Economics Bulletin

Volume 40, Issue 4

# Changes in the tax-spend nexus: Evidence from selected European countries

Fabricio Linhares
*Graduate Program in Economics, Federal University of Ceara (CAEN/UFC)*

Glauber Nojosa
*Department of Applied Economics, Federal University of Ceara (DEA/UFC)*

## Abstract

This study assesses the tax-spend nexus in Germany, the United Kingdom (UK), France, Italy and Spain using quarterly data for the period 1995–2019. Different from previous studies, we test for changes in the causality links between government expenditures and revenues because the recent economic crisis and political changes in these countries may have shifted their tax-spend nexus. Based on the recursive rolling test developed by Shi et al. (2018), we found that the direction of causality in the tax-spend nexus for these countries significantly changed after the 2008 crisis and a series of recovery and stability policies implemented over 2010–2012. While there are several periods before and during the crisis where institutional separation hypothesis prevailed, the after-crisis data support uninterrupted causality links between revenues and expenditures in all countries: tax-and-spend in Germany, the UK and Italy; spend-and-tax in France; and fiscal synchronization in Spain.

We would like to thank Gueorgui I. Kolev (the associate editor), and two anonymous referees for their comments and suggestions, which improved the quality of this paper.
**Citation:** Fabricio Linhares and Glauber Nojosa, (2020) "Changes in the tax-spend nexus: Evidence from selected European countries", *Economics Bulletin*, Volume 40, Issue 4, pages 3077-3087
**Contact:** Fabricio Linhares - flinhares@caen.ufc.br, Glauber Nojosa - glaubernojosa@ufc.br.
**Submitted:** May 28, 2020. **Published:** November 24, 2020.

# 1. Introduction

When the financial crisis hit European countries in 2008, the fiscal measures implemented to reduce its negative effects led to a sharp deterioration in their budget/debt positions. While public expenditures were successively failing to adjust to the level and trajectory of the revenues, the average debt-to-GDP ratio in the region reached 86.1% in 2010. To contain concerns of sovereign debt crisis and to convince the market that their leaders would protect the economic stability in the region, the European Financial Stability Facility (EFSF) was temporarily instituted in May 2010. On other fronts, the European Union took a series of measures to strengthen its economic governance, limit increases in the budget deficit and avoid excessive growth of debt as a percentage of GDP. Such actions were based on legislative changes to ratify and strengthen the Stability and Growth Pact (SGP), such as the Six Pack (December 2011), the Two Pack (May 2013) and the Fiscal Compact (January 2013).

This research assesses the effects of these events on the causality between government revenues and expenditures in Germany, the United Kingdom (UK), France, Italy and Spain. These countries form a fairly diverse portrait of the richest nations in Europe so far as macroeconomic performance and fiscal position are concerned, and they are the largest economies in Western Europe. Their fiscal performance is critical for the whole region. From an econometric viewpoint, besides providing recent evidence on the tax-spend nexus for these countries, this article expands the existing literature by testing for changes in the direction of causality between revenue and expenditure. To this end, we employ Shi *et al.*'s (2018) causality test in a bivariate vector autoregression (VAR) model using quarterly time series for the period 1995–2019.

There are four hypotheses regarding the tax-spend nexus: (1) the tax-and-spend hypothesis, when changes in revenues temporally precede changes in expenditure (Friedman 1978, Buchanan and Wagner 1978); (2) the spend-and-tax hypothesis, when expenditure decisions come first and government revenues eventually adjust to it (Peacock and Wiseman 1979, Barro 1979); (3) the fiscal synchronization hypothesis, when changes in revenues and expenditures are jointly determined (Musgrave 1966, Meltzer and Richard 1981); and (4) the institutional separation hypothesis, which states that there is no temporal ordering of revenue and expenditure changes (Cameron 1978, Wildavsky 1988).

The empirical literature on the tax-spend nexus is vast and covers many countries. For those countries considered in this study, we found 11 articles that use recent data. Studies based on VAR and vector error correction models yielded mixed results (Chang *et al.* 2002, Castro *et al.* 2004, Kollias and Paleologou 2006, Saunoris and Payne 2010, Lusinyan and Thornton 2012, Paleologou 2013). The ones employing panel data models, which include other European countries, supported the fiscal synchronization hypothesis in the region when using more traditional approaches (Chang and Chiang 2009, Vamvoukas 2012) and different unidirectional causality hypothesis across countries when the Kónya (2006) bootstrap panel techniques are applied (Afonso and Rault 2009, Mutascu 2015). Despite their methodological differences, the common feature in these studies is the time-invariant structure of their econometric model. This may be a strong assumption because economic cycles and political changes generally affect the fiscal framework and consequently the tax-spend nexus. In fact, as Afonso and Rault (2009) found in their analysis of the subsamples 1960–1985 and 1986–2006, there was some shifting in the direction of the causality patterns after 1985, which might have resulted from adjustments of fiscal behavior in the run-up to Economic and Monetary Union.

