THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 54_12

Gabriella Legrenzi
Keele University, UK
CESifo, Germany
The Rimini Centre for Economic Analysis (RCEA), Italy

Costas Milas
The University of Liverpool, UK
The Rimini Centre for Economic Analysis (RCEA), Italy
Eranistis.eu, Greece

FISCAL POLICY SUSTAINABILITY, ECONOMIC CYCLE AND FINANCIAL CRISES: THE CASE OF THE GIPS

The Rimini Centre for Economic Analysis
Legal address: Via Angherà, 22 – Head office: Via Patara, 3 - 47900 Rimini (RN) – Italy
www.rcfea.org - secretary@rcfea.org

Fiscal Policy Sustainability, Economic Cycle and Financial Crises: the Case of the GIPS.

Gabriella Legrenzi*
Keele University, CESifo and Rimini Centre for Economic Analysis

Costas Milas
Liverpool University, Rimini Centre for Economic Analysis and Eranistis.eu

July 6, 2012

Abstract

We extend previous work on the sustainability of the government's intertemporal budget constraint by allowing for non-linear adjustment of the fiscal variables, conditional on (i) the sign of budgetary disequilibria and (ii) the phase of the economic cycle. Further, our endogenously estimated threshold for the non-linear adjustment is not fixed; instead it is allowed to vary over time and during financial crises. Our analysis presents particular interest within the current economic scenario of financial crises, poor growth and debt crises. Our empirical analysis, applied to the GIPS, shows evidence of a threshold behaviour for the GIPS, that only correct "large" unbalances, which, in the case of Greece and Portugal, are higher than the EGSP criteria. Financial crises further relax the threshold for adjustment: during financial crises, only "very large" budgetary unbalances are corrected.

JEL Classification: H63, H20, H60, C22

Keywords: debt sustainability, fiscal adjustment, nonlinear models

1 Introduction and Background.

The recent (and arguing ongoing) financial crisis has given rise to remarkable fiscal stimuli packages in an attempt to revive the world economy. While the total impact of such packages on the global economy is still to be fully assessed,

*Previous versions of this paper have been presented at the International Finance and Banking Society Annual Congress, the CESifo Workshop in Public Sector Economics and the Rimini Centre for Economic Analysis Annual Conference. We thank participants for useful feedback. The usual disclaimer applies. Financial support from CESifo is gratefully adknowledged.

there is a growing concern that some countries are creating excessive deficits and debt, resulting in an unsustainable fiscal policy path. In particular, within the European Monetary Union four peripheral countries are currently considered to be in a weaker position with respect to the sustainability of their public finances, namely Greece, Ireland, Portugal and Spain, denoted by the acronym of GIPS.

Greece was bailed-out twice (for €110bn in May 2010 and then again for €109bn in July 2011). In October 2011 and then again in February 2012, Greece negotiated a new €130bn rescue package involving a voluntary "haircut" of some 53.5% on the face value of its bonds held by the private sector. Ireland was bailed-out once (for €85bn in November 2010) and Portugal was also bailed once (for €78bn in May 2011) by the European Union, the European Central Bank and the International Monetary Fund. Spain requested, in June 2012, financial assistance from the European Financial Stability Facility for €100bn.

At the same time, to relieve tensions in financial markets, the ECB, has purchased since 2010 over 200 billion euros in marketable debt instruments (both private and public), under the Securities Market Programme[1]. The ECB has also purchased, under the Covered Bonds Purchase Programme, 60 billion euro in European government bonds[2], announcing plans for a further 40 billion euros[3]. Tighter fiscal rules as well as a possible fiscal union are currently being discussed, and international support to GIPS has been conditional to increasing supranational control over national fiscal policies.

Despite the bail-outs, international markets remain extremely volatile and worried that the debt levels of all GIPS are unsustainable, posing a risk to the

[1]Decision of the ECB of 14 May 2010, establishing a Securities Markets Programme (ECB/2010/5)

[2]ECB Press Release, 30.06.2010.

[3]ECB Press Release, 03.11.11.

whole Eurozone.

As the GIPS economies, taken together, account for around 17% of Eurozone's GDP, a detailed analysis on the sustainability of their public finances, as well as a comprehensive account of the behavior of their fiscal policy variables in different phases of the economic cycle and during debt and financial crises, becomes particularly relevant. This in order to monitor the fiscal health of Eurozone countries and minimize, at an early stage, possible threats to Eurozone's stability, as well as to limit any spill-over effects to the rest of the world. On the contrary, existing literature on fiscal policy sustainability has mainly focussed on the intertemporal budget constraint (IBC) and the long-run properties of the fiscal variables (see, e.g. Hamilton and Flavin, 1986, Hakkio and Rush, 1991, Trehan and Walsh, 1991, Quintos, 1995). Further, most of the existing literature on sustainability has implicitly considered a linear adjustment process. This assumption is too restrictive, as fiscal variables might adjust in a discrete fashion (see on this, Bertola and Drazen 1993), only when the deficit becomes "too large", due to the difficulties to obtain the necessary consensus for reforms. A limited empirical literature (Arghyrou and Liuntel, 2007, Bajo Rubio *et al.*, 2006), has considered the fiscal adjustment process, restricting the fiscal adjustment to follow the same process during good and bad times.

As a consequence, existing empirical literature might prove inadequate to understand the current economic environment of fiscal adjustments which are taking place despite the period of very low economic growth.

The main contribution of our paper to the existing empirical literature on fiscal policy sustainability relies on introducing non-linear tests conditional on (i) the sign and magnitude of fiscal disequilibria and (ii) the phase of the economic cycle. Further, our endogenously estimated threshold for the fiscal adjustment is not fixed; instead it is allowed to vary over time, in particular with the incidence of financial crises. This allows us to investigate whether fiscal adjustment is relaxed during periods of financial stress.

Existing empirical evidence on the sustainability of the GIPS's IBC provides ambiguous results. Afonso (2005) shows that the revenues and expenditure of the GIPS (and most of the other European countries) are not cointegrated, pointing to IBC unsustainability. Similar results are obtained by Santos-Bravo and Silvestre (2002) for Ireland and Portugal. Greiner et al. (2007) show that the public finances of Portugal are sustainable, based on Bohn's (1998) fiscal reaction function. Based on a panel of the EU-15 countries, Afonso and Raut (2010) show that, taken together, the European fiscal policies are sustainable.

All above tests are conducted under the assumption of a linear relationship between government revenues and spending and/or debt and primary surpluses, which might prove too restrictive. Bajo-Rubio et al. (2006) show that the Spanish fiscal policy is sustainable, within a non-linear revenues/expenditures approach. Arghyrou and Luintel (2007), find that the Greek and Irish public finances satisfy the weak form sustainability, with positive structural shifts associated with the Maastricht Treaty.

This work is organized as follows. Section 2 reports our empirical analysis of the long-run sustainability. Section 3 presents our analysis of the year-to-year adjustments of the fiscal variables, considering their behavior in different phases of the economic cycle and during financial crises. Section 4 concludes and provides directions for further research. Variables used for the empirical analysis are defined in Appendix 1.

2 Are the GIPS respecting their IBC?

2.1 Long-run Estimates.

We address the following questions. First, do GIPS respect their IBC? Second, do taxes and expenditures equally carry the burden of fiscal policy adjustments?

