholesky decomposition of the variance-covariance matrix of the errors and we compute the implied impulse response functions to a shock of one standard deviation in the first variable.

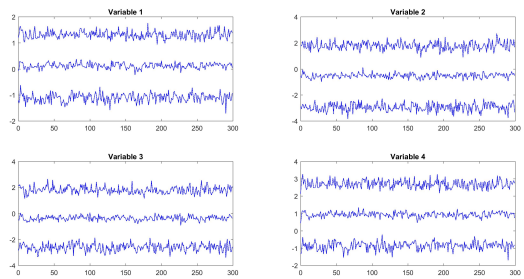
Figure 15: Artificial datasets



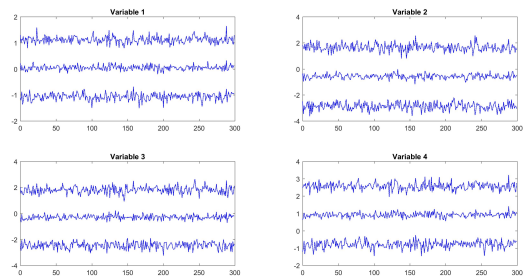
(a) A_1 and Σ with $T = 45$



(b) A_2 and Σ with $T = 45$



(c) A_1 and Σ with $T = 300$



(d) A_2 and Σ with $T = 300$

Here, we present the results and evaluate the performance of the panel and the VAR model by comparing the estimated reduced-form coefficients and the impulse response functions with the true ones presented above.

Let us start with the case with series of 45 observations.

Figures 16 and 17 in appendix A show the cumulated impulse response functions and their mean, the 5th and 95th percentiles of the distribution. These distributions come from the Monte Carlo simulations, i.e., they represent impulse response functions of different reduced-form parameters estimated from the artificial datasets. Not surprisingly, the responses from the panel and the VAR model are almost identical to the true ones in figure 14, when the dataset is generated by A_1 meaning that no distortion is introduced if we estimate a panel model in this case. However, when coefficients are heterogeneous, the responses of variables 2 and 3 are slightly different in the panel and VAR setup. Those from the VAR model seem to be more consistent with the true ones in figure 14. In this case the VAR model yields a better approximation of its system dynamics.

Furthermore, we compute the average matrices of reduced-form coefficients and error covariances across the Monte Carlo simulations (displayed in appendix A). Here, the subscript P and V specify the matrices from the panel and VAR model respectively, while 45 is the time-series dimension of the dataset.

In order to gauge to what extent the impulse response functions and parameter matrices of parameters differ from the true ones, we compute the following measure to capture the distance between two matrices:

$$M = \sqrt{\sum_i \sum_j (d_{ij,true} - d_{ij,est})^2}, \tag{21}$$

Table 3: Distance between estimated and true parameters and IRFs in the panel and VAR models when $T = 45$

	D_1	D_2	Σ_{D_1}	Σ_{D_2}	IRFs$_{D1}$	IRFs$_{D_1}$cum	IRFs$_{D2}$	IRFs$_{D_2}$cum
Panel	0.3146	0.274	0.0782	0.2751	0.2286	0.6711	0.1753	0.3585
VAR	0.5261	0.4586	0.0626	0.1711	0.2302	0.7258	0.047	0.0429

where d_{ij} is a generic element of the considered matrix. This indicates that we compute the square of the difference between each element of the true matrices compared to the estimated one, sum these elements and take their square root.

The results for the exercise with 45 observations are summarized in table 3. The values are derived comparing the true matrices D_1, D_2 and Σ with the estimated ones $\hat{D}_1$, $\hat{D}_2$ and $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$. As for the impulse response functions, we consider both the simple and cumulated impulse response functions, even if the results are of course consistent across these two transformations. In particular, we compare the true responses displayed in figure 14 with the average ones from the Monte Carlo exercise, which are the thick blue lines in figure 16 and 17. The coefficient matrices turn out to be more correctly estimated by the panel model while the opposite is true for the variance-covariance matrices. These discrepancies may be due to the fact that the VAR model needs to estimate a total of 20 coefficients with only 45 time series observations, so that estimates are more biased in the VAR setup. Nevertheless, when considering impulse response functions, the bias for the two models is similar if the true model has coefficients that are the same through equations while the VAR outperform the panel model to a large extent when coefficients are heterogeneous. These results are most probably affected by the better estimates of the variance-covariance matrix, which is crucial in determining the dynamics of the system after an exogenous shock. Furthermore, this is in line with what one would expect and with the graphical analysis of impulse response functions. Overall, we can conclude that if the focus of the analysis are structural relationships and reactions to exogenous shocks, the VAR model should be preferred to the panel model.