The next section of this article presents the methodology. Section 3 describes the data set. Section 4 reports the econometric results. Section 5 summarizes the concluding observations.

# 2. Methodology

We test for causality in the tax-spend nexus based on the following bivariate P-th order VAR,

$$\begin{cases} r_t = \sum_{i=0}^{1} \varphi_{10}(i) t^i . + \sum_{i=1}^{P} \varphi_{11}(i) . r_{t-i} + \sum_{i=1}^{P} \varphi_{12}(i) . s_{t-i} + e_{r,t-i} , \\ s_t = \sum_{i=0}^{1} \varphi_{20}(i) t^i . + \sum_{i=1}^{P} \varphi_{21}(i) . r_{t-i} + \sum_{i=1}^{P} \varphi_{22}(i) . s_{t-i} + e_{s,t-i} , \end{cases} \quad (1)$$

where $r_t$ and $s_t$ are government revenues and expenditures, $\varphi_{jk}(i)$ are coefficients, P is the lag length, and $e_{r,t-i}$ and $e_{s,t-i}$ are error terms, *t = 1, 2, …, T*. When expenditures are important in predicting future values of revenues, then expenditures are said to Granger-cause revenues, and vice versa. It is important to emphasize that Granger causality is based on the forecasting criteria of Granger and does not necessarily imply behavioral causation.

In the system of equation (1), when $r_t$ and $s_t$ are stationary, the null hypotheses of no causality are:

$$H_{(r \nrightarrow s),0} : \varphi_{21}(i) = 0 \; for \; all \; i \quad (r_t \; DNGC \; s_t), \quad (2)$$

$$H_{(s \nrightarrow r),0} : \varphi_{12}(i) = 0 \; for \; all \; i \quad (s_t \; DNGC \; r_t), \quad (3)$$

where $DNGC$ means does not Granger cause. Rejection of only $H_{(r \nrightarrow s),0}$ (or $H_{(s \nrightarrow r),0}$) is evidence of the tax-and-spend (or the spend-and-tax) hypothesis. Rejection of both $H_{(r \nrightarrow s),0}$ and $H_{(s \nrightarrow r),0}$ implies fiscal synchronization, and a failure to reject both null hypotheses means institutional separation.

We determine and date changes in Granger causality based on the recursive evolving supremum Wald test (RE) developed recently by Shi *et al.* (2018).[1] To explain the RE testing procedure, consider the VAR model (1) written as

$$y_t = \Phi x_t + e_t, \quad (4)$$

where $y_t = (r_t, s_t)'$, $x_t = (1, t, y'_{t-1}, y'_{t-2}, \dots, y'_{t-P})'$, $e_t = (e_{r,t}, e_{s,t})'$, and $\Phi_{2\times(2P+2)} = [\varphi_0, \varphi(1), \dots, \varphi(P)]$.

The test proceeds as follows. Suppose λ is some fractional observation of interest and the $\lambda_0$ is the minimum window size, as a percentage of the whole sample, set to estimate the VAR model (1). The first observation of interest is, therefore, $T_{\lambda_0} = \lfloor T\lambda_0 \rfloor$, where $\lfloor . \rfloor$ denotes the integer part function. The RE window test is based on the supremum of a series of recursively calculated Wald statistics $W_\lambda^{(i \nrightarrow j)}(l_\kappa)$, $i = r, s, j = r, s, i \neq j$. Considering the case of $H_{(r \nrightarrow s),0}$, a set of Wald statistics $\left\{ W_\lambda^{(r \nrightarrow s)}(l_1), \dots, W_\lambda^{(r \nrightarrow s)}\left(l_{(\lambda-\lambda_0)}\right) \right\}$ is computed for each $\lambda \in [\lambda_0, 1]$, where

$$W_\lambda^{(r \nrightarrow s)}(l_\kappa) = \left[R_{rs} vec\left(\widehat{\Phi}_\kappa\right)\right]' \left[R_{rs}\left(\widehat{\Omega}_\kappa \otimes (X_\kappa' X_\kappa)^{-1}\right) R'_{rs}\right]^{-1} \left[R_{rs} vec\left(\widehat{\Phi}_\kappa\right)\right], \quad (5)$$

[1] Shi et al. (2018) examined three similar time varying Granger causality tests: the forward, rolling and recursive evolving tests. Based on their simulations, for a sample size close to this study's, Shi et al. (2018) suggested the RR testing approach is the most balanced, with false detection proportions and successful detection rates generally between the other tests.