To answer the first question, we initially run linear cointegration tests, based on Quintos (1995). To allow for potential endogeneity of fiscal variables, cointegration tests are performed by estimating a Vector Error Correction Model (VECM; see Johansen, 1988) of the form:

$$\Delta y_t = \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-1} + \Pi y_{t-1} + \mu + \varepsilon_t \tag{1}$$

where $y_t = [TAX/GDP, G/GDP]$. TAX/GDP is the general government total revenues, G/GDP is the general government total outlays, both in GDP ratios; $\varepsilon_t \sim niid(0, \Sigma), \mu$ is a drift parameter, and Π is a $(p * p)$ matrix of the form $\Pi = \alpha\beta'$, where α and β are $(p * r)$ matrices of full column rank, with β containing the r cointegrating vectors and α carrying the corresponding loadings in each of the r vectors. For each country, the lag length k is set as to minimize the Akaike Information Criterion. The test for cointegration is conducted in each case using Johansen's (1988) maximal eigenvalue (λ-max) and trace (λ-trace) statistics. To account for our small sample, both tests use a small sample correction (for exact mathematical formulas, see e.g. Doornik and Hendry, 2000, p.282).

Following Quintos' 1995 work, cointegration with an estimated cointegrating vector of (1, -1) suggests IBC sustainability. Absence of cointegration, but with an estimated cointegrating vector of (1,-1) suggests weak form sustainability. Weak exogeneity tests are then performed, in order to ascertain which variable (i.e. the average tax rate or the government share) carries the burden of the fiscal adjustment. This is particularly important given the possibility of non-keynesian effects of spending cuts (as opposed to tax increases), as in Alesina and Ardagna (1998).

For all countries considered, we use the longest available annual time series data from AMECO. Annual data are preferred in this analysis, as they allow us to consider discretionary decisions of fiscal policy authorities, which are not captured by higher frequency data. A full description of the AMECO series used for each country is presented in Appendix 1. Figure 1 provides, for each country, a plot of the series against time.

[INSERT FIGURE 1 HERE]

Preliminary analysis of TAX/GDP and G/GDP for all countries considered, using different unit root tests, suggests that all series are non-stationary in levels for all the GIPS[4]. We now turn our attention to the empirical results of the sustainability tests, which are reported in Table 1.

[INSERT TABLE 1 HERE]

Starting with Greece, the Johansen's tests show no evidence of sustainability, based on the λ-max statistic, and a weak evidence of cointegration based on the λ-trace statistics[5]. The estimated $b < 1$ further corroborates the sustainability problems. On the other hand, the adjustment coefficient for taxes is negative and statistically significant, whilst the adjustment coefficient on government spending is statistically insignificant, resulting in a weakly exogenous government spending.

[4] To save space, these results are not reported but are available on request.

[5] When the results of the $\lambda - trace$ and $\lambda - \max$ conflict, it is usually preferred to follow the latter to pin down the number of cointegrating vectors, as it has the sharper alternative hypothesis (Enders, 2010).

For the case of Ireland, no cointegration is found between revenues and expenditures, and the estimated $b < 1$ point to a sustainability problem. Further, government spending results weakly exogenous.

For the case of Portugal, there is some evidence of cointegration (stronger for the λ-trace statistic), with an estimated $b < 1$. On the other hand, government spending results weakly exogenous.

For the case of Spain, both the λ-trace and the λ-max show the absence of a long-run relationship between revenues and expenditures, pointing to an unsustainable IBC. The estimated $b < 1$, accompanied by a statistically insignificant adjustment coefficient for taxes, further corroborate the sustainability problems. On the other hand, the adjustment coefficient for spending is positive and statistically significant.

Taken all together, these tests point to serious sustainability problems for all GIPS, pointing to a risk in eurozone's stability.

2.2 Recursive Estimates.

Further useful insight on the sustainability of the GIPS's IBC can be obtained from the analysis of the recursively estimated λ-max and λ-trace statistics, divided by their critical value and plotted against time, for each country, in Figure 2. A value higher than one implies sustainability over the sample considered. We also recursively estimate, for each country, the estimated cointegrating vector, and the estimated adjustment coefficient for taxes and government spending. These are plotted in Figures 3 to 5.

Our recursive analysis can also shed some light on the current debate regarding the behavior of credit rating agencies. It has been argued that some downgradings of sovereign debt in the GIPS were not motivated by economic fundamentals[6] and might have been responsible for a worsening of the GIPS public finances, by raising the risk premia required by financial markets.

[INSERT FIGURE 2 HERE]

For Greece, we notice a clear unsustainability of the IBC, arising since 2002, i.e. shortly after the country's admission into the eurozone in 2001 (rather than 1999). Spain's and Ireland's sustainability problems trace back to 1996 and 1990, respectively, whilst there is no evidence of sustainability for Portugal over the entire sample.

[INSERT FIGURE 3 HERE]

The estimated marginal response of taxes to spending for Greece clearly worsens after its admission to the euro. The estimated cointegrating vector for the rest of the GIPS remain rather stable over time.

[INSERT FIGURE 4 HERE]

[INSERT FIGURE 5 HERE]

The joint analysis of the estimated adjustment coefficients for taxes and spending provide further useful insights. Since entering the euro, Greece has increased its tax-adjustment whilst government spending adjustment has mainly remained statistically insignificant.

Spain's tax adjustment has only slightly increased in the last 3 years, whilst the spending adjustment has been largely insignificant.

Ireland's tax adjustment has increased until the early 1990s and dropped slightly since then, accompanied by an insignificant government adjustment, pointing to a weakly exogenous spending.

[6] Speaking to the European parliament in May 2010, Jose Manuel Barroso, the EU Commission President, criticised heavily the three main credit rating agencies noting that deficiencies in their working methods has led to ratings being too cyclical, too reliant on the general market mood rather than on fundamentals.

Portugal's tax adjustment has also increased over time, accompanied by a spending adjustment that is largely insignificant

Overall, our recursive estimates show a clear sustainability problem, arising much earlier than the successive downgrades. This sustainability problem was apparent at the time of the successive rulings of the European Council abrogating previous excessive deficit rulings (2005 for Portugal, 2007 for Greece, and 2010 for Ireland), pointing to some ineffective monitoring from the EU. The same also applies to credit rating agencies. These are supposed to monitor economic fundamentals. Although our analysis implies that recent credit rating downgrades might not be directly responsible for a worsening of the GIPS's public finances, we can also argue that credit rating downgrades could have been put in place much earlier in time.

We also uncover that admission to the euro for Greece has coincided with a remarkable relaxation of its fiscal policy. Our finding of weak exogeneity of government spending in the GIPS points to a spend-and-tax model, where spending is decided by the political process, regardless of the needs of IBC sustainability, and the burden of correcting budgetary disequilibria is entirely left to the tax instrument. This is bound to cause further detriment to the GIPS economies, not captured by the standard sustainability tests.

3 Government Solvency, Nonlinearities and the Business Cycle.

As discussed in Section 1, a notable drawback of existing empirical evidence on the GIPS is that it relies on linear models, making the implicit assumption of a continuous and state-invariant fiscal adjustment. This means that fiscal authorities are expected to correct every imbalance, no matter if positive/negative and/or large/small, and to correct it exactly in the same way, regardless of the phase of the economic cycle where such adjustments are taking place and/or the incidence of financial crises.

Bertola and Drazen (1993) suggest instead that fiscal policy authorities correct fiscal imbalances only when they are too large. Their motivation relies on the difficulties in reaching the necessary consensus for fiscal retrenchment. Complementary evidence is found in Alesina and Drazen (1991), within a "war of attrition" model, explaining that stabilizations are delayed, even when the entire society would benefit from them, due to the difficulties in reaching the necessary agreement on how to spread the costs.

A further drawback of the linear approach relies on the fact that linear cointegration tests have been shown to have low power to detect threshold cointegration (see, e.g. Kapetanios et al. 2003). As a consequence, applied to our fiscal policy set up, traditional linear tests might mistakenly suggest that given countries are on a unsustainable fiscal policy path, whereas in fact their intertemporal budget constraint holds, but corrections only take place beyond a given threshold.