We then move on to analyse the results when the series have 300 observations. This exercise allow us to compare the performance of the panel and VAR models in a setup with more degrees of freedom. This, in principle, should allow the VAR model to reduce the bias of the estimates of reduced-form coefficients.

As before, we use the artificial dataset to estimate a panel and a VAR model and figures 18 and 19 in appendix A show the impulse response functions by reporting and the mean, the 5th and 95th percentiles of their distributions. This graphical exercise leads to the same conclusions we drew from the analysis of the responses with $T = 45$.

The next step is to evaluate the ability of the panel and VAR models to estimate the model's parameters. In order to do so, we compute the average matrices of reduced-form coefficients and error covariances across the Monte Carlo simulations that can be found in appendix A.

We take these matrices and the average of the impulse responses displayed in figures 18 and 19 and compare them with the true quantities by computing the measure described in equation 21. Table 4 shows the results that are similar to those in table 3. The VAR model outperforms the panel model in estimating the variance-covariance matrices and the impulse response functions when the dataset is generated using D_2. Contrary to the previous exercise, when $T = 300$, the bias of the VAR estimates of the coefficient

Table 4: Distance between estimated and true parameters and IRFs in the panel and VAR models when $T = 300$

	D_1	D_2	Σ_{D_1}	Σ_{D_2}	IRFs$_{D_1}$	IRFs$_{D_1}$cum	IRFs$_{D_2}$	IRFs$_{D2}$cum
Panel	0.0196	0.2411	0.0166	0.0398	0.029	0.0828	0.1646	0.3534
VAR	0.0992	0.1016	0.0092	0.0193	0.0375	0.1228	0.0123	0.0194

matrix D_2 is smaller. Still, when the true coefficients are the same throughout equations, the panel model provides less biased estimates of the reduced-form coefficients. However, as before, the bias is greater in the D_1 estimate than in the implied impulse response functions. This table confirms that in this case, the ability of VAR to estimate contemporaneous relationships leads to impulse response functions that are practically indistinguishable from those from the panel setup.

Overall, thanks to this exercise, we can see that the VAR model always gives better estimates of errors' variance-covariance matrices. Consequently, if these matrices are used to identify contemporaneous structural relationships, VAR estimates give an advantage with respect to panel ones. Considering responses to exogenous shocks, which are a convolution of structural and reduced-form coefficients, VAR impulse responses are in the worst cases, indistinguishable from those from the panel model, when the true data generating process has coefficients that are the same through equations. However, when the reduced-form coefficients are equation-specific, the VAR outperform the panel model to a large extent in estimating responses. Therefore, if the objects of interest are the structural relationships, one should always opt for a VAR setup, regardless of the true structure of the economy. However, these differences are less pronounced the fewer data observations are available. On the contrary, if the interest of the researcher are the reduced-form coefficients, the panel model outperforms the VAR, when the time span is limited as it exploits the cross-sectional dimension of the dataset to estimate fewer parameters. As a consequence, the use of panel models may still be recommended in case of datasets with few observations and when the researcher strongly believes that the structure of the economy is similar across sample countries.

6 Conclusion

This work has dealt with empirical approaches used to study tax reaction functions. In doing so, we have departed from the standard empirical literature that only uses panel models, where contemporaneous relations are estimated with instrumental variables and coefficients are assumed to be the same across countries. As we have shown, such an approach may lead to relevant biases.

Two issues have been raised regarding the typical panel models used to study tax competition. Firstly, a researcher should identify exogenous variations in fiscal variables. The literature applies instrumental variables and relies on the twofold assumption that each country reacts to a weighted average of other countries' tax rates, where weights are exogenous. Secondly, panel models are useful to exploit the cross-sectional dimension of data, so several coefficients are constrained to be the same across countries. Typically, this structure implies that tax competition, if any, is assumed to be homogeneous across countries. As we have shown, these restrictions are too strong.

