Economic Issues, Vol. 16, Part 1, 2011

The Tax-Spend Debate with an Application to the EU

G A Vamvoukas[1]

ABSTRACT

The important contribution of this paper is to shed light on the validity of the tax-spend debate using panel data from 15 selected EU countries. The controversial tax and spend, spend and tax, fiscal synchronisation and institutional separation hypotheses compose the four theories in the field that have been formulated to describe the relationship between government expenditures and tax revenues. The goal of the present study is achieved through three steps. In the first step, the entire period 1970-2007 is split into two time frames, 1970-1991 and 1992-2007, in order to take into account the pre and post-Maastricht subperiods. In the second step, two-variable and three-variable panel models are estimated employing the TSLS and GMM techniques. In the third step, we perform a sensitivity analysis by conducting short-and long-run Granger causality tests to check the robustness of TSLS and GMM results. Our empirical findings in the case of EU countries are consistent with the theoretical framework of the fiscal synchronisation hypothesis.

1. INTRODUCTION

EXPENDITURES AND REVENUES represent two components of the government budget. The pattern of relationship between public expenditures and public revenues has been an interesting issue in economic research. Four hypotheses concerning this relationship between government spending and tax receipts have been analysed in the literature. First, the tax-spend hypothesis argues that changes in public outlays lead to changes in tax revenues. Second, the spend-tax hypothesis indicates that expenditures adjust to changes in revenues. Third, the fiscal synchronisation hypothesis suggests that government spending and revenue decisions are taken jointly. Fourth, the institutional separation hypothesis states that revenues and expenditures are independent of one another.

The purpose of this paper is to employ panel data from the European Union (EU) to investigate the validity of the four hypotheses. The use of panel data is very important in the case of the EU, as most countries which compose the EU have similar institutional features and share common goals in the application of fiscal policy. The EU member states constitute an interesting case because the fiscal discipline requirement is a primary target of economic policy. Given that a considerable number of EU countries are confronted with macroeconomic imbalances because of large government deficits, the relationship between public expenditures and tax revenues represents an important issue of empirical analysis in order to understand budget deficit movements.

The causal nexus between public outlays and receipts has attracted various researchers in the area and has been the subject of inconclusive empirical work. The evidence on the tax-spend debate is mostly related to G7 countries, mainly the USA, where time series data are readily available. The essential contribution of the present paper is that the four hypotheses are tested by employing panel data from EU countries. To our knowledge, no empirical study has used panel data techniques to investigate the pattern of relationship between government spending and tax receipts. In most cases the empirical studies on the tax-spend controversy have used time series techniques based on cointegration procedures, error correction modelling (ECM) and vector autoregression (VAR) analysis. Depending on the cointegrating properties between the two budgetary variables, a large part of the empirical literature has focused on their pattern of Granger causality.[2]

The rest of the paper is structured in the following way. The next Section discusses briefly the theoretical background of the four hypotheses, along with a review of the empirical literature. Section 3 presents the methodology used in the econometric analysis. Section 4 describes the data and reports TSLS and GMM results. Robustness checks based on short- and long-run Granger causality tests are the subject of the fifth Section. Section 6 draws some final conclusions.

2. A BRIEF SURVEY OF THE CONTROVERSY

Since the Maastricht Treaty in 1992 the convergence of fiscal policies has been a core element of the EU. The Maastricht convergence fiscal criteria focused on general government deficit and debt. Within the institutional framework of Economic and Monetary Union (EMU), there is the Excessive Deficit Procedure (EDP) which is described in the Maastricht Treaty. The EDP is clarified in more detail within the Stability and Growth Pact (SGP). According to the Treaty, achieving a general government deficit to GDP ratio of no more than 3 per cent is a necessary precondition for a country to participate in the final stage of EMU. Given the loss of the monetary instrument, the role of fiscal policy is central for the individual EMU country to achieve macroeconomic stability. Fiscal discipline is one of the principal goals of economic policy, indicating that the reduction of budget deficits depends on spending and tax changes.[3] Thus, the kind of relationship between public expenditures and tax receipts is a basic aspect of macroeconomic policy for fiscal authorities to attain budgetary balance over the medium term. The government, through its policy measures, can achieve this target by altering expenditures and revenues in the required direction.

Alternative hypotheses can be specified to investigate the relationship between public outlays and tax receipts. The spend-tax hypothesis argues that changes in expenditures lead to changes in taxes. This hypothesis is consistent with the rationale of the Keynesian model, in which fiscal policy on the sides of taxation and expenditure tries to manage aggregate demand and to achieve macroeconomic stability. The budget is considered as an impediment to fine-tuning the economy, namely, the continual increasing and decreasing of aggregate demand by fiscal policy in order to stabilise the economy. Albeit Keynesians support the idea of a balanced annual budget, they instead contend that government should run budget deficits during recession and surpluses only during expansion. Government spending has a key role in economic policy because it stimulates aggregate demand and contributes to macroeconomic growth and stability, by altering its level in the required direction. One of the prevailing views in the framework of the Keynesian model is that government expenditure determines macroeconomic activity and affects the level of taxation. In this way governments spend first and then increase taxes to finance these expenditures.

Peacock and Wiseman (1961, 1979) are the main proponents of the spend-tax hypothesis. They argue that temporary increases in public spending because of political and economic crises can cause permanent increases in tax revenue. Structural breaks such as wars, unstable political conditions, etc., may shift public outlays to higher levels causing tax increases.[4] Ricardian equivalence is in the spirit of the spend-tax hypothesis. Barro (1979), the leading proponent of Ricardian equivalence, argues that temporary increases in government expenditures today, financed by borrowing, will lead to permanent future tax increases for the public.

Supply-side economists are critical of the Keynesian emphasis on government spending.[5] The basic proposition of supply-side economics is that a substantial reduction of tax burdens, in general, and income tax rates in particular, will have significant effects on real GDP and the level of growth rates. Advocates of this school of economic thought formulated a new way of looking at government policy making, underlining that government economic policy should focus on aggregate supply rather than on aggregate demand. The general framework of supply-side tax policy is that the reduction of income tax rates would increase labour supply, savings and investment, bringing about sustained economic growth. Under the Laffer curve the reduction of tax rates would move the economy back from the prohibitive range, so as to contribute to an increase in tax revenues. Tax reductions contribute to an increase in real

GDP, giving the opportunity for government to collect more direct and indirect taxes. An efficient fiscal policy implies that tax cuts would actually increase government receipts rather than raise budget deficits. Taking into account that the primary goal of the supply-side fiscal policy is the lowering of marginal income tax rates, it is obvious that tax revenues lead government expenditures. In this way supply-side economics appears to accept the rationale of the tax-spend hypothesis. Friedman (1972, 1978), Buchanan and Wagner (1977, 1978) among others are advocates of the tax-spend theory. A balanced budget must be based on the principle that public revenues must determine the growth of public expenditures. The primary function of the budget is to ensure that the government increases the amount of revenues it considers necessary.

The fiscal synchronisation hypothesis, advanced mainly by Musgrave (1966), and Meltzer and Richard (1981), argues that government spending and revenue decisions are taken jointly. Policymakers compare the marginal benefits and marginal costs of fiscal policy with respect to the appropriate levels of outlays and receipts. Governments simultaneously determine the package of spending and revenue programmes so as to maximise the social welfare function. Governments, seeking macroeconomic stability and high growth rates, have to apply fiscal policy measures determining the budget programme, along with the taxes necessary to finance expenditures. The advocates of the fiscal synchronisation hypothesis do not assume that the government budget must be in balance every year. Whilst expenditures and revenues should move together, different targets of economic policy can bring the budget into transient surplus or deficit.