Existing non-linear tests, on the other hand, are based on a fixed threshold, and assume that the adjustment is invariant to economic cycle and financial crises incidences. We estimate our fiscal adjustment models relaxing the assumption of a fixed threshold, looking at the behavior of fiscal variables not only with respect to particular budgetary thresholds but also during different phases of the economic cycle and during financial crises. We believe this type of analysis will provide further insight on how "good" as opposed to "bad" times and incidences of financial crises have affected the adjustment of GIPS' fiscal policies.

3.1 Tax Policy Models.

To examine the issue of non-linear adjustment in the short run dynamics of taxes, we proceed by considering the non-linear model of the form

$$\begin{aligned}\Delta\left(\frac{TAX}{GDP}\right)_t &= \beta_0 + (\beta_{11}CV_{t-1} + \beta_{12}gap_{t-1})\,\theta_{t-1} + (\beta_{21}CV_{t-1} + \beta_{22}gap_{t-1})\,(1-\theta_{t-1}) + \\ &\quad + \beta_3 fincrisis_t + u_t \qquad (2)\end{aligned}$$

where CV_{t-1} are the residuals from the long-run relationship between $\frac{TAX}{GDP}$ and $\frac{G}{GDP}$ (i.e. $\frac{TAX}{GDP} - \beta\frac{G}{GDP}$), gap is the output gap (i.e. GDP detrended by a Hodrick-Prescott trend[7]), u_t is a stochastic error term, $u_t \sim i.i.d.\left(0, \sigma_u^2\right)$and

$$\theta_{t-1} = 1 - [1 + \exp(-\gamma^s(s_{t-1} - \tau^s)/\sigma_{s_{t-1}})]^{-1} \qquad (3)$$

is the logistic transition function discussed in e.g. van Dijk et al (2002)[8]. The variable $fincrisis$ is a composite measure of financial turmoil/crisis (which draws heavily on Reinhart and Rogoff, 2009). This is a world financial crisis measure which takes into account banking, currency, stock market, debt, and inflation incidences in the world. For a given country in a given year, the index is bounded between zero and five, emerging as the sum of the number of types of incidences the country experienced. Therefore, the index takes the value of 0 if the country did not experience any of the five incidences above and the value of 5 if it did experience all five incidences. The index (plotted in Figure 6) pools

[7] We have used a smoothing parameter λ equal to 100, as suggested by Hodrik and Prescott (1997) for annual data. For robustness, we also considered the Ravn and Uhligh (2002) suggested value of 6.25, obtained from $1600\left(\frac{1}{4}\right)^4$.

[8] In preliminary estimates we allowed for $\Delta\left(\frac{TAX}{GDP}\right)_t$ to depend (both in a regime-switching manner and by imposing common coefficients) on its lag, and the current and lagged value of $\Delta\left(\frac{G}{GDP}\right)_t$. With the exception of Portugal (results are reported below), we were unable to find any statistical effect of these variables for the remaining countries.

together world's 20 largest economies with country specific weights given by their relative GDP share of the total GDP (based on Purchasing Power Parity).

According to (2)-(3), tax policy exhibits regime-switching behavior which depends on whether the transition variable, s_{t-1}, is below or above an endogenously estimated threshold, τ^s with regime weights θ_t and $(1-\theta_t)$, respectively. When $(s_{t-1} - \tau^s) \to -\infty$, then $\theta_t \to 1$. In this case, the impact of CV_{t-1} and gap_{t-1} is given by β_{11} and β_{12}, respectively. When $(s_{t-1} - \tau^s) \to \infty$, then $\theta_t \to 0$. In this case, the impact of CV_{t-1} and gap_{t-1} is given by β_{21} and β_{22}, respectively. The parameter $\gamma^s > 0$ determines the smoothness of the transition regimes. We make γ^s dimension-free by dividing it by the standard deviation of s_{t-1}(Granger and Teräsvirta, 1993). We consider two possible candidates for s_{t-1}: CV_{t-1} and gap_{t-1}. In the first case, and under the assumption $\tau^{CV} < 0$, we assess how taxes adjust in periods of a rising deficit-to-GDP ratio (when $CV_{t-1} < \tau^{CV}$) as opposed to periods of a falling deficit-to-GDP ratio (when $CV_{t-1} > \tau^{CV}$). In the second case, we assess whether taxes adjust differently during periods of economic downturns (when $gap_{t-1} < \tau^{gap}$) and during periods of economic expansions (when $gap_{t-1} > \tau^{gap}$).

Assuming that budgetary corrective action in periods of a rising versus falling deficit-to-GDP ratio (CV_{t-1}) is dependent on a fixed threshold might be too restrictive; rather, corrective action might vary with the occurrence of financial crises. In this case,

$$\tau_t^{CV} = \tau_0^{CV} + \tau_1^{CV} fincrisis_t \tag{4}$$

where τ_0^{CV} is a fixed threshold and ($\tau_1^{CV} \gtreqless 0$). In this case, both $\tau_0^{CV} < 0$ and $\tau_1^{CV} < 0$ suggest that policymakers, driven by the fear of a deep and lasting recession in periods of financial crises, might be more willing to relax the threshold triggering a correction.

We start by reporting in column (i) of Tables 2-5 linear tax revenues error correction models for the GIPS. Our results suggest a low, but nevertheless significant, error correction for Greece (see Table 2(i)), significant, and twice as fast, error correction for Ireland (see Table 3(i)) and for Portugal (see Table 4(i)) and weak evidence of error correction for Spain (see Table 5(i)). Business cycle effects are significant only in the case of Greece; when the economy expands, tax revenues rise. The financial crisis variable suggests a significant negative effect for Spain and less so for Ireland.

For the linear models, we also report, at the bottom of each Table, the p-value of Hamilton's (2001) λ-test, and the bootstrapped p-value[9] of the λ_A and g-tests proposed by Dahl and González-Rivera (2003). Under the null hypothesis of linearity, these are Lagrange Multiplier test statistics following the χ^2 distribution. These tests are powerful in detecting nonlinear regime-switching behavior like the one considered in our paper. For all countries, all three tests reject linearity.

Column (ii) of Tables 2-5 reports the non-linear models (2)-(3) using CV_{t-1} as the transition variable.

For Greece (see Table 2(ii)), in line with our theoretical predictions, Greece fiscal authorities correct deficits only when they exceeded 4% of national GDP. On the other hand, below the 4% threshold, no corrective action is taken, but taxes only respond, positively, to the output gap.

For Ireland (see Table 3(ii)), we find evidence of budgetary corrections only when the deficit-to-GDP ratio exceeds a 1.2% threshold, while the adjustment

[9] Based on 1000 resamples.

below the threshold is statistically insignificant. The output gap is insignificant in both regimes. This points to a threshold behavior of fiscal policy authorities.

For Portugal (see Table 4(ii)), taxes adjust when the deficit-to-GDP ratio exceeds the 4.9% threshold. The output gap is statistically insignificant in both regimes. Once again, this points to a threshold behavior of fiscal policy authorities.

For Spain (see Table 5(ii)), we find no evidence of tax adjustments for deficit correction, as the endogenously estimated threshold is positive. That is, corrective action is taken for budgetary surpluses only. Fiscal policy is acyclical when above the threshold, and countercyclical below.

Overall, the non-linear model allowing for the deviations from budgetary equilibria (in GDP ratios) as possible transition variable provides evidence of threshold behavior in the fiscal policy of all GIPS. In line with the theoretical literature on fiscal adjustments, the GIPS do not correct deficits when they are "too low". The endogenously estimated threshold for adjustment varies considerably from country to country, implying a deficit to GDP ratio of 4.9% for Portugal, 4.1% for Greece and 1.20% for Ireland. We also note that the threshold for correction is, for Greece and Portugal, higher than what required by the European Growth and Stability Pact.