$l_\kappa$ is the subsample $[\kappa, T_\lambda]$, $\kappa = 1,2, \dots, \lfloor T(\lambda - \lambda_0) \rfloor$, $R_{rs}$ is the P×(2P+2) matrix of values 0 and 1 that set the corresponding coefficients of Φ to zero under $H_{(r \nrightarrow s),0}$, $vec(\widehat{\Phi}_\kappa)$ contains the row vectorized 2(2P+2)×1 coefficients of Φ estimated for $l_\kappa$, $X_\kappa$ is the matrix of the regressors in (4) observed in $l_\kappa$, and $\widehat{\Omega}_\kappa$ is the least squares estimate of the error covariance matrix in $l_\kappa$.

The sup Wald test statistic of $H_{(r \nrightarrow s),0}$ for the observation $T_\lambda$ is, therefore, the supremum taken over all the test statistics in the set $\left\{W_\lambda^{(r \nrightarrow s)}(l_1), \dots, W_\lambda^{(r \nrightarrow s)}(l_{(\lambda-\lambda_0)})\right\}$,

$$SW_\lambda^{(r \nrightarrow s)} = sup\left\{W_\lambda^{(r \nrightarrow s)}(l_\kappa) : \kappa \in [1, (\lambda - \lambda_0)]\right\}. \quad (6)$$

As the observation of interest moves from $T_{\lambda_0}$ to $T$, a sequence of supremum Wald statistics $SW_\lambda^{(r \nrightarrow s)}$ is calculated for $\lambda \in [\lambda_0, 1]$. In the case of $H_{(s \nrightarrow r),0}$, the statistics $W_\lambda^{(s \nrightarrow r)}(l_\kappa)$ and a sequence of supremum Wald statistics $SW_\lambda^{(s \nrightarrow r)}$ are computed in the same fashion. The sequences of $SW_\lambda^{(r \nrightarrow s)}$ and $SW_\lambda^{(s \nrightarrow r)}$ are then employed to determine the intervals in which $H_{(r \nrightarrow s),0}$ and/or $H_{(s \nrightarrow r),0}$ are rejected. For instance, if the sequence of only $SW_\lambda^{(r \nrightarrow s)}$ (or $SW_\lambda^{(s \nrightarrow r)}$) exceeds its corresponding critical values in the interval $[T_a, T_b]$, we reject $H_{(r \nrightarrow s),0}$ (or $H_{(s \nrightarrow r),0}$), and thus have support for the tax-and-spend (or the spend-and-tax) hypothesis in this period.

The minimum window sizes $\lambda_0$ are set equal to fractions that correspond to 20 quarters of data for all countries. The VAR lag order for each subsample $l_\kappa$ used in the Wald statistic calculations is allowed to change and is selected by the Akaike information criterion from a maximum of four lags. We use the heteroskedastic consistent version of the Wald statistics,

$$SWH_\lambda^{(i \nrightarrow j)} = sup\left\{WH_\lambda^{(i \nrightarrow j)}(l_\kappa) : \kappa \in [1, (\lambda - \lambda_0)]\right\}, \quad (7)$$

where

$$WH_\lambda^{(i \nrightarrow j)}(l_\kappa) = T(\kappa)\left[R_{ij} vec(\widehat{\Phi}_\kappa)\right]'\left[R_{ij}\left(\hat{V}_\kappa^{-1} \widehat{W}_\kappa \hat{V}_\kappa^{-1}\right) R'_{ij}\right]^{-1}\left[R_{ij} vec(\widehat{\Phi}_\kappa)\right], \quad (8)$$

$i = r, s, j = r, s, i \neq j$, $\hat{V}_\kappa \equiv I_n \otimes \hat{Q}_\kappa$ with $\hat{Q}_\kappa \equiv \left(\sum_{t=\kappa}^{T_\lambda} x_t x'_t\right)/T(\kappa)$, $\widehat{W}_\kappa \equiv \left(\sum_{t=\kappa}^{T_\lambda} \hat{\varepsilon}_t \hat{\varepsilon}'_t\right)/T(\kappa)$ with $\hat{\varepsilon}_t \equiv \hat{e}_t \otimes x_t$, and $T(\kappa)$ is the size of the subsample $l_\kappa$. We also use the residual-based bootstrap critical values (estimated from 1,000 replications). Although the bootstrap critical values are greater than the standard asymptotic critical values of the Wald statistic, Shi *et al.* (2018) suggest using the former in small samples. The VAR model and sup Wald statistics are estimated by running the MATLAB codes provided by the authors.

# 3. Data and Unit Root Tests

The data comprise seasonally adjusted quarterly time series for total government revenues and expenditures (including interest payments) as a share of GDP collected from the Eurostat online database (http://www.econstat.com). Following other studies, we use the revenues and expenditures as a share of GDP to account for fluctuations of aggregate economic activity. The data cover the periods January 1995-June 2019 for France, Spain and the UK;

October 1999-June 2019 for Italy; and October 2000-June 2019 for Germany. A plot of the variables results is presented in Figure 2 (Appendix).