In order to assess the robustness of the empirical evidence on international tax competition, we have

applied a VAR approach, which is widely used in macroeconomic analysis and disregarded by the tax competition literature. The advantage of VAR models is that reduced-form coefficients are equation-specific and, thanks to the identification of structural relations, they allow us to estimate impulse response functions capturing how variables react to exogenous shocks. The VAR approach can therefore be used to obtain more reliable estimates of tax reaction functions. In particular, we identified structural parameters with a Cholesky decomposition of the errors' variance-covariance matrix, which amounts to imposing zero restrictions on its upper extra-diagonal elements. From an economic point of view, this implies that the ordering of the variables affects their contemporaneous correlations. Our modelling strategy allows each tax rate to react simultaneously to exogenous shocks in other countries' rates. We are therefore in the position to verify whether countries play a Nash or Stackelberg game.

As an initial step, we have compared the panel and the VAR model from a theoretical point of view. In doing so, we have also shown that assumptions over exogenous weights can influence impulse response functions. The second step of our study has taken our models to the data by considering four European countries and exploring how impulse response functions change in different frameworks.

In our methodological part, we have considered panel models that embed various constraints on the coefficients of their autoregressive terms or on the lagged tax rates of other countries and where contemporaneous relations are identified with a Cholesky decomposition. The qualitative interpretation of our results is as follows: here, tax reactions functions are sometimes positive and sometime negative. This means that tax competition, when present, is heterogeneous. Furthermore and more interestingly, we have found that results are consistent between two countries: in other words, if the tax rate of country A is a strategic complement (substitute) for the tax rate of country B, the converse is also true. In regards to the type of competition, all significant responses are contemporaneous to the shock, so that it is possible to conclude that our results support the hypothesis that countries play a Nash game. As for the methodological implications, results are similar across models. This means that, in our application, such constraints play a limited role in correctly identifying reaction functions. However, this may not be so when dealing with larger systems where the country-specific dynamics play a more relevant role.

To reconcile our approach with the literature, we have also studied the effects of using an identification strategy based on zero short-run restrictions instead of applying instrumental variables. Here, we have considered a panel model where contemporaneous relations are identified with instrumental variables and then compared its matrix of contemporaneous parameters with the one from a panel model identified with the Cholesky decomposition. This comparison has confirmed that the two identification strategy lead to contrasting results. Namely, with instrumental variables, tax rates are all strategic complements while under our approach, we found that certain contemporaneous coefficients are positive and others are negative.

Subsequently, we have estimated VAR and BVAR models and found that impulse response functions are again different from the panel models identified in the same way. This highlights the fact that constraints on the exogenous regressors are also relevant to correctly identify structural relations, so that homogeneity assumptions should be applied only when strongly supported by theoretical and empirical evidence. Overall, we can state that the choice between a panel and a VAR model is a relevant one.

The set of dependent variables of the VAR and BVAR models may also be expanded to include other

variables that are supposed to be endogenous to tax rates. We have also tested the effect of including FDI inflows as dependent variables. Once again, the impulse response functions are affected by this strategy, even if only marginally. However, these results cast doubt on the effectiveness of tax competition as a policy to attract FDIs, as we found that FDI inflows increase after a rise in the tax rate. In order to confirm this finding, it would be necessary to try to extend the model to properly account for all forces shaping FDI inflows. Overall, our exercise has shown how VAR can be applied to study tax competition and in particular, Bayesian VAR is the the best tool when dealing with systems characterised by many endogenous variables.

Given the crucial role of the identification strategy to model responses to exogenous shocks, we have also analysed the performance of panel and VAR models identified with the Cholesky decomposition of the errors' variance-covariance matrix when the true data generating process is a VAR model. We have used a Monte Carlo exercise, where two dataset are simulated assuming that the reduced-form coefficients are either the same or different across equations. We have then evaluated the ability of the panel and VAR model to estimate both the parameters and the responses to exogenous shocks. We have found that the VAR always outperform the panel model in estimating structural contemporaneous coefficients and it leads to estimates of the impulse response functions that are at least as good as those from the panel model under the least favorable scenario. However, these differences are less pronounced if the number of observations is smaller. Furthermore, the panel model tends to give less biased estimates of the reduced-form coefficients when the time span is limited because it exploits the cross-sectional dimension of the dataset and it estimates fewer parameters. As a consequence, although the structural contemporaneous coefficients are less biased if estimated with a VAR, impulse response functions from a panel model may still be reliable and useful if few observations are available and the structure of the economy is similar across the sample.