Column (iii) of Tables 2-5 reports the non-linear models (2)-(3) using the output gap as possible transition variable. The near zero estimates of the threshold parameter suggest regime-switching with respect to positive versus negative deviations from trend output.

For Greece (see Table 2(iii)), the response of the average tax ratio to the output gap is positive and very similar during good times (characterized by the regime where the output gap is positive) and bad times. At the same time, during good times the error correction adjustment is statistically insignificant. This means that when the Greek economy grows above trend, no correction for budgetary disequilibria is taken.

For Ireland, the correction of budgetary disequilibria appears similar in both regimes, but taxes become procyclical during good times; for Portugal, we provide evidence of budgetary correction during bad times only, whilst the output gap is statistically insignificant in both regimes; for Spain, budgetary disequilibria are corrected only during good times, and taxes respond positively to the economic cycle.

Overall the non-linear model using the output gap as transition variable uncovers further interesting features of the GIPS fiscal policies. Fiscal policy is countercyclical in Greece and Spain, and acyclical for Portugal, but, for Ireland, it becomes procyclical during good times. With the exception of Spain, all GIPS correct budgetary disequilibria during bad times. Surprisingly, only Ireland and Spain correct budgetary disequilibria during good times, pointing to some degree of irresponsibility of Greece's and Portugal's fiscal policy authorities.

Column (iv) of Tables 2-5 reports the non-linear models (2)-(4) using CV_{t-1} as the transition variable and introducing a time-varying threshold. For Greece (see Table 2(iv)) and for Portugal (see Table 4(iv)), τ_1^{CV} is statistically insignificant and the estimated model is inferior to the corresponding model with the fixed threshold reported in column (ii). For Ireland (see Table 3(iv)), we find a statistically negative τ_1^{CV} . This model (which is superior to the model with the fixed threshold but inferior to the model in column (iii)) suggests that, during a financial crisis, corrective action is only taken for much larger deficits (this because the threshold that triggers action is relaxed). For Spain (see Table 5(iv)) we find a statistically positive τ_1^{CV} . This model (which is superior to the model with the fixed threshold but inferior to the model in column (iii)) suggests that, during a financial crisis, corrective action is only taken for much larger surpluses.

Amongst the three estimated models, the non-linear model in column (ii) delivers the lowest standard error and the highest adjusted R^2 for Greece and Portugal. For Ireland and Spain, the non-linear model in column (iii) provides the best fit.

3.2 Government expenditure models

Short-run models were also considered for government expenditure. Results, not reported for space considerations (but available on request) are summarized as follows.

For all GIPS, own lags have a significant and positive effect on current spending, pointing to self-perpetuating spending growth dynamics. Budgetary disequilibria are insignificant in explaining government spending dynamics, as evidenced by the statistical insignificance of CV_{t-1}, with the only exception of Spain, for which we report very weak evidence of a positive impact from CV_{t-1} (i.e. coefficient=0.10; t-ratio=1.72). Financial crises are also statistically insignificant, whilst output gap is insignificant for Greece and Portugal, but positive for Ireland and Spain. We fail to find evidence of non-linear effects in any of the countries. These results further corroborate our initial finding of weak exogeneity of government spending in the GIPS, as budgetary unbalances are corrected via the tax instrument, and government spending is mainly determined by its previous lags. The positive output gap effect for Ireland and Spain points nevertheless to a procyclical use of government spending.

4 Conclusions and Directions for further Research.

We have considered the fiscal policies of the GIPS, the four euro peripheral countries which are currently being scrutinized amid fears of insolvency and are receiving (or feared to need) financial aid from the EU/IMF. We confirm, through a formal IBC sustainability testing, the weak position of these countries. We fail to find evidence of IBC sustainability both in the "strong" and "weak" form for all the GIPS. This result holds not only when looking at the most recent years, but for most of the sample considered. Further, we show that most of the adjustments have taken place using the tax instrument alone, with government spending weakly exogenous for all GIPS (except Spain).

Allowing for non-linear corrections, we provide evidence of threshold behavior for all GIPS fiscal authorities. In line with the theoretical predictions, budgetary unbalances are not corrected, unless they become "too large". Further, our endogenously estimated thresholds for correction show that Greece and Portugal implicitly target a rather high level of deficit-to-GDP ratio, starting to correct well above the 3% ratio implied by the EGSP. Incidences of financial crises has the effect of increasing such threshold, further delaying corrective action: i.e. during financial crises only "very large" deficits are corrected. In principle, threshold adjustment of the fiscal variables would point to non-explosive debt dynamics and therefore to a sustainable path of fiscal policy. On the other hand, markets find it difficult to consider the sustainability of such policies as credible; indeed, in the presence of larger debt/deficits, financial markets require higher interest premia on government bonds, rendering more problematic the servicing of the existing stock of debt. Further, in these circumstances, credit rating agencies proceed by downgrading the rating of the sovereign debt, making more problematic the marketing of the debt. Indeed, the past three years have witnessed both high interest rate premia and successive sovereign downgrades.

Looking at the effects of the economic cycle, we note that Greece's and Portugal's fiscal policy authorities do not correct budgetary unbalances during good times, pointing to some degree of fiscal irresponsibility, as corrections during bad times will become more costly. At the same time, we document a procyclical government spending in Ireland and Spain. Announced policies in the GIPS to reduce government spending seem, consequently, to be a step in the right direction, provided they do not endanger future growth prospects.

A full assessment of fiscal policies during good and bad times, in conjunction with the effects of political cycles might prove an interesting extension. Further, the possibility of a "twin deficit" and its consequences for the GIPS fiscal policies has not been explicitly considered in our analysis; this important issue will be explored in future research.

References

[1] Afonso A. (2005), Fiscal Sustainability: the Unpleasant European Case. *FinanzArchiv*, 61 (1), 19-44.

[2] Afonso A., Rault C. (2008), What Do We Really Know about Fiscal Sustainability in the EU? A Panel Data Diagnostic. *Review of World Economics,* 145 (4), 731-755.

[3] Alesina, A. and S. Ardagna (1998), Fiscal adjustments. Why they can be expansionary. *Economic Policy*, 13, 489-545.

[4] Alesina, Alberto & Drazen, Allan, 1991. "Why Are Stabilizations Delayed?," *American Economic Review*, American Economic Association,

vol. 81(5), 1170-88.

[5] Arghyrou, M. and K. Luintel (2007). Government Solvency: Revisiting some EMU countries, *Journal of Macroeconomics*, 29, 387-410.

[6] Bajo-Rubio O., Diaz-Roldan C. and Esteve V. (2006), "Is the Budget Deficit Sustainable when Fiscal Policy is Nonlinear? The Case of Spain", *Journal of Macroeconomics* 28, pp.596-608.

[7] Bertola, G. and A. Drazen (1993), Trigger points and budget cuts: explaining the effects of fiscal austerity. *American Economic Review*, 83, 11-26.

[8] Bohn H. (1998), "The Behavior of U.S. Public Debt and Deficits", *Quarterly Journal of Economics,* 113, pp.949-963.

[9] Dahl, C.M. and G. González-Rivera (2003). Testing for neglected nonlinearity in regression models based on the theory of random fields. *Journal of Econometrics*, 114, 141-164.

[10] van Dijk, D., T. Teräsvirta, and P.H. Franses (2002). Smooth Transition Autoregressive Models – a Survey of Recent Developments, *Econometric Reviews*, 21, 1-47.