We assess the stationarity properties of $r_t$ and $s_t$ to verify whether the restrictions imposed in $H_{(r \nrightarrow s),0}$ and $H_{(s \nrightarrow r),0}$ and their corresponding test statistics are adequate to test the hypotheses of no causality in the VAR model (1). Note, however, that the presence of structural breaks in these series, which can be anecdotally perceived in Figure 2 for the 2008 financial crisis, would bias traditional unit root tests toward a false unit root null hypothesis. Therefore, we employ the two most traditional unit root tests, augmented Dickey-Fuller (ADF) and Philips-Perron (PP) tests and the Perron and Vogelsang (1992) (PV) unit root test with one unknown structural break. The tests are performed considering the intercept plus linear trend model and, in the case of PV, one innovation outlier structural break. The unit root test statistics are reported in Table 1. Overall, the results suggest that all revenue series are stationary. For the expenditure series, however, stationarity is found only when accounting for one structural break. The break dates indicated by the tests were mostly concentrated near or in 2008.

**Table 1.** Unit root test results

| | Revenues | | | Expenditures | | |
|---|---|---|---|---|---|---|
| Country | ADF | PP | PV | ADF | PP | PV |
| Germany | −4.26* | −4.32* | −6.00* (12/2008) | −2.17 | −3.27*** | −5.46* (06/2008) |
| Spain | −3.32*** | - | −5.68* (09/2007) | −2.42 | −1.86 | −5.04** (06/2008) |
| France | −2.42 | −4.49* | −4.95** (12/2010) | −2.89 | −2.25 | −5.32* (03/2008) |
| Italy | −2.91 | −4.42* | −5.65* (03/2012) | −2.77 | −3.62** | −6.32* (09/2008) |
| United Kingdom | −4.68* | −4.56* | −6.50* (6/2008) | −0.65 | −0.93 | −7.99* (3/2008) |

Note: *, ** and *** denotes rejection of the unit root null hypothesis for 1%, 5% and 10% levels. The dates (month/year) below the PV statistics are break dates selected by minimizing the Dickey-Fuller *t*-statistic.