A fourth hypothesis, suggested by Wildavsky (1975) and Wilensky (1975) among others, argues that taxes and expenditures are independent of one another. This view, known in the literature as the fiscal separation hypothesis, underlines the institutional separation of the spending and taxation decision-making processes of government. Government's power to spend and levy taxes is a result of the legal framework that was previously established. The government, setting the institutional and legal framework within which markets operate, is the fundamental mechanism for determining the function of the economic system. Having decided the ceiling under which expenditure and revenue should be constrained, the government has to fulfill three criteria: allocative efficiency, distribution, and macroeconomic stability. Any economic system has to determine how to allocate economic resources, what goods and services to produce, how to produce them and to whom to distribute them. Considering that public outlays and tax receipts follow an independent path in the long run, the solution to these problems depends substantially on the application of fiscal policy measures.

All four hypotheses have been investigated empirically by various authors using data from developed and developing countries. The empirical studies on different countries have yielded inconclusive results. The sensitivity of results is related to the degree of temporal aggregation, the time period used in the empirical analysis, the kind of instruments, and the choice of econometric methodology. It should be noted that most papers have employed VAR analysis based mainly upon Granger causality techniques.[6] Miller and Russek (1990), employing annual and quarterly data from 1946 to 1987 for the US and using an ECM strategy, found that a two-way feedback relationship exists between taxes and spending. The evidence of bidirectional Granger causality between receipts and outlays is in line with the fiscal synchronisation hypothesis, indicating that government decides on the appropriate expenditure level and adjusts taxes to cover the cost.

Baffes and Shah (1994) examined Granger causality between government expenditures and revenues for Argentina, Brazil and Mexico, using annual data for 1913-1984, 1908-1985 and 1895-1984, respectively. Their results gave strong evidence of two-way Granger causality for Argentina and Mexico. For Brazil the type of Granger causality was from revenues to spending. Koren and Stiassny (1998) estimated structural VAR models to check the relation between government spending and taxation decisions in nine developed countries. In their empirical procedure, the VAR model for each country includes, besides outlays and revenues, GDP as an additional variable. The authors, interpreting the budget-making process as an ECM, provided evidence in accordance with the spend-tax hypothesis for Austria, France and Italy. The tax-spend theory appears to be true for Germany, the Netherlands, the UK and the US. Their results for Sweden and Switzerland tend to support the institutional separation hypothesis.

Owoye (1995) investigated the directions of causality between expenditures and taxes in G7 countries for the 1961-1990 period. This paper utilised ECM and Granger causality tests. The results indicated that bidirectional causality exists between taxes and expenditures in the USA, Germany, the UK, France and Canada, showing that the decisions to tax and spend or vice versa are made jointly. The evidence for Japan and Italy suggested that causality runs from taxes to expenditures, implying that higher public revenues lead to higher public outlays. Darrat (1998) examined the causal relationship between expenditure and revenue for Turkey over the period 1967-1994. The methodology was based on cointegration, ECM approach and Granger causality. In his empirical analysis he incorporated real GNP and the rate of interest as two additional variables as determining the behaviour of revenue and spending. The results provided strong support for the contention that tax revenues Granger-cause government expenditures in Turkey, both in the short-run and the long-run. Chang, Liu and Caudill (2002), employing time series techniques and using annual data over the period 1951-1996 for ten countries, found conflicting results regarding the four hypotheses. Their findings supported the spend-tax hypothesis for Australia and South Africa. For Japan, South Korea, Taiwan, the UK and the US, they provided evidence consistent with the tax-spend hypothesis. For Canada their results appear to be in line with the fiscal synchronisation hypothesis. Finally, for New Zealand and Thailand the results seem to support the institutional separation hypothesis.

Narayan and Narayan (2006) examined the pattern of causality between government spending (*G*) and government revenue (*T*) within a multivariate framework, by modelling *G* and *T* with *Y* (real GDP). Adopting the Toda and Yamamoto (1995) procedure to test for Granger causality, they found support for the tax-spend hypothesis in the case of Chile, El Salvador, Mauritius and Venezuela. Their results for Haiti indicated the acceptance of the fiscal synchronisation hypothesis. For Peru, South Africa, Guatemala, Uruguay and Ecuador, there is evidence of neutrality, suggesting no causality between *G* and *T*; thus the institutional separation hypothesis is accepted in these cases. Kollias and Paleologou (2006) investigated the long-term relationship between taxes and expenditures in fifteen European countries, using annual data in a vector error correction modelling (VECM) framework. The four hypotheses were tested in the context of cointegration and Granger causality. For Denmark, Greece, Ireland, the Netherlands, Portugal and Sweden, the evidence was consistent with the fiscal synchronisation hypothesis. For Finland, France, Italy, Spain and the UK, the findings point out the tax-spend hypothesis. The spend-tax view is supported in the case of Luxembourg. Finally, the institutional separation hypothesis seems to be the case for Austria, Belgium and Germany. Ho and Huang (2009) tested the interaction between revenues and expenditures for 31 Chinese provinces using panel data covering 1999 to 2005. Based on panel vector error correction analysis, their findings support the fiscal synchronisation hypothesis for all 31 provinces.

3. Methodology

Panel models of various types that can be estimated by using different estimation techniques may be formulated by the following specification:

$$Y_{it} = \alpha + X'_{it} b_{it} + \delta_i + \gamma_t + \varepsilon_{it} \tag{1}$$

where $i = 1,2,\ldots,N$ and $t = 1,2,\ldots,T$ denote cross-sectional units and dated periods respectively; Y_{it} is the dependent variable; X_{it} is a *k*-vector of explanatory variables; *b* is the vector of coefficients to be estimated; the parameter α indicates the overall constant; δ_i and γ_t reflect cross-sectional or period-specific effects which can be fixed or random; and ε_{it} are the error terms. The *b* coefficients may be classified into groups of cross-section specific, period-specific, and common regressor parameters. The structure of vector *b* allows for *b* coefficients to differ across cross sections or periods. Estimating model 1, one can create interaction variables in order to generate variations for *b* coefficients across periods or cross section countries. In panel model 1 the terms

δ_i and γ_t may be handled adopting fixed or random effects procedures. The fixed effects specifications employ orthogonal projections which contain a proper approach to remove cross section means from the dependent variable and exogenous regressors. Given that in the estimation procedure instrumental variables are formulated with fixed effects, orthogonal projections are also applied to instruments. The random effects approaches suppose that the terms δ_i and γ_t are realisations of independent random variables with mean zero and finite variance, and the effects are not correlated with the idiosyncratic residuals $\hat{\varepsilon}_{it}$.

Based on panel model 1 the OLS estimator is given as

$$\hat{b}_{OLS} = \left(\sum_i X'_i X_i \right)^{-1} \left(\sum_i X'_i Y_i \right) \tag{2}$$

Taking into account that TSLS is a straightforward extension of the standard OLS estimator, the TSLS estimator is written as

$$\hat{b}_{TSLS} = \left(\sum_i X'_i P_{Z_i} X_i \right)^{-1} \left(\sum_i X'_i P_{Z_i} Y_i \right) \tag{3}$$

where the orthogonal projection matrix P_{Z_i} for the instruments Z_i is

$$\mathbf{P}_{\mathbf{Z}_\mathbf{i}} = \left[\mathbf{Z}_\mathbf{i} \left(\mathbf{Z'}_\mathbf{i} \mathbf{Z}_\mathbf{i} \right)^{-1} \mathbf{Z'}_\mathbf{i} \right] \tag{4}$$

The covariance matrix of the composite error V_{it}(V_{it}=δ_{it}+γ_{it}+ε_{it}) can be calculated by various procedures, such as the Swamy-Arora, Wallace-Hussain and Wansbeek-Kapteyn estimators.[7] In large samples the Wansbeek-Kapteyn estimator of the component variances, Var(V_{it}), leads to similar results compared with Swamy-Arora and Wallace-Hussain estimators. The purpose of TSLS approach is to generalised least squares (GLS) transform the dependent and regressor data included in the instruments prior to estimation. TSLS permits for different structures of correlations between the residuals. Contemporaneous covariances, period-specific heteroskedasticity and between-period covariances constitute essential specifications that allow for various patterns of residual variance.