To su up, this paper has proposed an alternative tool to study tax competition. As shown, the panel approach is characterized by relatively restrictive assumptions and can lead to unreliable estimates. In turn, VAR models coupled with an identification based on zero short-run restrictions are more flexible and look helpful to study tax competition. A crucial role is played by the identification strategy, so that coupling panel models with an identification based on a Cholesky decomposition of the errors' variance-covariance matrix is recommended to exploit the cross-sectional dimension of data without imposing unrealistic assumptions over structural coefficients.

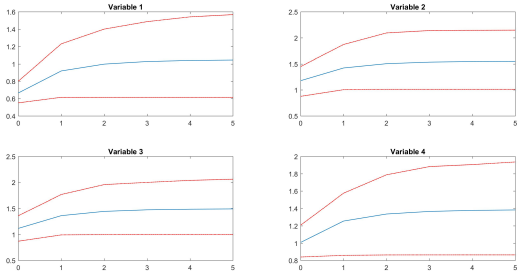
Appendix 1. Monte Carlo Exercise

A Impulse response functions

This section showss the impulse response functions of the Monte Carlo exercise. The responses are generated by different datasets and the figures display the mean, the 5th and 95th percentiles of distributions for each point in time. The distributions refers to all impulse responses derived from 100 series in each dataset.

Figure 16: IRFs of artificial dataset generated by A_1 and Σ when $T = 45$

(a) IRFs from the panel model



(b) IRFs from the VAR model

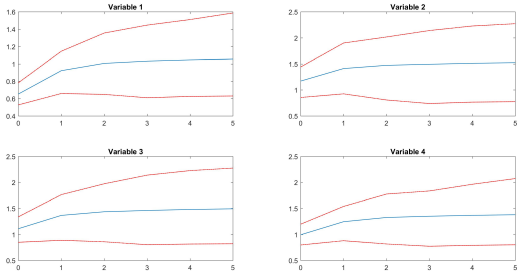
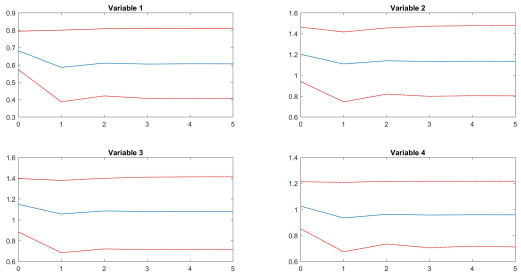


Figure 17: IRFs of artificial dataset generated by A_2 and Σ when $T = 45$

(a) IRFs from the panel model



(b) IRFs from the VAR model

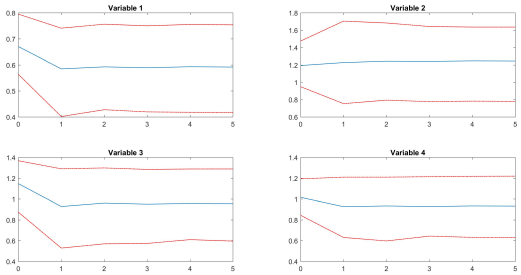
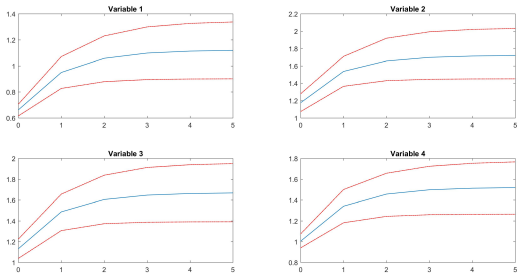