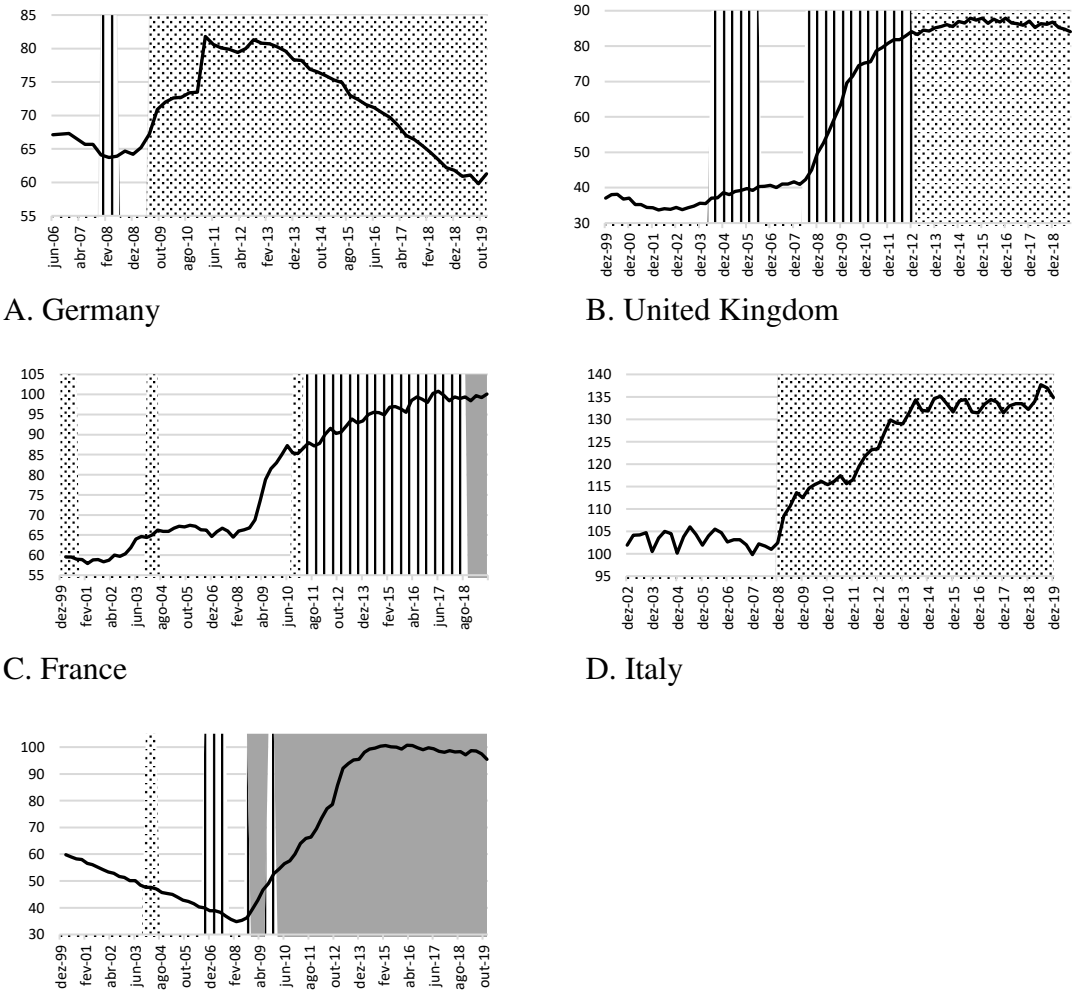
# 4. Results

For the sake of simplicity and space, instead of presenting the RE test results for $H_{(r \nrightarrow s),0}$ and $H_{(s \nrightarrow r),0}$ separately, Figure 1 displays the time intervals with their corresponding tax-spend hypotheses determined by combining the conclusions of each null hypothesis, shown as texturized areas (the non-texturized area refers to the institutional separation hypothesis). For instance, the full grey areas in panel E (Spain) show the time intervals in which both $H_{(r \nrightarrow s),0}$ and $H_{(s \nrightarrow r),0}$ were rejected, supporting the fiscal synchronization hypothesis. In another example, the non-texturized area in panel D (Italy) indicates that the institutional separation hypothesis is supported for the period before the 2008 crisis ($H_{(r \nrightarrow s),0}$ and $H_{(s \nrightarrow r),0}$ are not rejected), and the dotted area indicates that the tax-and-spend hypothesis is supported in the remainder of the sample ($H_{(r \nrightarrow s),0}$ is rejected, and $H_{(s \nrightarrow r),0}$ is not rejected).

The graphs also show the evolution of the debt-to-GDP ratios, collected from the Eurostat online database, to confront the results with the fiscal condition in each country or at least one aspect of it. Note that the debts of France, Spain, Italy and the UK show similar evolution patterns. There is a slowdown in the pace of increase in the debt-to-GDP ratios probably due to the 2014 Fiscal Compact, which introduced a permanent numeric budget rule

into government budget national laws. The main difference in Germany (panel A) is the pronounced reversal of its debt-to-GDP ratio trend already apparent by 2011.

**Figure 1.** Recursive evolving causality test results and debt-to-GDP ratio evolution




Notes: The graphs show the results for the period starting at $T_{\lambda_0}$, where the calculation of RE test statistics begins.To simplify the exposition of the results even further, we "clear out" the graphs from isolated very short periods (one/two quarters) in which $H_{(r \nrightarrow s),0}$ or $H_{(s \nrightarrow r),0}$ were rejected. These occurrences were very rare in the graphs. The vertical stripes indicate the intervals in which only $H_{(s \nrightarrow r),0}$ is rejected, indicating that only expenditures Granger cause revenues (the Spend-and-Tax hypothesis), the dotted bars indicate the intervals in which only $H_{(r \nrightarrow s),0}$ is rejected, indicating that only revenues Granger cause expenditures (the Tax-and-Spend hypothesis), and the grey bars indicate the intervals in which $H_{(r \nrightarrow s)0}$ and $H_{(s \nrightarrow r),0}$ are rejected, supporting the fiscal synchronization hypothesis. The periods without texturizes bars support the institutional separation hypothesis. The numbers in vertical axes are percentages.

For most of the countries, there were several periods in which government revenues and expenditures have no Granger-causal relation, mostly in the years before 2009. In this case, it seems that the institutional separation hypothesis probably dominated the budgetary processes in these countries as they were (loosely?) trying to adjust their fiscal policy to meet the restriction conditions established by the SGP in 1997. Perhaps, led by the good growth prospects between 1997 and 2007, the executive efforts in search of fiscal consolidation (to ensure proper functioning of the European Union) were to some extent moderated, while the

increasing demand for public expenditures dominated the actions taken by politicians and legislative bodies. Some rules in the SGP were even softened in 2003, with greater discretion, leniency and political control written into the agreement (Shucknecht *et al.* 2011). The government institutions or entities affected the budgetary process more independently, as Wildavsky (1988) and Cameron (1978) have argued. After the 2008 crises, when governments needed greater fiscal coordination and control over the budgetary process, causality links in the relation between government revenues and expenditures emerged for all countries and stayed active henceforth.