[11] Doornik, J.A. and D.F. Hendry (2000). *Modelling Dynamic Systems Using PcGive*, Volume II, Timberlake Consultants Ltd.

[12] Enders W. (2010), *Applied Econometrics of Time Series,* Wiley,

[13] Granger, C.W.J., and T. Teräsvirta (1993). *Modelling Nonlinear Economic Relationships*. Oxford: Oxford University Press.

[14] Greiner A, Köller U. and Semmler W. (2007), "Debt sustainability in the European Monetary Union: Theory and empirical evidence for se-

lected countries," *Oxford Economic Papers*, Oxford University Press, vol. 59(2), 194-218.

[15] Hakkio, G. and M. Rush (1991), Is the budget deficit too large?. *Economic Inquiry*, 29, 429-445.

[16] Hamilton, J.D. (2001). A parametric approach to flexible nonlinear inference. *Econometrica*, 69, 537-573.

[17] Hamilton, J. and M. Flavin (1986), On the limitations of government borrowing: a framework for empirical testing. *American Economic Review*, 76, 808-819.

[18] Hodrick, R. J. and E. C. Prescott (1997). "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, 29, 1–16.

[19] Johansen, S. (1988), Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control,* 12, 231-254.

[20] Kapetanios G., Shin Y. and Snell A. (2003), "Testing for Cointegration in Nonlinear STAR Error Correction Models", *Journal of Econometrics*, 112(2), 359-379.

[21] MacKinnon, J.G., A.A. Haug and L. Michelis (1999), Numerical distribution functions of likelihood ratio tests for cointegration. *Journal of Applied Econometrics*, 14, 563-577.

[22] Polito V. and M. Wickens (2007), Measuring the Fiscal Stance. *Discussion Paper in Economics*, University of York, 14/2007.

[23] Quintos, C.E. (1995), Sustainability of the deficit process with structural shifts. *Journal of Business & Economic Statistics*, 13, 409-417.

[24] Ravn, Morten O. and Harald Uhlig (2002). "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, 84, 371-375.

[25] Reinhart C.M. and K.S. Rogoff (2009). This Time is Different. Eight Centuries of Financial Folly. Princeton University Press, Princeton and Oxford.

[26] Santos Bravo A.B. and Silvestre A.L. (2002), "Intertemporal Sustainability of Fiscal Policy: Some tests for European Countries". *European Journal of Political Economy* 18, 517-528.

[27] Trehan, B. and C. Walsh (1991), Testing Intertemporal Budget Constraints: Theory and Applications to the U.S. Federal Budget and Current Account Deficits. *Journal of Money, Credit and Banking* 23, 210-223.

5 Data Appendix

The data used for this paper are taken from the annual macro-economic database of the European Commission's Directorate General for Economic and Financial Affairs (AMECO).

The statistical definitions of the series are:

TAX: Total revenue; general government; ESA 1995 (URTG).

URTG includes: sales of market output (ESA 95-code P.11) and output-for own final use (P.12) + Payments for other non-market output (P.131) + Other subsidies on production (D.39), receivable + Taxes on production and imports (D.2), receivable + Property income (D.4), receivable + Current taxes on income and wealth (D.5), receivable + Social contributions (D.61), receivable + Other current transfers (D.7), receivable + Capital transfers (D.9), receivable.

G:Total expenditure; general government; ESA 1995 (UUTG).

Total general government expenditure is the sum of: Intermediate consumption (P.2) + Gross capital formation (P.5) + Compensation of employees (D.1), payable + Other taxes on production (D.29), payable + Subsidies (D.3), payable + Property income (D.4), payable + Current taxes on income and wealth (D.5), payable + Social benefits other than social transfers in kind (D.62), payable + Social transfers in kind related to expenditure on products supplied to households via market producers (D.6311 + D.63121 + D.63131), payable + Other current transfers (D.7), payable + Adjustment for the change in the net equity of households on pension funds reserves (D.8)[10] + Capital transfers (D.9), payable + Acquisitions of non-produced non-financial assets (K.2)

GDP: gross domestic product at current market prices, reference level for excessive deficit procedure, ESA 1995 (UVGD).

[10] The adjustment for the change in net equity of households in pension funds reserves (D.8) represents the adjustment needed to make appear in the saving of households the change in the actuarial reserves on which households have a definite claim. Accordingly, it is part of the expenditure of the insurance enterprises sector and other sectors administering non-autonomous pension funds (see ESA 1995, paragraph 4.141 and 4.144).

Figure 1.*TAX/GDP* and *G/GDP*

Figure 2. Recursively estimated λ-max and λ-trace statistics divided by their critical values.

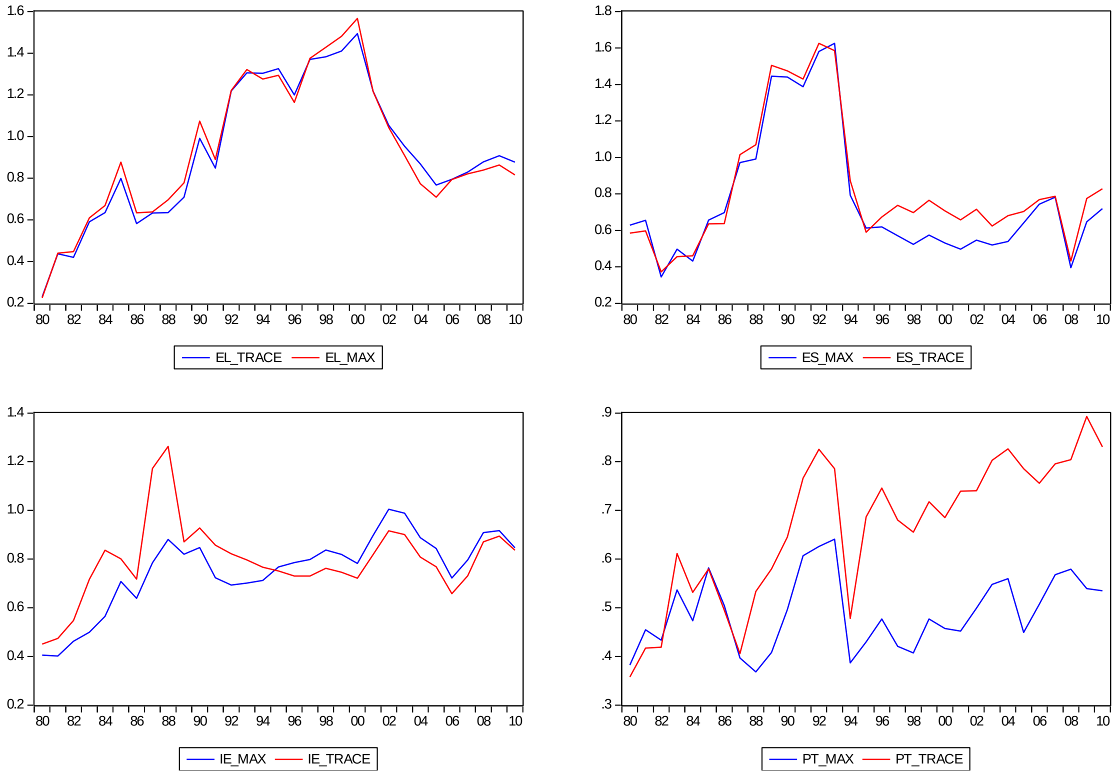

Figure 3. Recursively estimated cointegrating vector (plus/minus 2 s.e.)