Figure 18: IRFs of artificial dataset generated by A_1 and Σ for $T = 300$

(a) IRFs from the panel model



(b) IRFs from the VAR model

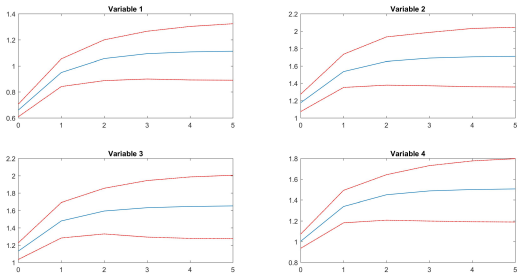
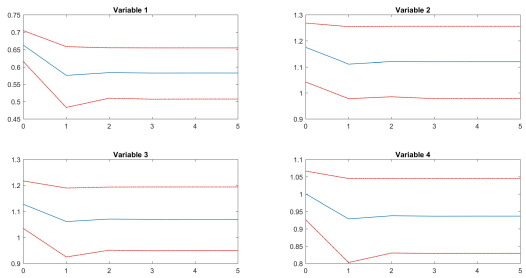
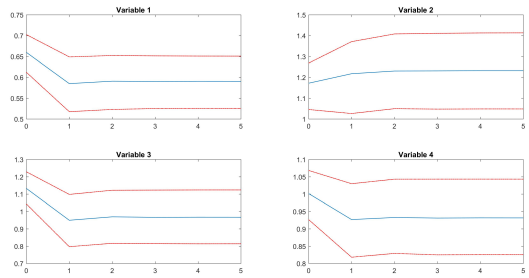


Figure 19: IRFs of artificial dataset generated by A_2 and Σ for $T = 300$

(a) IRFs from the panel model



(b) IRFs from the VAR model



B Estimated matrices

This section reports the average matrices of reduced-form coefficients and errors' variance-covariances computed across the Monte Carlo simulations.

When the time-series dimension of the dataset is 45 observation, the matrices are as follows:

$$\hat{D}_{1,P45} = \begin{bmatrix} 0.0884 & -0.4912 & -0.3269 & 0.7877 \\ 0.0475 & 0.0663 & 0.0663 & 0.0663 \\ 0.0663 & 0.0475 & 0.0663 & 0.0663 \\ 0.0663 & 0.06625 & 0.0475 & 0.0663 \\ 0.0663 & 0.06625 & 0.0663 & 0.0475 \end{bmatrix} \quad \hat{\Sigma}_{D_1,P45} = \begin{bmatrix} 0.4483 & 0.7933 & 0.752 & 0.679 \\ 0.7933 & 1.9490 & 1.5327 & 1.2153 \\ 0.7520 & 1.5327 & 1.6560 & 1.0942 \\ 0.6790 & 1.2153 & 1.0942 & 1.0616 \end{bmatrix}$$

$$\hat{D}_{1,V45} = \begin{bmatrix} 0.0402 & -0.6214 & -0.4647 & 0.7206 \\ 0.0853 & -0.1347 & -0.0914 & 0.0507 \\ 0.0578 & 0.03814 & 0.05344 & 0.0597 \\ 0.0327 & 0.03421 & 0.00914 & 0.0252 \\ 0.106 & 0.2336 & 0.2378 & 0.1158 \end{bmatrix} \quad \hat{\Sigma}_{D_1,V45} = \begin{bmatrix} 0.4315 & 0.7734 & 0.7341 & 0.6571 \\ 0.7734 & 1.8958 & 1.4946 & 1.1832 \\ 0.7341 & 1.4946 & 1.6147 & 1.0684 \\ 0.6571 & 1.1832 & 1.0684 & 1.0272 \end{bmatrix}$$

$$\hat{D}_{2,P45} = \begin{bmatrix} 0.0522 & -0.5614 & -0.2856 & 0.9213 \\ -0.0165 & -0.025 & -0.025 & -0.025 \\ -0.025 & -0.0165 & -0.025 & -0.025 \\ -0.025 & -0.025 & -0.0165 & -0.025 \\ -0.025 & -0.025 & -0.025 & -0.0165 \end{bmatrix} \quad \hat{\Sigma}_{D_2,P45} = \begin{bmatrix} 0.4681 & 0.8237 & 0.7901 & 0.7053 \\ 0.8237 & 2.026 & 1.5948 & 1.2513 \\ 0.7901 & 1.5948 & 1.7691 & 1.1568 \\ 0.7053 & 1.2513 & 1.1568 & 1.1029 \end{bmatrix}$$