In Germany (panel A), the direction of causality shifted from the spend-and-tax hypothesis to the tax-and-spend hypothesis after a brief period of institutional separation around 2008/2009. During the recent tax-and-spend regime, the German government adopted a fiscal course of great commitment to budget discipline. The "debt brake" policy in 2009, which limited the deficit-to-GDP ratio to 0.35%, was reinforced with several administrative reforms and Parliamentary engagement to ensure government spending remained under tight control while avoiding tax increases. This, in fact, is the course of action recommended by Friedman (1978) when revenues temporally precede expenditures. It may be a coincidence, but German austerity measures in the tax-and-spend regime greatly reduced its debt-to-GDP ratio after 2009.

Recent data for the UK (Panel B) and Italy (Panel D) also support the tax-and-spend hypothesis after the 2008 crisis. Whereas in Italy this change occurred after an institutional separation period before the crisis, in the UK, the new tax-and-spend regime started after another period of spend-and-tax emphasis, approximately from 2008 to 2012. Both countries also committed to fiscal austerity, but with more moderate measures than were employed in Germany. The UK, with the new conservative-led government in 2010, implemented a deficit-reduction program that cut spending in several areas but preserved the National Health Service and the education system, responsible for a large part of their expenditures.

Despite having one of the highest debt-to-GDP ratios in Europe, Italy also initially opted for a slower fiscal adjustment process. Ratifying the Fiscal Compact, the Italian Parliament approved a new set of budget rules with constitutional status in April 2012 (which came into force in January 2014). It stipulates not only a deficit and debt rule but also an expenditure rule, whereby the annual growth rate of general government expenditures must not exceed the limit laid down by EU regulations. While the focus of their policies was on spending cuts, as Friedman (1978) proposed, it seems that the austerity measures in the UK and Italy were only enough to control their increasing deficits, as their debt-to-GDP ratios remained relatively high.

Panel C shows that, in France, three short periods support the tax-and-spend hypothesis, approximately around 1999-2000, 2003–2004 and 2010-2011, and a post-2012 period supports the spend-and-tax hypothesis, a picture quite the opposite of what was found for the UK. The spend-and-tax dynamic seems to have started with the new fiscal framework implemented by the former Hollande government in 2012–2017, which intended to manage the budget deficits through higher taxes while attending the current demand for government spending. This revenue adjustment in response to changes in spending is compatible with the tax smoothing hypothesis of Barro (1979). However, in the context of the spend-and-tax hypothesis, according to Peacock and Wiseman (1979) and Barro (1979), the government should implement policies focusing on spending cuts. In fact, after missing deficit targets, the French government revaluated its fiscal policy and start working more on spending cuts, but this seemed to be insufficient to revert the positive trend in the public debt-to-GDP ratio. It is worth pointing out that France implemented the Fiscal Compact by adopting an organic law in late 2012, which imposed low barriers to alter or override the law. In addition, the reform did

not codify a clear rule to prevent government deficit or to impose a yearly balanced budget (Fabbrini, 2013).

In Spain (panel E), there were a few brief occurrences of Granger causality between revenues and expenditures before the 2008 crisis. Although the Spanish economy grew consistently at that time, it is still interesting to find that in this period when the institutional separation hypothesis dominated, the Spanish government managed to control its deficit and significantly reduce its debt-to-GDP ratio. For almost the entire period after 2008, Spanish data support the fiscal synchronization hypothesis (Musgrave 1966 Meltzer and Richard 1981). This is in line with fiscal policies implemented during the early premiership of Mariano Rajoy, which initiated an austerity program consisting of tax increases, as well as cuts in wages and benefits, combined with decreases in the public budget and the execution of several legal reforms to reduce public spending, such as reforms in the pension system and labor market. In addition, the Spanish Parliament enshrined budgetary discipline in a constitutional reform (January 2012) which imposed limits on general government debt and structural deficits. Such coordination between expenditure-reducing and revenue-raising policies was only able to slightly decrease the debt-to-GDP ratio past 2013.

Finally, to examine the sensitivity of the RE test results to a larger minimum window size parameter, we performed further analysis considering values of $\lambda_0$ that correspond to 28 quarters of data. Shi *et al.* (2018) found that increasing $\lambda_0$ may result in fewer detected causality episodes, but it may also lead to lower bootstrap critical values, decreasing the problem of "non rejection of the false null." The results are presented in Figure 3 (Appendix). We verify that the additional tests using larger minimum window sizes are rather consistent with the initial test.

## 5. Conclusion

Using the recursive evolving supremum Wald test (RE) test of Shi *et al.* (2018), we found several changes in the direction of causality between government revenues and expenditures for Germany, the UK, France, Italy and Spain. Before the 2008 financial crisis, the institutional separation hypothesis seems to have dominated the fiscal framework in all countries analyzed. Nevertheless, after 2008, as the countries were restructuring their public finances based on fiscal policies implemented to mitigate the negative effects of the crisis, lasting causality links developed for all countries. The recent data support the tax-and-spend hypothesis for Germany, the UK and Italy; the spend-and-tax hypothesis for France; and the fiscal synchronization hypothesis for Italy.