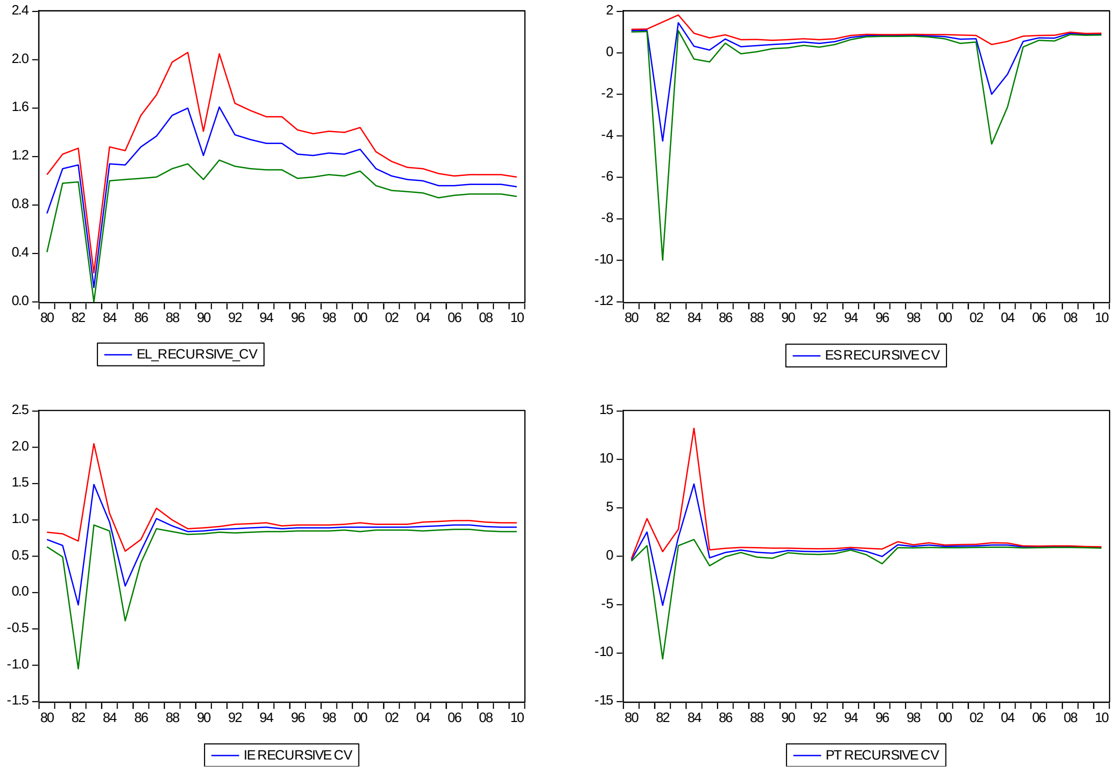

Figure 4. Recursively estimated adjustment coefficient for taxes (+/-2s.e.)

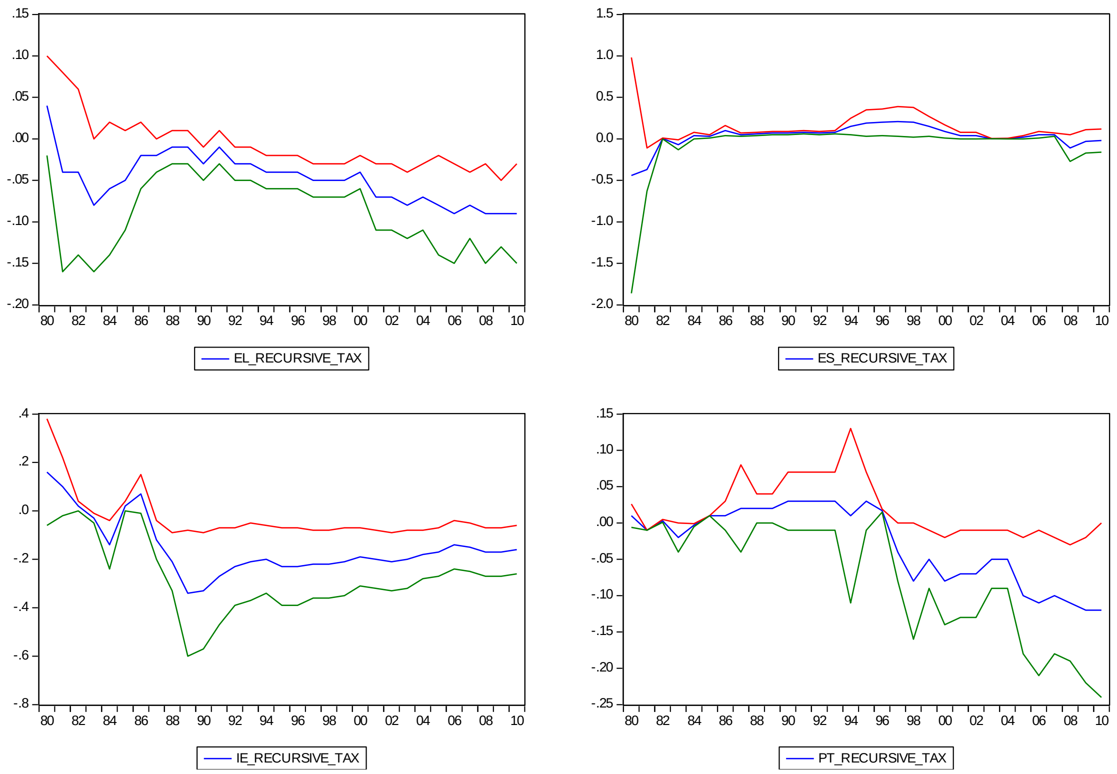

Figure 5. Recursively estimated adjustment coefficient for spending (plus/minus 2 s.e.).