$$\hat{D}_{2,V45} = \begin{bmatrix} 0.0629 & -0.6402 & -0.532 & 0.7876 \\ 0.0125 & -0.0309 & -0.2099 & 0.0571 \\ -0.0433 & -0.0097 & -0.1123 & -0.1121 \\ 0.0031 & 0.0031 & -0.098 & -0.046 \\ -0.0425 & 0.0713 & 0.171 & 0.0615 \end{bmatrix} \quad \hat{\Sigma}_{D_2,V45} = \begin{bmatrix} 0.454 & 0.8068 & 0.777 & 0,688 \\ 0.8068 & 1.9691 & 1.5864 & 1.2287 \\ 0.777 & 1.5864 & 1.7053 & 1.1285 \\ 0.688 & 1.2287 & 1.1285 & 1.0706 \end{bmatrix}$$

When the time-series dimension of the dataset is 300 observation, the matrices are as follows:

$$\hat{D}_{1,P300} = \begin{bmatrix} 0.1003 & -0.448 & -0.3033 & 0.8058 \\ 0.1913 & 0.0483 & 0.0483 & 0.0483 \\ 0.0483 & 0.1913 & 0.0483 & 0.0483 \\ 0.0483 & 0.0483 & 0.1913 & 0.0483 \\ 0.0483 & 0.0483 & 0.0483 & 0.1913 \end{bmatrix} \quad \hat{\Sigma}_{D_1,P300} = \begin{bmatrix} 0.4394 & 0.7799 & 0.7509 & 0.6666 \\ 0.7799 & 1.9 & 1.5258 & 1.1909 \\ 0.7509 & 1.5258 & 1.6576 & 1.0907 \\ 0.6666 & 1.1909 & 1.0907 & 1.0405 \end{bmatrix}$$

$$\hat{D}_{1,V300} = \begin{bmatrix} 0.0922 & -0.4722 & -0.2967 & 0.7828 \\ 0.1685 & -0.0003 & 0.0327 & -0.0068 \\ 0.0522 & 0.1842 & 0.0514 & 0.0505 \\ 0.0474 & 0.0566 & 0.1946 & 0.0488 \\ 0.062 & 0.0789 & 0.0457 & 0.2242 \end{bmatrix} \quad \hat{\Sigma}_{D_1,V300} = \begin{bmatrix} 0.4374 & 0.7778 & 0.7488 & 0.6642 \\ 0.7778 & 1.8942 & 1.5219 & 1.1875 \\ 0.7488 & 1.5219 & 1.6527 & 1.0875 \\ 0.6642 & 1.1875 & 1.0875 & 1.0367 \end{bmatrix}$$

$$\hat{D}_{2,P300} = \begin{bmatrix} 0.0581 & -0.5171 & -0.2513 & 0.8972 \\ 0.0137 & -0.0289 & -0.0289 & -0.0289 \\ -0.0289 & 0.0137 & -0.0289 & -0.0289 \\ -0.0289 & -0.0289 & 0.0137 & -0.0289 \\ -0.0289 & -0.0289 & -0.0289 & 0.0137 \end{bmatrix} \quad \hat{\Sigma}_{D_2,P300} = \begin{bmatrix} 0.4394 & 0.7796 & 0.7482 & 0.6639 \\ 0.7796 & 1.9078 & 1.5081 & 1.185 \\ 0.7482 & 1.5081 & 1.6807 & 1.0936 \\ 0.6639 & 1.185 & 1.0936 & 1.0381 \end{bmatrix}$$

$$\hat{D}_{2,V300} = \begin{bmatrix} 0.1065 & -0.4517 & -0.2983 & 0.8207 \\ -0.0043 & 0.0577 & -0.0187 & -0.0296 \\ -0.0138 & 0.0821 & -0.0489 & -0.0645 \\ 0.0043 & -0.0245 & -0.0856 & -0.0452 \\ -0.0611 & -0.0624 & -0.0184 & 0.0707 \end{bmatrix} \quad \hat{\Sigma}_{D_2,V300} = \begin{bmatrix} 0.4360 & 0.7745 & 0.7484 & 0.6619 \\ 0.7745 & 1.8863 & 1.5177 & 1.1827 \\ 0.7484 & 1.5177 & 1.6541 & 1.0869 \\ 0.6619 & 1.1827 & 1.0869 & 1.0326 \end{bmatrix}$$

References

[1] Altshuler R., and T. Goodspeed (2002), "Follow the Leader? Evidence on European and U.S. Tax Competition", Rutgers University, Departmental Working Papers 200226.