Based on equation 1, GMM estimators may indicate moments of the following form:

$$\hat{b}_{GMM} = \sum_{i=1}^{m} Z'_i \, \varepsilon_i \left(b \right) \tag{5}$$

where $\boldsymbol{Z_i}$ is a $\boldsymbol{T_i}$ x $\boldsymbol{P}$ matrix of instruments for cross-section *i*, and

$$\varepsilon_i(b) = \left[Y_i - f(X_{it}, b)\right] \tag{6}$$

Having estimated the coefficient vector $\hat{b}$ the coefficient covariance matrix is computed as below:

$$Var(\hat{b}) = (\Phi' H \Phi)^{-1} (\Phi' H \Lambda H \Phi)(\Phi' H \Phi)^{-1} \tag{7}$$

where $\boldsymbol{H}$ is a $\boldsymbol{p}$ x $\boldsymbol{p}$ weighting matrix, $\boldsymbol{\Phi}$ is a derivative matrix with dimensions $\boldsymbol{T_i}$ x $\boldsymbol{K}$ and $\boldsymbol{\Lambda}$ is an estimator of $E\left[\Phi_i(b)\Phi_i(b')\right] = E\left[Z_i'\varepsilon_i(b)\varepsilon_i(b')Z_i\right]$. The GMM estimation technique follows three main stages: (i) determining the instruments $\boldsymbol{Z}$; (ii) choosing the weighting matrix $\boldsymbol{H}$; and (iii) specifying an estimator for $\boldsymbol{\Lambda}$. One may formulate alternative specifications for $\boldsymbol{H}$ and $\boldsymbol{\Lambda}$. It should be noted that using GLS-transformed data we can compute GMM estimators. In this case, the GMM estimator is expressed by the GLS weighting

$$\hat{b}_{GMM} = \sum_{i=1}^{M} Z'_i \hat{\Omega}^{-1} \varepsilon_i(b) \tag{8}$$

where $\hat{\Omega}$ is an estimator of the contemporaneous variance-covariance matrix. Efficient GMM estimators can be computed using dynamic panel data approaches. Consider the formulation

$$Y_{it} = \sum_{j=1}^{p} \pi_j Y_{it-j} + X_{it}' b + \delta_i + \varepsilon_{it} \tag{9}$$

First differencing (9) produces

$$\Delta Y_{it} = \sum_{j=1}^{p} \pi_j \Delta Y_{it-j} + \Delta X_{it}' b + \Delta\varepsilon_{it} \tag{10}$$

The individual effect δ_i has been eliminated by first-differencing. Specification 10 reflects a dynamic panel model which may be estimated employing GMM methods. In GMM model 10, the period-specific instruments are related to lagged values of the dependent and predetermined variables. In the estimation procedure, along with a group of strictly exogenous variables, various instruments for each period will be used to produce efficient GMM coefficients. Given that in model 10 the disturbances are not autocorrelated Arellano and Bond (1991), for the two-step estimator, proposed the following weighting matrix:

$$H = \left(M^{-1} \sum_{i=1}^{M} Z_i' \Delta\varepsilon_i \Delta\varepsilon_i' Z_i \right)^{-1} \tag{11}$$

where Z_i includes a group of strictly exogenous and predetermined instruments. Arellano and Bover (1995) introduced the alternative approach of transforming the residuals using orthogonal deviations, where the optimal first-stage weighting matrix corresponds to the following weighting matrix:

$$H = \left(M^{-1} \sum_{i=1}^{M} Z_i' Z_i \right)^{-1} \tag{12}$$

4. Data and estimation results

All data come from Eurostat databases and form a balanced annual panel for 1970 to 2007 for 15 EU countries: Austria, Belgium, Germany, Denmark, Greece, Spain, United Kingdom, Finland, France, Ireland, Italy, Luxembourg, Netherlands, Portugal and Sweden. The remaining 12 EU countries which joined in 2004 and 2007 were excluded from the analysis because of incomplete data. It was necessary to apply the various panel data techniques with the same data set of EU member states in order to produce reliable results. The choice of the entire period 1970-2007 was influenced by the fact that the data set for general government is available on a comparable basis for the 15 EU countries. The empirical analysis employs data on total expenditure of general government (G), total revenue of general government (T), the unemployment rate (U), and the interest rate (E).[8] G is measured by the share of general government expenditures in GDP. T denotes the ratio of general government revenues in GDP. U and E are adopted in the empirical procedure as critical variables in order to exploit the pattern of relation between G and T. Public expenditures and revenues seem to be sensitive to U and E changes.

In Tables 1, 2 and 3, TSLS panel models are estimated, using either G or T as the dependent variable. We split the entire period 1970-2007 into two subperiods 1970-1991 and 1992-2007, presenting the TSLS estimates for robustness tests. There are various reasons for choosing the subperiod 1992-2007 to evaluate the four hypotheses. After 1992, when the Maastricht Treaty was established, the EU followed a path towards EMU. The Maastricht Treaty created its own economic and political dynamics, inducing EU countries to undertake efficient macroeconomic policies in order to participate later on in EMU. Protocol 5 of the Maastricht Treaty formulates the fiscal convergence criteria for an individual EU country to join EMU. A general government debt-to-GDP ratio of not more than 60 per cent and a ratio of the general government deficit to GDP of not more than 3 per cent in each EU member state represent the so-called fiscal convergence criteria.

Table 1. Panel TSLS results

variables	G_{it}	T_{it}	G_{it}	T_{it}	G_{it}	T_{it}
Fixed effects						
Constant	9.677 (4.191)	0.305 (0.089)	11.388 (3.797)	-0.904 (0.239)	5.817 (1.483)	2.262 (0.588)
T_{it}	0.841 (16.742)	–	0.805 (11.848)	–	0.918 (11.796)	–
G_{it}	–	0.934 (13.108)	–	0.946 (11.397)	–	0.910 (12.003)
R^2	0.841	0.820	0.830	0.794	0.854	0.837
F	74.0	63.5	68.2	53.6	81.7	71.5
obs	555	555	315	315	240	240
Random effects						
Constant	9.349 (3.532)	1.171 (0.344)	10.770 (3.138)	0.884 (2.263)	5.803 (1.216)	2.845 (3.692)
T_{it}	0.849 (18.429)	–	0.820 (12.851)	–	0.918 (11.847)	–
G_{it}	–	0.915 (14.528)	–	0.906 (12.899)	–	0.898 (13.085)
R^2	0.782	0.780	0.759	0.761	0.827	0.825
F	197.8	195.3	97.5	99.7	113.8	112.3
Hau	0.141	0.429	0.749	0.813	0.002	0.141
obs	555	555	315	315	240	240

Notes: The data consist of a balanced annual panel for 15 EU countries: Austria, Belgium, Germany, Denmark, Greece, Spain, United Kingdom, Finland, France, Ireland, Italy, Luxembourg, Netherlands, Portugal and Sweden. Absolute values of t-ratios are in parentheses. The instruments in TSLS pooled estimates are lagged explanatory variables. R^2 is the within-R^2 for fixed effects and overall- R^2 for random effects. The F tests check the joint statistical significance of the fixed or random effects TSLS estimates. Hau is the Hausman statistic which evaluates the null hypothesis that there is no misspecification in the random effects estimation. obs is the number of observations.