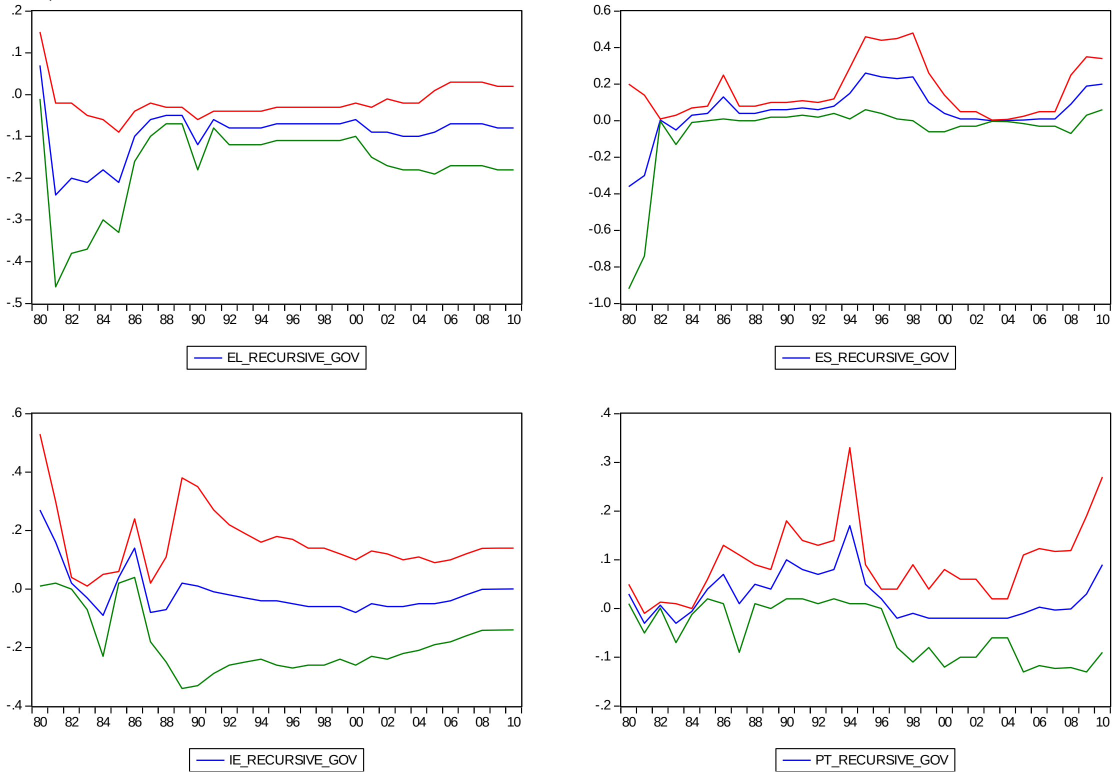

Figure 6: Financial crisis variable

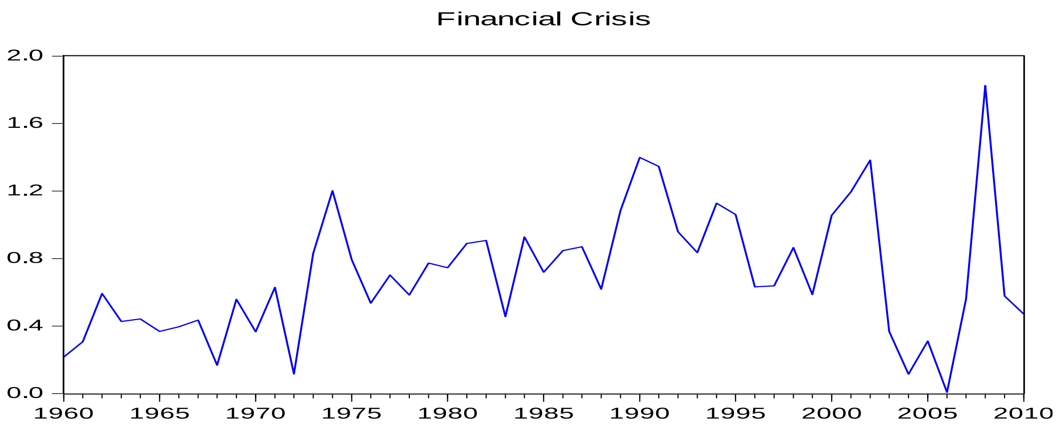

Table 1. Johansen Cointegration Test Results.

	CV	λ-max	λ-trace	αTAX	αGOV
GREECE	(1 –0.95)	9.16 (.11)	10.81 (.08)	-.09 (.03)	-.08 (.05)
IRELAND	(1 –0.90)	6.05 (.34)	7.00 (.32)	-.16 (.05)	.0002 (.007)
PORTUGAL	(1 –0.91)	9.48 (.10)	13.01 (.04)	-.12 (.06)	.09 (.09)
SPAIN	(1 –0.90)	7.98 (.17)	9.12 (.16)	-.002 (.007)	.21 (.07)

For each country we report the results of the Johansen cointegration: the estimated cointegrating vector (CV), the estimated λ-max and λ-trace statistics for the null hypothesis of zero cointegrating vectors (MacKinnon-Haug-Michelis, 1999 p-values are in parentheses), and the estimated adjustment coefficients of TAX and GOV (αTAX and αGOV respectively. Standard errors in parenthesis.

Table 2: **GREECE-**OLS estimates of alternative error correction models for $\Delta(TAX/GDP)$

	(i)	(ii)	(iii)	(iv)
	Linear model	Logistic model $s_{t-1} = CV_{t-1}$	Logistic model $s_{t-1} = gap_{t-1}$	Logistic model $s_{t-1} = CV_{t-1}$
Constant	0.202 (1.39)	0.193 (0.72)	0.190 (0.59)	0.309 (1.13)
$fincrisis_t$	0.050 (0.57)	-0.008 (-0.02)	-0.007 (-0.02)	-0.115 (-0.27)
CV_{t-1}	-0.078 (-2.37)			
gap_{t-1}	0.137 (3.24)			
		$\boldsymbol{CV_{t-1} <} \tau^{CV}$ **Regime**	$\boldsymbol{gap_{t-1} <} \tau^{gap}$ **Regime**	$\boldsymbol{CV_{t-1} <} \tau_t^{CV}$ **Regime**
CV_{t-1}		-0.110 (-2.72)	-0.130 (-2.30)	-0.128 (-2.78)
gap_{t-1}		0.060 (0.72)	0.151 (2.10)	0.010 (0.70)
		$\boldsymbol{CV_{t-1} >} \tau^{CV}$ **Regime**	$\boldsymbol{gap_{t-1} >} \tau^{gap}$ **Regime**	$\boldsymbol{CV_{t-1} >} \tau_t^{CV}$ **Regime**
CV_{t-1}		0.051 (0.58)	-0.040 (-0.83)	-0.030 (-0.62)
gap_{t-1}		0.167 (3.32)	0.179 (2.09)	0.161 (3.53)
τ^{CV}		-4.011 (-3.27)		
γ^{CV}		50.12 (-)*		48.34 (-)*
τ^{gap}			-0.101 (-0.97)	
γ^{gap}			20.31 (-)*	
τ_0^{CV}				-3.991 (-3.35)
τ_1^{CV}				-4.121 (-0.96)
Diagnostics				
Regression s.e.	0.85	0.82	0.87	0.83
$\overline{R}^2$	0.27	0.30	0.25	0.28
F*ar* (*p*-value)	0.87	0.88	0.83	0.87
F*arch* (*p*-value)	0.68	0.63	0.66	0.62
$\chi^2 nd$ (*p*-value)	0.95	0.84	0.90	0.83
λ-test (p-value)	0.01			
λ_A-test (p-value)	0.00			
g-test (p-value)	0.01			

Notes: t-ratios in parentheses. $\overline{R}^2$ is the adjusted coefficient of determination. *Imposed value. van Dijk et al. (2002) argue that the likelihood function is very insensitive to γ, suggesting that precise estimation of this parameter is unlikely. For this reason, we run a grid search in the range [0.1, 250] and fix the γ parameter to the one that delivers the best fit of the estimated models. F*ar* is the Lagrange Multiplier F-test for 2nd order serial correlation. F*arch* is the 1st order ARCH F-test. $\chi^2 nd$ is a Chi-square test for normality.

Table 3: **IRELAND-**OLS estimates of alternative error correction models for Δ(*TAX*/*GDP*)

	(i)	(ii)	(iii)	(iv)
	Linear model	Logistic model $s_{t-1} = CV_{t-1}$	Logistic model $s_{t-1} = gap_{t-1}$	Logistic model $s_{t-1} = CV_{t-1}$
Constant	0.739 (1.51)	0.617 (1.12)	1.643 (2.92)	0.723 (1.55)
fincrisis $_t$	-0.987 (-1.66)	-0.960 (-1.57)	-0.970 (-1.77)	-1.124 (-1.95)
CV_{t-1}	-0.134 (-2.77)			
gap_{t-1}	0.040 (1.02)			
λ-test (p-value)	0.01			
λ_A-test (p-value)	0.02			
g-test (p-value)	0.01			
		$CV_{t-1} < \tau^{CV}$ **Regime**	$gap_{t-1} < \tau^{gap}$ **Regime**	$CV_{t-1} < \tau_t^{CV}$ **Regime**
CV_{t-1}		-0.159 (-2.10)	-0.180 (-1.89)	-0.208 (-3.06)
gap_{t-1}		0.010 (0.14)	0.236 (2.94)	0.050 (0.59)
		$CV_{t-1} > \tau^{CV}$ **Regime**	$gap_{t-1} > \tau^{gap}$ **Regime**	$CV_{t-1} > \tau_t^{CV}$ **Regime**
CV_{t-1}		-0.099 (-0.81)	-0.114 (-2.23)	-0.040 (-0.59)
gap_{t-1}		0.032 (0.54)	-0.154 (-1.88)	0.023 (0.64)
τ^{CV}		-1.20 (-2.35)		
γ^{CV}		40.30 (-)*		35.64 (-)*
τ^{gap}			-0.010 (-0.78)	
γ^{gap}			25.43 (-)*	
τ_0^{CV}				-1.30 (-2.27)
τ_1^{CV}				-6.110 (-2.34)
Diagnostics				
Regression s.e.	1.49	1.46	1.35	1.43
$\overline{R}^2$	0.17	0.20	0.28	0.22
F*ar* (*p*-value)	0.11	0.18	0.40	0.19
F*arch* (*p*-value)	0.68	0.28	0.77	0.30
χ^2*nd* (*p*-value)	0.15	0.49	0.29	0.50
λ-test (p-value)	0.01			
λ_A-test (p-value)	0.02			
g-test (p-value)	0.01			