[2] Altshuler R., and T. Goodspeed (2015), "Follow the Leader? Evidence on European and US Tax Competition", *Public Finance Review*, 43, pp. 485-504.

[3] Banbura M., D. Giannone, and L. Reichlin (2010), "Large Bayesian Vector Auto Regressions", *Journal of Applied Econometrics*, 25, pp. 71-92.

[4] Baskaran T., and M. Lopes da Fonseca (2014), "The Economics and Empirics of Tax Competition, A Survey and Lessons for the EU", Erasmus Law Review, n.1.

[5] Besley T., R. Griffith, and A. Klemm (2001), "Empirical Evidence on Fiscal Interdependence in OECD countries", World Tax Competition Conference, London, UK.

[6] Bretschger L., and F. Hettich (2002), "Globalisation, Capital Mobility and Tax Competition: Theory and Evidence for OECD Countries", *European Journal of Political Economy*, 18, pp. 695-716.

[7] Brett C., and J. Pinkse (2000), "The Determinants of Municipal Tax Rates in British Columbia", *Canadian Journal of Economics*, 33, pp. 695-714.

[8] Bucovetsky S. (1991), "Asymmetric Tax Competition", *Journal of Urban Economics*, 30, pp. 167-181.

[9] Canova F., Ciccarelli M. (2013), "Panel Vector Autoregressive Models. A Survey.", Working Paper n. 1507, European Central Bank.

[10] Chirinko R.S., and D.J. Wilson (2018), "Tax Competition Among U.S. States: Racing to the Bottom or Riding on a Seesaw?", *Journal of Public Economics*, forthcoming.

[11] Davies R.B., and J. Voget (2011), "Tax Competition in an Expanding European Union", GEE Papers n. 33.

[12] de Mooij R., and S. Ederveen (2003), "Taxation and Foreign Direct Investment: A Synthesis of Empirical Research", *International Tax and Public Finance*, 10, pp. 673-693.

[13] Devereux M., and S. Loretz (2013), "What Do We Know about Corporate Tax Competition?", *National Tax Journal*, 66, pp. 745-73.

[14] Devereux, M.P., R. Griffith, and A. Klemm (2002), "Corporate Income Tax Reforms and International Tax Competition", *Economic Policy: A European Forum*, 17, pp. 450-495.

[15] Devereux, M.P., B. Lockwood, and M. Redoano (2008), "Do Countries Compete over Corporate Tax Rates?", *Journal of Public Economics*, 92, pp. 1210-1235.

[16] Egger P., and H. Raff (2015), "Tax Rate and Tax Base Competition for Foreign Direct Investment", *International Tax and Public Finance*, 22, pp. 777-810.

[17] Eugster B., and R. Parchet (2011), "Culture and Taxes: Towards Identifying Tax Competition", Cahiers de Recherches Economiques du Departement d'Econometrie et d'Economie Politique (DEEP) 11.05.

[18] Fuest C., B. Huber, and J. Mintz (2005), "Capital Mobility and Tax Competition", *Foundations and Trends in Microeconomics*, 1, pp. 1-62.

[19] Garbarino C., and P.M. Panteghini (2007), "Corporate Taxation in Europe: Competitive Pressure and Cooperative Targets", in Gregoriou, G.N. and C. Read (eds.), *International Taxation Handbook*, Elsevier.

[20] Ghinamo M., P.M. Panteghini, and F. Revelli (2010), "FDI Determination and Corporate Tax Competition in a Volatile World", *International Tax and Public Finance*, 17, pp. 532-555.

[21] Giannone D., M. Lenza, and G. Primiceri (2015), "Prior Selection for Vector Autoregressions", *The Review of Economics and Statistics*, 2, pp. 436-451,

[22] Gibbons S., and H. Overman (2012), "Mostly Pointless Spatial Econometrics?", *Journal of Regional Science*, 52, pp. 172-191

[23] Gordon R. (1992), "Can Capital Income Taxes Survive in Open Economics?", *Journal of Finance*, 47, pp. 1159-1180.