To test the validity of the four hypotheses, we employ the TSLS technique estimating fixed and random effects models. The common part of the models includes a constant term. Applying TSLS we introduce in the instrument list lags of G, T, U and E as well as the constant term. The values of R^2 and F-statistics suggest the explanatory power of the entire model. In both fixed and random effects estimation procedures, we specify settings for White cross-section to allow for general contemporaneous correlation between the country residuals. White cross-section permits for TSLS estimators to correct for both cross-section heteroskedasticity and contemporaneous correlation. In the random effects specifications, we employ the Wansbeak-Kapteyn (1989) method in order to calculate the estimates of random component variances. The null hypothesis that the instruments are uncorrelated with the error terms is tested by using the Hausman (1978) test of overidentifying restrictions. The Hausman statistic follows asymptotically the Chi-square with k degrees of freedom equal to the number of estimated coefficients.

Table 2. Panel TSLS results

variables	G_{it}	T_{it}	G_{it}	T_{it}	G_{it}	T_{it}
Fixed effects						
Constant	5.718 (2.270)	3.616 (1.181)	6.331 (2.708)	3.834 (0.954)	3.653 (0.874)	3.718 (2.022)
T_{it}	0.874 (15.396)	–	0.861 (14.348)	–	0.924 (11.823)	–
G_{it}	–	0.927 (15.250)	–	0.920 (12.017)	–	0.918 (13.905)
U_{it}	0.383 (2.577)	-0.456 (3.822)	0.480 (2.939)	-0.623 (3.607)	0.249 (1.481)	-0.243 (1.682)
R^2	0.860	0.848	0.853	0.838	0.866	0.850
F	83.3	75.7	77.2	68.7	84.7	74.0
obs	555	555	315	315	240	240
Random effects						
Constant	5.379 (2.233)	3.755 (1.254)	4.986 (2.365)	3.970 (1.594)	3.450 (2.725)	4.434 (2.234)
T_{it}	0.878 (16.765)	–	0.883 (16.288)	–	0.924 (11.868)	–
G_{it}	–	0.923 (17.251)	–	0.915 (18.215)	–	0.907 (15.358)
U_{it}	0.413 (3.315)	-0.451 (4.212)	0.562 (3.751)	-0.607 (4.023)	0.271 (1.853)	-0.264 (2.135)
R^2	0.811	0.815	0.830	0.821	0.842	0.839
F	118.2	121.1	66.5	71.6	63.1	61.9
Hau	0.165	0.051	1.619	0.001	0.924	0.159
obs	555	555	315	315	240	240

Note: see Table 1 for a detailed discussion of the various test statistics.

The TSLS fixed and random effects results reported in Tables 1, 2 and 3 indicate that there is an interplay between G and T in all time periods. The variables G and T carry positive and highly significant coefficients, implying the validity of the fiscal synchronisation hypothesis. Introducing U and E as control variables, we find that in most regressions the explanatory variables G and T are statistically significant at better than the 5 per cent level. Focusing on the robustness of TSLS estimates, we observe that the estimates tend to remain highly significant over the entire period 1970-2007 and the two subperiods 1970-1991 and 1992-2007. It is interesting to note that the point estimates for all the time periods appear to be relatively close. The goodness of fit as calculated by R^2 and F-statistics is satisfactory in all specifications. In the random effects estimates, the Hau statistics indicate that the null hypothesis is not rejected, confirming that there is no misspecification in the estimated models. The essential conclusion from panel data analysis of Tables 1, 2, and 3 is that using either fixed or random effects in the estimation procedure, TSLS results for 15 EU countries are broadly in line with the fiscal synchronisation hypothesis.

Table 3: Panel TSLS results

variables	G_{it}	T_{it}	G_{it}	T_{it}	G_{it}	T_{it}
Fixed effects						
Constant	4.468 (1.234)	7.227 (1.649)	6.903 (1.526)	8.582 (2.445)	0.086 (1.002)	5.643 (2.391)
T_{it}	0.901 (15.097)	–	0.862 (11.257)	–	0.972 (13.200)	–
G_{it}	–	0.888 (11.481)	–	0.868 (8.375)	–	0.918 (12.948)
E_{it}	0.306 (1.855)	-0.530 (3.562)	0.219 (1.323)	-0.523 (2.927)	0.547 (3.062)	-0.610 (4.620)
R^2	0.831	0.836	0.822	0.809	0.870	0.863
F	67.1	75.8	60.4	59.7	91.4	98.9
obs	524	524	284	284	240	240
Random effects						
Constant	2.740 (1.897)	8.268 (2.699)	4.308 (2.101)	10.733 (3.309)	-1.127 (2.269)	6.572 (1.993)
T_{it}	0.925 (17.429)	–	0.900 (13.293)	–	0.980 (15.016)	–
G_{it}	–	0.864 (15.529)	–	0.840 (12.456)	–	0.910 (15.555)
E_{it}	0.374 (2.589)	-0.521 (5.353)	0.309 (2.140)	-0.593 (4.724)	0.667 (3.499)	-0.701 (4.540)
R^2	0.789	0.817	0.765	0.803	0.849	0.853
F	98.2	131.6	48.1	63.2	70.7	87.3
Hau	0.914	0.613	0.863	0.001	0.334	0.004
obs	524	524	284	284	240	240

Note: see Table 1 for a detailed discussion of the various test statistics.

To evaluate the TSLS results presented in Tables 1, 2 and 3, we continue by using the GMM estimator to test the validity of the four hypotheses. The GMM method reflects an instrumental variables estimation of dynamic panel model 10. Since our GMM model is dynamic, we have to describe its basic characteristics. First, in the group of regressors we include one lag of the dependent variable. Second, the variables U and E are specified as explanatory variables because they express overall economic activity and affect the behaviour of both G and T. Third, orthogonal deviations are chosen as a transformation to remove the cross section fixed effect. Period dummy variables are included in the system and they are entered in first differences. Fourth, lag lengths of the dependent and explanatory variables are specified as predetermined instruments. Fifth, the weighting matrix H, depicted by formula 11, is chosen in order to allow the GMM estimates to be robust to possible serial correlation and heteroskedasticity of unknown type in the error terms. The final characteristic of our GMM estimation procedure is that the Arellano-Bond two-step estimator is used to specify the GMM weighting and coefficient covariance computation choices. In this case, the GMM estimator permits for computing White period robust standard errors.

Having described the characteristics of the GMM methodology, we discuss the various test statistics used in the empirical analysis. We use the R^2 statistics to determine the explanatory power of the entire estimated models. We perform an F test to check the null hypothesis that the lagged independent variables are redundant. A J-statistic is calculated to evaluate the validity of over-identifying restrictions. The J-statistic is simply the Sargan statistic which is distributed as a $X_{ip\text{-}k}$, where *IP* is the instrument rank and *k* is the number of estimated coefficients. To check if the error terms are serially correlated we use the AR test statistics. AR(1) and AR(2) are statistics for first-and second-order serial correlation.