# 6. References


Afonso, A. and Rault, C. (2009) "Spend-and-tax: a panel data investigation for the EU" *Economics Bulletin*. **29(4)**, 2542–2548.

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

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

Cameron, D. R. (1978) "The expansion of the public economy: A comparative analysis" *American Political Science Review* **72(4)**, 1243-1261. doi:10.2307/1954537

Castro, Francisco; González-Páramo, José; Cos, Pablo (2004) "Fiscal consolidation in Spain: Dynamic interdependence of public spending and revenues" *Investigaciones Economicas* **28**, 193-207.

Chang, Tsangyao and Chiang, Gengnan (2009) "Revisiting the government revenue-expenditure nexus: Evidence from 15 OECD countries based on the panel data approach" *Czech Journal of Economics and Finance* **59(2)**, 165-172.

Chang, Tsangyao; Liu, Wen Rong; Caudill, Steven B. (2002). "Tax-and-spend, spend-and-tax, or fiscal synchronization: new evidence for ten countries" *Applied Economics* **34(12)**, 1553-1561. DOI: 10.1080/00036840110103265

Fabbrini, Federico. (2013) "The Fiscal Compact, the "Golden Rule," and the Paradox of European Federalism*" Boston College International and Comparative Law Review* **36(1)**. https://lawdigitalcommons.bc.edu/iclr/vol36/iss1/1/

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

Kollias, C. and Paleologou, S. (2006) "Fiscal policy in the European Union: Tax and spend, spend and tax, fiscal synchronization or institutional separation?" *Journal of Economic Studies* **33(2)**, 108 - 120. https://doi.org/10.1108/01443580610666064

Kónya, L. (2006) "Exports and growth: Granger causality analysis on OECD countries with a panel approach" *Economic Modelling* **23(6)**, 978-992.

Lusinyan, L. and Thornton, J. (2012) "The intertemporal relation between government revenue and expenditure in the United Kingdom, 1750 to 2004" *Applied Economics* **44(18)**, 2321–2333. DOI:10.1080/00036846.2011.564142

Meltzer, A. H. and Richard, S. F. (1981) "A rational theory of the size of government" *Journal of Political Economy **89***, 914–927.

Musgrave, R. (1966) "Principles of budget determination" In *Public finance: Selected readings* by H. Cameron, & W. Henderson, Eds., Random House: New York, 15–27.

Mutascu, Mihai (2015) "A bootstrap panel Granger causality analysis of government revenues and expenditures in the PIIGS countries" *Economics Bulletin* **35(3)**, 2000-2004.

Owoye, O. and Onafowora, O. A. (2011) "The relationship between tax revenues and government expenditures in European Union and Non-European Union OECD Countries" *Public Finance Review* **39(3)**, 429–461. https://doi.org/10.1177%2F1091142110386211

Paleologou, S-M. (2013) "Asymmetries in the revenue–expenditure nexus: A tale of three countries" *Economic Modelling* **30**, 52–60.

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

Perron, P., and Vogelsang, T. (1992) "Testing for a unit root in a time series with a changing mean: Corrections and extensions" *Journal of Business & Economic Statistics* **10(4)**, 467-470. DOI:10.2307/1391823

Saunoris, J. and Payne, J. E. (2010) "Tax more or spend less? Asymmetries in the UK revenue-expenditure nexus" *Journal of Policy Modeling* **32**, 478-487.

Shi, S.; Phillips, P.C.B; Hurn, S. (2018) "Change Detection and the Causal Impact of the Yield Curve" *J. Time Ser. Anal*., **39**, 966-987. DOI:10.1111/jtsa.12427

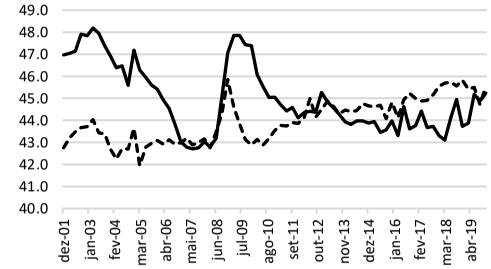
Schuknecht, L., Moutot, P., Rother, P., Stark, J. (2011) "The Stability and Growth Pact: Crisis and Reform" *Occasional Paper Series* Nº 129, European Central Bank, September.

Vamvoukas, G. A. (2012) "Panel data modelling and the tax-spend controversy in the Euro Zone" *Applied Economics* **44(31)**, 4073–4085.