[24] Haufler A., and I. Wooton (1999), "Country Size and Tax Competition for Foreign Direct Investment", *Journal of Public Economics*, 71, pp. 121-139.

[25] Heineman F., M. Overesch, J. Rincke (2010), "Rate-Cutting Tax Reforms and Corporate Tax Competition in Europe", *Economics & Politics*, 22, pp. 498-518.

[26] Hory M.-P. (2017), "Delayed Mimicking: the Timing of Fiscal Interactions in Europe", *European Journal of Political Economy*, forthcoming.

[27] Im K.S., M.H. Pesaran, and Y. Shin (2003), "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, 115, pp. 53-74.

[28] Kanbur R., and M.J. Keen (1993), "Jeux Sans Frontieres: Tax Competition and Tax Coordination When Countries Differ in Size.", *American Economic Review*, 83, pp. 877-892.

[29] Keen M., and K. Konrad (2013), "The Theory of International Tax Competition and Coordination.", in *Handbook of Public Economics*, 5, pp. 259-330, edited by Alan Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, Elsevier.

[30] Leibrecht M., and C. Hochgatterer (2012), "Tax Competition As a Cause of Falling Corporate Income Tax Rate, A Survey of Empirical Literature", *Journal of Economic Surveys*, 26, pp. 616–648.

[31] Levin A., C.-F. Lin, and C.-S.J. Chu (2002), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties.", *Journal of Econometrics*, 108, pp. 1-24.

[32] Mintz J., and H. Tulkens (1986), "Commodity Tax Competition Between Member States of a Federation: Equilibrium and Efficiency.", *Journal of Public Economics*, 29, pp. 133-172.

[33] Overesch M., and J. Rincke (2011), "What Drives Corporate Tax Rates Down? A Reassessment of Globalization, Tax Competition, and Dynamic Adjustment to Shocks", *Scandinavian Journal of Economics*, 113, pp. 579-602.

[34] Panteghini P.M., and G. Schjelderup (2006), "To Invest or not to Invest: A Real Options Approach to FDIs and Tax Competition", *International Tax and Public Finance*, 13, pp. 643-660.

[35] Redoano M. (2014), "Tax Competition among European Countries: Does the EU Matter?", *European Journal of Political Economy*, 34, pp. 353-371.

[36] Revelli F. (2005), "On Spatial Public Finance Empirics", *International Tax and Public Finance*, 12, pp. 475-492,

[37] Sims C. (1980), "Macroeconomics and Reality", *Econometrica*, 48, pp.1-48.

[38] Slemrod J. (2004), "Are Corporate Tax Rates or Countries Converging?", *Journal of Public Economics*, 88, pp. 1169-1186.

[39] Stöwhase S. (2005), "Asymmetric Capital Tax Competition with Profit Shifting", *Journal of Economics*, 85, pp. 175-196.

[40] Vrijburg H., and R.A. de Mooij (2016), "Tax rates as strategic substitutes," *International Tax and Public Finance*, 23, pp. 2-24.

[41] Westerlund J. (2007), "Testing for Error Correction in Panel Data", *Oxford Bulletin of Economics and Statistics*, 69, pp. 709-748.

[42] Wildasin D. E. (1988), "Nash Equilibria in Models of Fiscal Competition.", *Journal of Public Economics*, 35, pp. 229-240.

[43] Wilson J.D. (1986): "A Theory of Interregional Tax Competition", *Journal of Urban Economics*, 19, pp. 296-315.

[44] Wilson J.D. (1991), "Tax Competition with Interregional Differences in Factor Endowments", *Regional Science and Urban Economics*, 21, pp. 423-451.

[45] Wilson J., and D. Wildasin (2004), "Capital tax competition: bane or boon", *Journal of Public Economics*, 88, pp. 1065-1091.

[46] Zodrow G.R., and P. Mieszkowski (1986), "Pigou, Tiebout, Property Taxation, and the Underprovision of Local Public Goods.", *Journal of Urban Economics*, 19, pp. 356-370.