Table 4 reports GMM results. Our GMM estimates are compatible in terms of sign and statistical significance with the equivalent estimates obtained by TSLS in Tables 1, 2 and 3. In case ΔG and ΔT are used as explanatory variables, they have a positive and highly significant sign indicating the acceptance of the fiscal synchronisation hypothesis. The coefficients of ΔU and ΔE appear to be significant in all time periods. Lag values of ΔG, ΔT, ΔU and ΔE are very reliable, suggesting that the explanatory variables are highly significant. R^2 statistics show that the estimated models have satisfactory goodness of fit. The high F-statistic values imply the rejection of the null hypothesis, suggesting that $\Delta G_{it\text{-}1}$, $\Delta T_{it\text{-}1}$, $\Delta U_{it\text{-}1}$ and $\Delta E_{it\text{-}1}$ are not redundant in the unrestricted GMM models. J-test statistics indicate that the error terms and explanatory variables are independent, lead-

Table 4. GMM estimates

variables	ΔG_{it}	ΔT_{it}	ΔG_{it}	ΔT_{it}	ΔG_{it}	ΔT_{it}
Constant	-0.011	0.125	0.035	-0.025	-0.062	0.199
	(0.177)	(1.819)	(0.362)	(0.234)	(2.124)	(2.348)
ΔT_{it}	0.697		0.703		0.569	
	(7.183)	–	(6.399)	–	(3.040)	–
ΔG_{it}		0.254		0.332		0.575
	–	(2.048)	–	(2.229)	–	(3.976)
ΔU_{it}	0.861	0.542	0.968	0.708	0.974	-0.338
	(2.723)	(1.801)	(5.849)	(2.001)	(2.884)	(1.942)
ΔT_{it-1}	-0.005	0.030	0.034	0.023	-0.197	0.050
	(0.083)	(0.703)	(0.522)	(0.407)	(2.929)	(1.093)
ΔG_{it-1}	0.002	-0.004	-0.035	0.006	0.148	0.005
	(0.012)	(0.084)	(0.662)	(0.113)	(1.596)	(0.090)
ΔU_{it-1}	-0.194	-0.173	-0.172	-0.225	-0.437	0.108
	(1.847)	(1.827)	(2.157)	(1.863)	(2.797)	(1.410)
R^2	0.669	0.405	0.769	0.532	0.489	0.555
F	16.8	15.9	17.7	22.4	31.5	27.1
J	0.017	0.016	0.036	0.022	0.060	0.037
AR(1)	1.146	3.999	1.684	1.123	0.791	0.778
AR(2)	1.163	4.214	1.896	1.455	0.878	1.144
obs	566	566	325	325	234	234
Constant	0.006	0.149	0.192	0.030	-0.043	0.214
	(0.089)	(2.613)	(1.649)	(0.292)	(1.949)	(2.188)
ΔT_{it}	0.944		0.827		0.864	
	(5.211)	–	(8.024)	–	(3.367)	–
ΔG_{it}		0.305		0.541		0.723
	–	(3.530)	–	(6.521)	–	(4.339)
ΔE_{it}	-0.383	0.358	-0.413	0.405	0.617	-0.475
	(1.122)	(2.125)	(1.571)	(2.373)	(4.374)	(2.911)
ΔT_{it-1}	-0.058	-0.069	0.211	-0.077	0.035	-0.004
	(0.875)	(1.659)	(2.581)	(1.092)	(0.400)	(0.047)
ΔG_{it-1}	0.127	0.089	-0.185	0.092	-0.005	0.005
	(1.802)	(2.601)	(2.939)	(1.636)	(0.060)	(0.069)
ΔE_{it-1}	0.464	-0.106	0.678	-0.304	0.288	-0.191
	(5.433)	(1.755)	(5.024)	(3.335)	(3.633)	(2.133)
R^2	0.612	0.498	0.591	0.618	0.677	0.660
F	35.1	26.1	29.5	31.6	34.4	41.2
J	0.027	0.025	0.128	0.048	0.036	0.055
AR(1)	0.404	0.718	1.839	2.107	2.341	0.993
AR(2)	1.112	0.926	2.447	2.214	1.877	1.514
obs	564	564	323	323	232	232

Notes: For the EU countries included in the data set, see Table 1. t-ratios are in parentheses. Δ is the difference operator. Cross-section fixed effects are used by employing Arellano-Bond 2-step estimation. The instruments in the Arellano-Bond dynamic panel GMM estimators are lags of explanatory variables. R^2 is the coefficient of multiple determinations. The F-test statistics check the null hypothesis that the lagged explanatory variables are redundant. J is the Sargan statistic of overidentifying restrictions of the instruments used in the estimation procedure. obs is the number of observations. AR (1) and AR (2) are LM tests of first- and second-order serial correlation following asymptotically the χ^2 distribution.

ing to the conclusion that the instruments are exogenous and, thus, appropriately chosen. The low values of AR(1) and AR(2) show that the estimated GMM models pass the test for the absence of serial correlation in the residuals. Overall, the results in favour of the fiscal synchronisation theory do not change either using and as control variables, or employing alternative time frames.

5. ROBUSTNESS CHECKS

In this Section, we provide alternative estimates to evaluate the results in Section 4. The empirical analysis uses Pedroni's (1999, 2004) panel error correction modelling to infer Granger causality relationships between government revenues and expenditures within the framework of bivariate- and trivariate- models.[9] Our econometric methodology proceeds in three stages. First, we perform panel unit root tests to ascertain the order of integration of the variables. Second, conditional on finding that each variable under study is integrated of order one, we check for panel cointegration using the technique proposed by Pedroni (1999, 2004). Third, conditional on finding panel cointegration, we test for Granger causality between *G* and *T* (model 1), between *G, T* and *U* (model 2) and between *G,T* and *E* (model 3).

The Im, Pesaran and Shin (2003), or IPS, panel unit root is used in our empirical procedure. Consider the following AR (1) model for panel data:

$$Y_{it} = \alpha_i Y_{it-1} + X_{it}\gamma_i + \varepsilon_{it} \tag{13}$$

Y_{it} denotes a stochastic process for a panel of cross section countries; $i=1,\ldots,N$ and each cross section contains $t=1,\ldots,T$ time series observations. The X_{it} indicate the exogenous variables, α_i are the autoregressive coefficients and the errors ε_{it} are supposed to be *iid* ~(0, σ_ε^2). If $|\alpha_i|<1$, Y_i is trend stationary and if $|\alpha_i|=1$, Y_i contains a unit root. The IPS test assumes that α_i vary freely across cross section countries, allowing for different orders of autocorrelation. The IPS test employs a null hypothesis of a unit root and starts by considering the following Augmented Dickey-Fuller (ADF) formulation:

$$\Delta Y_{it} = \rho Y_{it-1} + \sum_{j=1}^{p_i} b_{ij}\Delta Y_{it-j} + X_{it}^{'}\delta + \varepsilon_{it} \tag{14}$$

where $\rho=\alpha-1$. The lag length for the difference terms, ,can vary across cross-sectional individuals. The IPS test is based on the average ADF statistics, calculated for each group in the panel, referred to as the t-bar test procedure. Taking into account that the IPS test permits for individual unit root processes, the null and alternative hypotheses are specified as follows:

$$H_0 : \rho_i = 0 \quad \text{for all } i \tag{15}$$

$$H_1 : \begin{cases} \rho_i = 0 & \text{for } i = N_1+1,\ N_1+2,\dots,N \\ \rho_i < 0 & \text{for } i = 1,2,\dots,N_1 \end{cases} \tag{16}$$