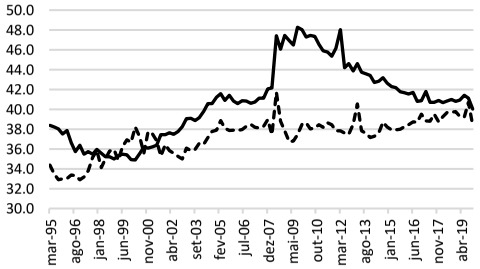
Wildavsky, A. (1988). *The New Politics of Budgetary Process*, Glenview: Scott Foresman.

# Appendix

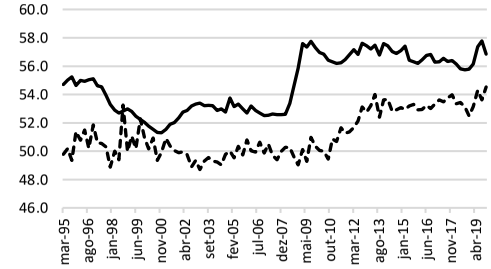
**Figure 2.** Time series for total government revenues and expenditures as shares of GDP



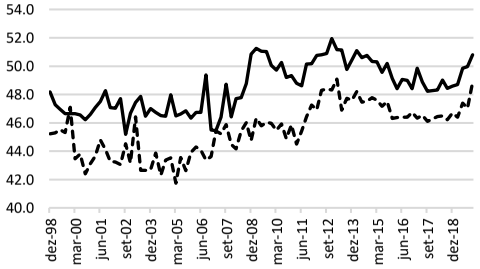

A. Germany



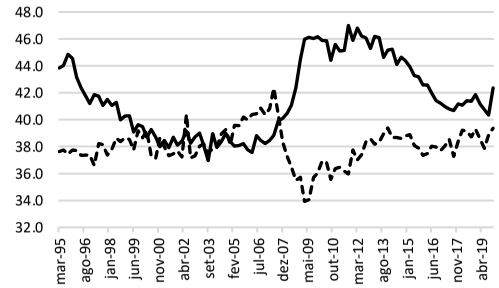

B. United Kingdom




C. France




D. Italy




D. Spain

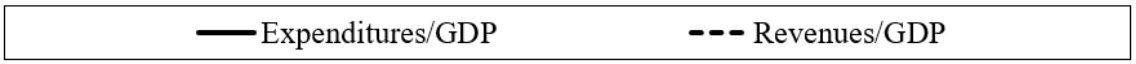
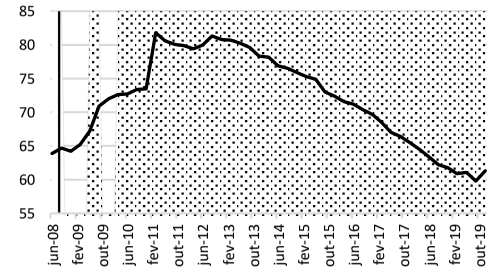
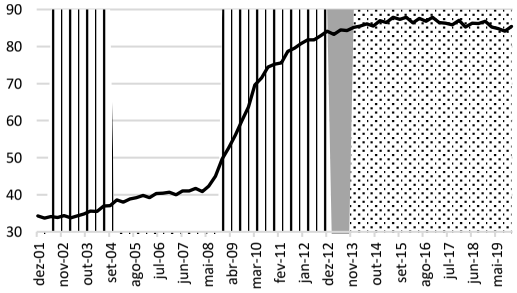


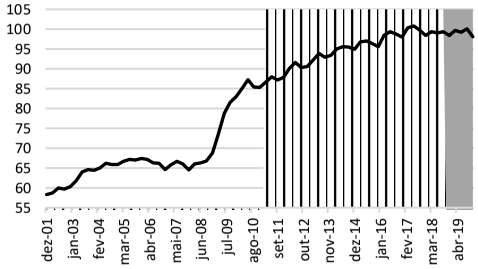
**Figure 3.** Recursive evolving causality test results for minimum window sizes ($\lambda_0$) that correspond to 28 quarters of data



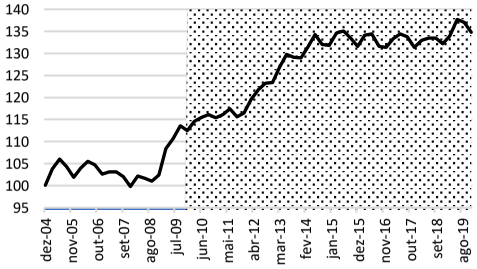

A. Germany



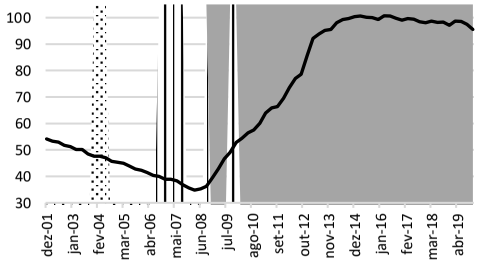

B. United Kingdom




C. France




D. Italy




D. Spain