Notes: t-ratios in parentheses. $\overline{R}^2$ is the adjusted coefficient of determination. *Imposed value. van Dijk et al. (2002) argue that the likelihood function is very insensitive to γ, suggesting that precise estimation of this parameter is unlikely. For this reason, we run a grid search in the range [0.1, 250] and fix the γ parameter to the one that delivers the best fit of the estimated models. F*ar* is the Lagrange Multiplier F-test for 2nd order serial correlation. F*arch* is the 1st order ARCH F-test. χ^2*nd* is a Chi-square test for normality.

Table 4: **PORTUGAL-**OLS estimates of alternative error correction models for $\Delta(TAX/GDP)$

	(i)	(ii)	(iii)	(iv)
	Linear model	Logistic model $s_{t-1} = CV_{t-1}$	Logistic model $s_{t-1} = gap_{t-1}$	Logistic model $s_{t-1} = CV_{t-1}$
Constant	1.581 (2.85)	1.094 (1.46)	1.356 (2.17)	1.102 (140.)
$fincrisis_t$	-0.046 (-0.11)	-0.027 (-0.81)	-0.101 (-0.22)	-0.04 (-0.83)
CV_{t-1}	-0.194 (-2.99)			
gap_{t-1}	-0.002 (-0.57)			
$\Delta(G/GDP)_{t-1}$	-0.073 (-2.79)	-0.059 (-2.70)	-0.054 (-2.45)	-0.044 (-2.45)
		$CV_{t-1} < \tau^{CV}$ Regime	**$gap_{t-1} < \tau^{gap}$ Regime**	**$CV_{t-1} < \tau_t^{CV}$ Regime**
CV_{t-1}		-0.205 (-3.21)	-0.185 (-2.15)	-0.207 (-3.13)
gap_{t-1}		-0.010 (-0.34)	0.030 (0.17)	-0.005 (-0.24)
		$CV_{t-1} > \tau^{CV}$ Regime	**$gap_{t-1} > \tau^{gap}$ Regime**	**$CV_{t-1} > \tau_t^{CV}$ Regime**
CV_{t-1}		-0.101 (-1.57)	-0.111 (-1.37)	-0.090 (-1.35)
gap_{t-1}		-0.036 (-0.43)	0.046 (0.20)	-0.027 (-0.22)
τ^{CV}		-4.910 (-2.27)		
γ^{CV}		30.30 (-)*		25.37 (-)*
τ^{gap}			0.002 (0.77)	
γ^{gap}			5.33 (-)*	
τ_0^{CV}				-4.881 (-2.33)
τ_1^{CV}				-2.271 (0.28)
Diagnostics				
Regression s.e.	0.96	0.89	0.95	0.91
$\overline{R}^2$	0.29	0.35	0.30	0.33
F*ar* (*p*-value)	0.49	0.58	0.50	0.56
F*arch* (*p*-value)	0.02	0.12	0.10	0.11
$\chi^2 nd$ (*p*-value)	0.82	0.90	0.84	0.88
λ-test (p-value)	0.01			
λ_A-test (p-value)	0.00			
g-test (p-value)	0.01			

Notes: t-ratios in parentheses. $\overline{R}^2$ is the adjusted coefficient of determination. *Imposed value. van Dijk et al. (2002) argue that the likelihood function is very insensitive to γ, suggesting that precise estimation of this parameter is unlikely. For this reason, we run a grid search in the range [0.1, 250] and fix the γ parameter to the one that delivers the best fit of the estimated models. F*ar* is the Lagrange Multiplier F-test for 2nd order serial correlation. F*arch* is the 1st order ARCH F-test. $\chi^2 nd$ is a Chi-square test for normality.

Table 5: **SPAIN-**OLS estimates of alternative error correction models for Δ(*TAX*/*GDP*)

	(i)	(ii)	(iii)	(iv)
	Linear model	Logistic model $s_{t-1} = CV_{t-1}$	Logistic model $s_{t-1} = gap_{t-1}$	Logistic model $s_{t-1} = CV_{t-1}$
Constant	1.630 (4.47)	1.731 (4.67)	1.929 (5.80)	1.788 (5.01)
fincrisis $_t$	-1.420 (-3.35)	-1.486 (-4.04)	-1.670 (-4.52)	-1.493 (-4.10)
CV $_{t-1}$	-0.130 (-1.91)			
gap $_{t-1}$	0.050 (1.18)			
		***CV*$_{t-1}$ <** τ^{CV} **Regime**	***gap*$_{t-1}$ <** τ^{gap} **Regime**	***CV*$_{t-1}$ <** τ_t^{CV} **Regime**
CV $_{t-1}$		-0.018 (-0.11)	0.053 (0.06)	0.040 (0.38)
gap $_{t-1}$		0.151 (2.12)	0.110 (2.69)	0.120 (2.09)
		***CV*$_{t-1}$ >** τ^{CV} **Regime**	***gap*$_{t-1}$ >** τ^{gap} **Regime**	***CV*$_{t-1}$ >** τ_t^{CV} **Regime**
CV $_{t-1}$		-0.205 (-2.22)	-0.360 (-3.87)	-0.244 (-2.66)
gap $_{t-1}$		-0.050 (-0.96)	0.119 (2.70)	-0.060 (-0.94)
τ^{CV}		1.300 (2.58)		
γ^{CV}		20.42 (-)*		17.45 (-)*
τ^{gap}			-0.010 (-0.99)	
γ^{gap}			40.23 (-)*	
τ_0^{CV}				1.277 (2.65)
τ_1^{CV}				0.991 (2.12)
Diagnostics				
Regression s.e.	0.94	0.84	0.82	0.83
$\overline{R^2}$	0.24	0.38	0.40	0.39
F*ar* (*p*-value)	0.07	0.86	0.86	0.87
F*arch* (*p*-value)	0.91	0.92	0.92	0.93
χ^2*nd* (*p*-value)	0.05	0.74	0.60	0.61
λ-test (p-value)	0.01			
λ_A-test (p-value)	0.00			
g-test (p-value)	0.00			

Notes: t-ratios in parentheses. $\overline{R^2}$ is the adjusted coefficient of determination. *Imposed value. van Dijk et al. (2002) argue that the likelihood function is very insensitive to γ, suggesting that precise estimation of this parameter is unlikely. For this reason, we run a grid search in the range [0.1, 250] and fix the γ parameter to the one that delivers the best fit of the estimated models. F*ar* is the Lagrange Multiplier F-test for 2nd order serial correlation. F*arch* is the 1st order ARCH F-test. χ^2*nd* is a Chi-square test for normality.