This specification of the alternative hypothesis permits the coefficients ρ_i to differ across cross-sectional units. Thus, the IPS testing procedure allows for some of the individual variables to have unit roots under the alternative hypothesis. In cases where the disturbances in the dynamic panel are autocorrelated, the t-bar tests require that both T and N should be sufficiently large. In models with serially correlated errors, IPS proposed the W_{tbar}-statistic, which is asymptotically equivalent to the t-bar statistic. The W_{tbar}-test seems to perform very well even for relatively small samples.[10]

Table 5. IPS panel unit root tests

	1970-2007		*1970-1991*		*1992-2007*	
Variable	*Levels*	*First order difference*	*Levels*	*First order difference*	*Levels*	*First order difference*
T_{it}	0.268	-7.318*	-0.227	-3.980*	-0.972	-3.384*
G_{it}	-0.919	-5.113*	1.346	-2.075**	-0.837	-3.204*
U_{it}	-0.998	-6.365*	1.568	-2.505*	0.585	-4.240*
E_{it}	0.079	-8.845*	0.282	3.317*	1.149	-5.207*

*, ** indicates significance at the 1%, 5% levels respectively.
Notes: W_{tbar} test statistics are computed for the null hypothesis of a unit root. The critical values of IPS tests are taken from Im, Perasan and Shin (2003). All the statistics are left-tailed tests. The lag length for the IPS regressions is specified using the AIC (Akaike information criterion). Δ ιs the difference operator.

Given that our panel data analysis aims at evaluating the robustness of the results, we conduct IPS panel unit root tests for the entire period 1970-2007 and the two subperiods 1970-1991 and 1992-2007. In Table 5, IPS is the W_{tbar}-statistic, which takes explicit account of the underlying ADF orders in calculating the mean and the variance adjustment factors. All unit root tests were carried out employing regressions with individual constants and trends at level forms. For the first order differences only constant terms were included in the regressions, since differencing usually removed the deterministic terms. The lag length specification is determined using the Akaike Information Criterion (AIC).[11] It is clear from Table 5 that all the individual variables are found to be nonstationary. The null hypothesis of a unit root cannot be rejected at either the 1 per cent or the 5 per cent level of significance. Thus, all of the variables have unit roots. Moreover, when IPS tests are conducted on first order differences, the null hypothesis of nonstationarity is easily rejected, suggesting that each of the variables T_{it}, G_{it}, U_{it} and E_{it} is I(1).

Once the existence of a panel unit root has been established, the next step is to test whether a long-run equilibrium relationship exists in the cases of the bivariate and trivariate models:

$$T_{it} = \alpha_i + \delta_i t + \beta_i G_{it} + \varepsilon_{it} \tag{17}$$

$$T_{it} = \alpha_i + \delta_i t + \beta_i G_{it} + \gamma_i U_{it} + v_{it} \tag{18}$$

$$T_{it} = \alpha_i + \delta_i t + \beta_i G_{it} + \varphi_i E_{it} + \mu_{it} \tag{19}$$

Here $t = 1,\ldots,T$ and $i = 1,\ldots N$, where T denotes the number of observations and N refers to the number of individual countries in the panel. The bivariate and trivariate models (T,G), (T,G,U) and (T,G,E), allow for cointegrating vectors between EU countries. Table 6 presents Pedroni's panel cointegration estimation results. Pedroni (1999, 2004) proposed a cointegration panel test based on the Engle-Granger (1987) two-step cointegration procedure. Pedroni's cointegration technique overcomes the problem of small samples, allows for heterogeneity in the intercepts and slopes of the cointegrating equation and permits for trend coefficients across cross sections. Regarding equations 17, 18 and 19, the null hypothesis of no cointegration suggests that ε_{it}, v_{it} and μ_{it} will be I(1).

Pedroni (1999, 2004) constructs various statistics for testing the null hypothesis. The panel cointegration test results are based on Pedroni's seven test statistics presented in Table 6.[12] All the tests are conducted by considering deterministic intercept and trend in the estimation procedure. In Panel A, we present the results for the bivariate model (T,G), while Panels B and C report the findings for the trivariate models (T,G,U) and (T,G,E). In all the models and examining the entire period 1970-2007, we find that six of the seven test statistics reveal evidence for panel cointegration at the 5 per cent level. In the time frame 1992-2007, the results show that five of the seven tests suggest panel cointegration, leading to the main conclusion that both the bivariate and the trivariate models are cointegrated. Regarding the sample 1970-1991, in the case of model 17, only the group rho-statistic is not significant, while the other six statistics imply the rejection of the null hypothesis. Evaluating models 18 and 19 over 1970-1991, the results show that the null of no cointegration is rejected. Overall, Pedroni's panel cointegration tests suggest that T and G, T, G and U, T, G and E follow a strong long-run equilibrium relationship.

Table 6. Panel cointegration tests

A. Model 17	1970-2007		1970-1991		1992-2007	
Panel variance-stat	2.826	(0.0024)*	2.553	(0.0053)*	-0.742	(0.7710)
Panel rho-stat	-2.706	(0.0034)*	-2.158	(0.0155)*	-1.847	(0.0342)*
Panel PP-stat	-3.666	(0.0001)*	-4.614	(0.0000)*	-2.348	(0.0094)*
Panel ADF-stat	-3.340	(0.0004)*	-6.099	(0.0000)*	-2.794	(0.0026)*
Group rho-stat	-1.396	(0.0813)	-0.3110	(0.3779)	1.721	(0.9574)
Group PP-stat	-3.338	(0.0004)*	-3.894	(0.0000)*	-1.828	(0.0338)*
Group ADF-stat	-3.321	(0.0004)*	-5.749	(0.0000)*	-3.198	(0.0007)*
B. Model 18						
Panel variance-stat	3.212	(0.0007)*	1.612	(0.0535)	-0.2830	(0.6114)
Panel rho-stat	-2.794	(0.0026)*	-2.245	(0.0104)*	-1.978	(0.0164)*
Panel PP-stat	-4.839	(0.000)*	-5.144	(0.0000)*	-2.052	(0.0201)*
Panel ADF-stat	-3.868	(0.0001)*	-6.168	(0.0000)*	-4.250	(0.0000)*
Group rho-stat	-1.660	(0.0485)	0.3441	(0.6346)	2.520	(0.9941)
Group PP-stat	-4.856	(0.0000)*	-4.545	(0.0000)*	-1.895	(0.0291)*
Group ADF-stat	-4.042	(0.0000)*	-6.253	(0.0000)*	-5.397	(0.0000)*
C. Model 19						
Panel variance-stat	2.713	(0.0033)*	2.003	(0.0226)*	-1.041	(0.8510)
Panel rho-stat	-2.439	(0.0074)*	-1.177	(0.1196)	-1.824	(0.0281)*
Panel PP-stat	-4.073	(0.000)*	-4.813	(0.0000)*	-2.115	(0.0178)*
Panel ADF-stat	-3.432	(0.0003)*	-5.986	(0.0000)*	-2.553	(0.0053)*
Group rho-stat	-0.937	(0.1743)	0.336	(0.6316)	3.052	(0.9989)
Group PP-stat	-3.528	(0.0002)*	-4.532	(0.0000)*	-2.878	(0.0023)*
Group ADF-stat	-3.976	(0.0000)*	-6.648	(0.0000)*	-2.894	(0.019)*

Notes: * rejects the null of no cointegration at the 5% level. P-values are in parentheses. Statistics are asymptotically distributed as normal. The variance ratio test is right-sided while the others are left-sided. The trend assumption considers deterministic intercept and trend in the estimation procedure. Lag length selection is based on AIC (Akaike information criterion) with a maximum lag of 8.

Having established that a long-run equilibrium relationship exists among (*T,G*), (*T,G,U*) and (*T,G,E*), the next stage of our empirical analysis is to explore for panel Granger causality. Engle and Granger (1987) argue that if two nonstationary variables are cointegrated, then a causal relation will exist between them. Our methodology employs the Holtz-Eakin, Newey and Rosen (1989) approach, which uses a panel-based VECM. This indicates the traditional panel VAR model is expanded with a one period lagged error correction term, which is obtained from the cointegrated system. The panel Granger causality tests will be based on the following bivariate model:

$$\Delta T_{it} = \pi_{1j} + \sum_k \pi_{11ik} \Delta T_{it-k} + \sum_k \pi_{12k} \Delta G_{it-k} + \lambda_{1i} EC_{it-1} + u_{1t} \tag{20}$$

$$\Delta G_{it} = \pi_{2j} + \sum_k \pi_{21ik} \Delta G_{it-k} + \sum_k \pi_{22ik} \Delta T_{it-k} + \lambda_{2i} E_{it-1} + u_{1t} \tag{21}$$

Introducing *U* and *E* as third variables, in the sense that *U* and *E* affect the behaviour of both *T* and *G*, we receive the trivariate models of the following type:

$$\Delta T_{it} = \pi_{1j} + \sum_k \pi_{11ik} \Delta T_{it-k} + \sum_k \pi_{12ik} \Delta G_{it-k} + \sum_k \pi_{13k} \Delta U_{it-k} + \lambda_{1i} EC_{it-1} + u_{1t} \tag{22}$$

$$\Delta G_{it} = \pi_{2j} + \sum_k \pi_{21ik} \Delta G_{it-k} + \sum_k \pi_{22ik} \Delta T_{it-k} + \sum_k \pi_{23k} \Delta U_{it-k} + \lambda_{2i} EC_{it-1} + u_{1t} \tag{23}$$

and

$$\Delta T_{it} = \pi_{1j} + \sum_k \pi_{11ik} \Delta T_{it-k} + \sum_k \pi_{12ik} \Delta G_{it-k} + \sum_k \pi_{13k} \Delta E_{it-k} + \lambda_{1i} EC_{it-1} + u_{1t} \tag{24}$$

$$\Delta G_{it} = \pi_{2j} + \sum_k \pi_{21ik} \Delta G_{it-k} + \sum_k \pi_{22ik} \Delta T_{it-k} + \sum_k \pi_{23k} \Delta E_{it-k} + \lambda_{2i} EC_{it-1} + u_{1t} \tag{25}$$

Here, EC_{it-1} are the estimated residuals lagged one period. Thus, $\lambda_{1i}EC_{it-1}$ and $\lambda_{2i}EC_{it-1}$ indicate the long-run equilibrium relationship, where the parameters λ_{1i} and λ_{2i} show the speed of adjustment. Δ is the first difference operator and denotes the lag length specification. Panel Granger short-run causality tests are conducted by checking, via a standard F-test, whether all the coefficients of ΔT_{it-k} or ΔG_{it-k} are statistically different from zero as a group. Long-run Granger causality is examined by testing the significance of the λ_{1i} and λ_{2i} coefficients, based on a t-test. Since the Granger causality tests appear to be sensitive to the lag order selection, the lag lengths are specified using the AIC.

Table 7 presents the results on short-run and long-run causality for the bivariate and trivariate models 20 to 25. For models (ΔT, ΔG), (ΔT, ΔG, ΔU) and (ΔT, ΔG, ΔE), the key conclusion from our empirical analysis is that between ΔT and ΔG there is strong bidirectional long-run Granger causality in 1970-2007 and 1992-2007 periods.[13] This result is consistent with the fiscal synchronisation theory, leading to the inference that fiscal policy in the 15 EU member countries, in the pre-Maastricht and post-Maastricht periods, suggests that spending measures are conducted with revenues decisions. Efforts to achieve low budget deficits should be accomplished by the joint determination of government revenues and expenditures. The right balance between revenues and expenditures can be found by focusing fiscal rules on avoiding budgetary instability and on ensuring long-term fiscal sustainability.

Table 7. Panel Granger causality tests

Dependent variable	1970-2007 s-run causality ΔG	ΔT	l-run causality λ_{1i}	λ_{2i}	1992-2007 s-run causality ΔG	ΔT	l-run causality λ_{1i}	λ_{2i}
Model (ΔT, ΔG)								
ΔT	0.604 (0.612)	—	-0.098* (3.451)	—	1.208 (0.308)	—	-0.205* (2.740)	—
ΔG	—	2.824* (0.060)	—	-0.085* (2.444)	—	1.131 (0.325)	—	-0.262* (3.180)
Model (ΔT, ΔG, ΔU)								
ΔT	0.279 (0.756)	—	0.134* (4.175)	—	2.099** (0.101)	—	-0.266* (3.259)	—
ΔG	—	2.231* (0.108)	—	-0.037 (0.947)	—	0.407 (0.666)	—	-0.243** (2.392)
Model (ΔT, ΔG, ΔE)								
ΔT	0.578 (0.561)	—	0.060** (1.757)	—	1.133 (0.342)	—	0.247* (2.812)	—
ΔG	—	1.058 (0.377)	—	-0.125* (3.242)	—	1.016 (0.364)	—	-0.180** (1.722)

*, ** indicates significance at the 1%, 5% levels, respectively.
Notes: Asymptotic t-statistics are in brackets and p-values in parentheses. The null hypothesis of no short-run Granger causality is tested using the Wald F-statistics. The evaluation of the long-run Granger causality is based on the significance of t-statistics.

6. Conclusions

This paper has attempted to contribute to the spend-tax debate by testing the four hypotheses in this field with EU panel data. To the best of our knowledge, no empirical study has tackled the issue of the causal relationship between general government revenues and expenditures in the EU by adopting the framework of panel data modelling. Almost all empirical papers have used VAR techniques based on cointegration, an ECM strategy and Granger causality tests. Our aim was to set up alternative panel models with *G* and *T* as well as a limited number of key macroeconomic variables, which affect the behaviour of both public revenues and expenditures, seeking to test the validity of the four hypotheses. TSLS results, based on a procedure which includes either fixed or random effects for EU member states, are broadly in line with the fiscal synchronisation hypothesis. Estimates appear to be robust in the entire period 1970-2007 and in the two subperiods 1970-1991 and 1992-2007.

In order to improve the TSLS estimation, we apply the GMM technique. The relationship between *G* and *T* is investigated using cross-section fixed effects and employing Arellano-Bond 2-step estimation. GMM estimates are obtained by introducing either ΔU or ΔE in the group of explanatory variables. Our GMM estimates are consistent in terms of sign and statistical significance with the TSLS findings reported in Tables 1, 2 and 3. Note that estimated models pass both the Sargan test and the test for the absence of AR (1) and AR (2) in the residuals. A comparison of the point estimates in the 1970-1991 and 1992-2007 subperiods shows that the coefficients for ΔG and ΔT are positive and significant, indicating acceptance of the fiscal synchronisation theory. In order to reinforce the evidence provided by TSLS and GMM techniques, we examine the direction of causation between ΔG and ΔT. Panel Granger tests suggest that in all time periods there is a powerful two-way long-run Granger causality between ΔG and ΔT.

It is well known that a considerable number of EU countries face problems with large budget deficits. One of the fundamental requirements of macroeconomic stability is that government deficit and debt should be under control. This requirement is of heightened importance in the EU, because the Maastricht Treaty defines 'excessive deficits' in order for an EU country to participate in EMU: deficits that exceed 3 per cent of GDP or general government debt above 60 per cent of GDP. The target of fiscal discipline must be rooted in realistic short-and long-term changes of tax receipts and government expenditures. In this way, the validity of the fiscal synchronisation hypothesis, in the case of the EU, suggests that for countries with high debt-to-GDP ratios, government decision-makers should plan fiscal policy having in mind the links between public spending and revenue in order to reduce deficits.

Accepted for publication: 1 July 2010

ENDNOTES

1. Associate Professor in Economics, The Athens University of Economics and Business, 76 Patission Street, Athens 104 34, Greece. Email: gvamv@ aueb.gr I wish to thank an anonymous referee of this journal for useful comments. All remaining errors and deficiencies are the responsibility of the author.

2. For a detailed and critical survey of the relevant literature on the spend-tax debate see Vamvoukas (1997) and Payne (2003).

3. In the context of the SGP, eurozone countries that have a deficit above 3 per cent of GDP are obliged to correct it in accordance with the recommendations of articles 104(6), 104(7) and 104(12) of the Maastricht Treaty. With respect to the year in which the excessive deficits occurred, EMU member states have up to two years to decrease the deficit below 3 per cent of GDP. However, under special circumstances such as the current international financial crisis, EMU countries with excessive deficit positions may need more than two years to address budgetary imbalances.

4. The spend-tax hypothesis is also known as the 'displacement effect' because the impact of structural breaks displaces the growth pattern of public outlays.

5. Supply-side economics originated in the United States in 1970s as a reaction to government economic policies proposed by Keynesians. The main propositions of supply-side economics come from the writings of the Classical and Neoclassical Schools. For a lucid and critical analysis of supply-side economics, see Hailstones (1982).

6. Payne (2003) mentions fifty-three empirical studies on the spend-tax controversy which explore the four hypotheses using VAR analysis and Granger causality tests.

7. Baltagi (2005) provides a comprehensive discussion on this point.

8. For an extended discussion on the specific definitions of the variables *G, T, U* and *E* used in our econometric analysis, see Statistical Annex of European Economy, Autumn 2008.

9. I would like to thank the referee's constructive suggestions for raising the issue to employ Pedroni's (1999, 2004) panel error correction framework to check the causal links between public revenues and outlays.

10. For a detailed analysis on the computation of t-bar and W_{tbar} statistics, see Im, Pesaran, and Shin (2003).

11. For an extended discussion on AIC, see Akaike (1987).

12. See Pedroni's original papers for more details on the construction of these test statistics.

13. To conserve space we do not present Granger test results over 1970-1991. Note that between 1970 and 1991 the findings are in line with the fiscal synchronisation hypothesis, suggesting mainly two-way long-run Granger causality between tax receipts and public outlays (the results are available upon request).

REFERENCES

Akaike H (1987) 'Factor analysis and AIC', *Psychometrica*, 52, 317-332.

Arellano M and Bond S R (1991) 'Some tests on specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies*, 58, 277-297.

Arellano M and Bover O (1995) 'Another look at the instrumental variables estimation of error-components models', *Journal of Econometrics,* 68, 29-51.

Baffes J and Shah A (1994) 'Causality and comovement between taxes and expenditures: Historical evidence from Argentina, Brazil, and Mexico', *Journal of Development Economics*, 44, 311-331.

Baltagi B H (2005) *Econometric Analysis of Panel Data,* 3e, New York: John Wiley & Sons.

Barro R (1979) 'On the determination of public debt', *Journal of Political Economy*, 87, 940-971.

Buchanan J M and Wagner R W (1977) *Democracy in Deficit,* New York: Academic Press.

Buchanan J M and Wagner R W (1978) 'Dialogues concerning fiscal religion', *Journal of Monetary Economics,* 4, 627-636.

Chang T, Liu W R and Caudill S B (2002) 'Tax-and-Spend, spend-and-tax, or fiscal synchronisation: New evidence from 10 countries', *Applied Economics,* 34, 1551-1561.

Darrat A F (1998) 'Tax and Spend, or Spend and Tax? An inquiry into the Turkish budgetary process', *Southern Economic Journal,* 64, 940-956.

Engle R F and Granger C W J (1987) 'Co-integration and error correction: representation, estimation and testing', *Econometrica*, 55, 251-276.

Friedman M (1972) *An Economist's Protest,* New Jersey: Horton & Co.

Friedman M (1978) 'The limitations of tax limitation', *Policy Review*, Summer, 7-14.

Hailstones T J (1982) *A Guide to Supply-Side Economics,* Reston VA: Reston PC.

Hausman J A (1978) 'Specification tests in econometrics', *Econometrica,* 46, 1251-1271.

Ho Y H and Huang C J (2009) 'Tax-spend, spend-tax, or fiscal synchronisation: a panel analysis of the Chinese provincial real data', *Journal of Economics and Management,* 5, 257-272.

Holtz-Eakin D, Newey W and Rosen H (1989) 'The revenues-expenditure nexus: evidence from local government data', *International Economic Review* 30, 415-429.

Im K S, Pesaran M H and Shin Y (2003) 'Testing for unit roots in heterogeneous panels', *Journal of Econometrics,* 115, 53-74.

Kollias C and Paleologou S-M (2006) 'Fiscal policy in the European Union: Tax and spend, spend and tax, fiscal synchronisation or institutional separation?', *Journal of Economic Studies,* 33, 108-120.

Koren S and Stiassny A (1998) 'Tax and spend, or spend and tax? An international study', *Journal of Policy Modeling,* 20, 163-191.

Meltzer A and Richard S (1981) 'A rational theory of the size of government', *Journal of Political Economy,* 89, 914-927.

Miller S and Russek F (1990) 'Co-integration and error correction models: The temporal causality between government taxes and spending', *Southern Economic Journal*, 56, 221-229.

Musgrave R (1966) 'Principles of budget determination' in Cameron H and Henderson W (eds) *Public Finance: Selected Readings,* New York: Random House.

Narayan P K and Narayan S (2006) 'Government revenue and government expenditure nexus: evidence from developing countries', *Applied Economics,* 38, 285-291.

Owoye O (1995) 'The causal relationship between taxes and expenditures in the G7 countries: cointegration and error-correction models', *Applied Economics Letters,* 2, 19-22.

Payne J E (2003) 'A survey of the international empirical evidence on the tax-spend debate', *Public Finance Review,* 31, 302-324.

Peacock A and Wiseman J (1961) *The Growth of Public Expenditures in the United Kingdom,* Princeton NJ: Princeton UP.

Peacock A and Wiseman J (1979) 'Approaches to the analysis of government expenditure growth', *Public Finance Quarterly,* 7, 3-23.

Pedroni P (1999) 'Critical values for cointegration tests in heterogeneous panels with multiple regressors', *Oxford Bulletin of Economics and Statistics,* 61, 653-670.

Pedroni P (2004) 'Panel cointegration; asymptotic and finite sample properties of pooled time series tests with an application to PPP hypothesis', *Econometric Theory*, 20, 597-625.

Toda H Y and Yamamoto T (1995) 'Statistical inference in vector autoregressions with possibly integrated processes', *Journal of Econometrics,* 66, 225-250.

Vamvoukas G A (1997) 'Budget expenditures and revenues: An application of error-correction modeling', *Public Finance*, 52, 125-138.

Wansbeek T and Kapteyn A (1989) 'Estimation of the error components model with incomplete panels', *Journal of Econometrics,* 41, 341-361.

Wildavsky (1975) *The Politics of the Budgetary Process,* Boston: Little Brown & Company

Wilensky H L (1975) *The Welfare State and Equality,* Berkeley-Los Angeles: University of California Press.