An early-warning indicator for sovereign debt sustainability

Jasper Lukkezen*[1,2] and Hugo Rojas-Romagosa[1]

[1] *CPB Netherlands Bureau for Economic Policy Analysis*
[2] *Utrecht University*

7 August 2013

Abstract

We propose an early-warning indicator to assess government debt sustainability. This indicator combines the effect of economic uncertainty –represented by stochastic simulations of interest and growth rates– with the expected fiscal response that provides information on the long-term country specific attitude towards fiscal sustainability. We apply our framework on post-war data for nine OECD countries and find that our indicator –the width of the band that contains 95% of the simulated debt paths in 10 years– distinguishes countries that have sustainability concerns: Italy, Spain, Portugal and Iceland, from those that do not: US, United Kingdom, Netherlands, Belgium and Germany. Moreover, our indicator performs well as an early-warning instrument: its 2007 value predicts quite well the market assessment of fiscal sustainability in the ensuing sovereign debt crisis.

Keywords: public debt, fiscal policy, debt sustainability, stochastic simulations
JEL Classification: H6, H3, E6

1 Introduction

Whether government debt –and fiscal policy in general– is sustainable in the medium and long-term has been one of the main topics of debate in the current Euro crisis. An assessment of debt sustainability is a key input in decisions concerning the speed of fiscal consolidation, the need for reform and the determination of risk premia on

*Corresponding author. Email: Lukkezen@cpb.nl; Telephone: +31 703383408; Mail: Centraal Planbureau, Postbus 80510, 2508 GM Den Haag, The Netherlands.
The authors would like to thank Nico van Leeuwen for excellent research assistance; Oscar Bajo Rubio, Henning Bohn, Frits Bos, Carlos Marinheiro and Jan Luijten van der Zanden for making their data available; and Leon Bettendorf, Adam Elbourne, Casper van Ewijk, Clemens Kool, Catherine Mathieu, Ruud Okker, Bert Smid, Paul Veenendaal and participants at CPB, UU seminars and the Euroframe 2012 conference, for providing suggestions for improvement. All errors are our own.

government debt. Furthermore, as the crisis has made clear, fiscal surveillance is a key concern within a monetary union as unsustainable public finance may cause significant cross-border spillovers.[1] There is considerable debate, however, on how to measure debt sustainability. The original sustainability norms envisaged at the creation of the European Monetary Union (EMU) were to follow the Maastricht Treaty criteria: ceilings of 3% and 60% on government deficits and debt-to-GDP ratios, respectively. However, these criteria have proven to be inadequate as several countries violated these criteria without consequences, while others that met them have been nonetheless hit by the crisis. In particular, Spain had debt-to-GDP ratios and budget deficits well below these Maastricht limits, but has still suffered sovereign debt problems.

The objective of this paper is to find more informative economic indicators that provide guidance on medium- and long-term fiscal sustainability. We develop a dynamic framework for the assessment of stability of public finances and focus on the question whether governments can be expected to be "in control" of public finances. Our approach entails a simple and practical stochastic simulation which takes into account the response of fiscal policy to the state of public finances.

Our methodology has two steps. First, we follow Bohn (1998, 2008) and estimate a fiscal reaction function (FRF), which provides information on the long-term country-specific behavior of that country's government and its attitudes towards fiscal sustainability. A positive and significant FRF coefficient denotes a country that has been committed to reduce or maintain steady debt-to-GDP ratios conditional on short-term economic fluctuations and temporary government expenditures.[2] In the second step, the estimated FRF is combined with the stochastic debt simulation method proposed by Budina and van Wijnbergen (2008), which uses historic volatility of interest and growth rates to generate a distribution of future expected debt levels. We then simulate this model ten thousand times to generate a distribution of debt paths, which can be used to analyse the effect of fiscal responses and interest and growth rate volatility on debt-to-GDP ratios.[3]

The main contribution of this paper is that using this framework we derive a comprehensive early-warning indicator for the assessment of sustainability of government debt. Our indicator provides insight into the question whether a country is "in control" of its public finances. In practical terms, debt can be regarded as sustainable if it does not lead to ever diverging debt ratios in the long run –i.e. if the

[1] For an overview of direct spillovers see Lejour *et al.* (2011), for spillovers via contagion see Arezki *et al.* (2011) and for spillovers via monetary policy see Beetsma and Giuliodori (2010) and references therein.

[2] In particular, a positive and significant FRF coefficient can be interpreted as a government that engages in fiscal austerity to reduce debt levels even when markets are not specifically concerned about those debt levels, nor is there international pressure (e.g. EU institutions) to reduce them. A reason for this might be that in advanced economies fiscally responsible politicians at the national level have larger re-election probabilities (Brender and Drazen, 2005, 2008).

[3] Note that a similar approach has been followed by Medeiros (2012) and Berti (2013). These authors, however, rely on shorter quarterly time series –instead of long-term annual time series– and they do not present sustainability indicators.

simulated debt-to-GDP distribution is properly defined and bounded (Hall, 2013). In particular, our sustainability indicator is defined as the bandwidth that captures 95% of the simulated debt paths 10-years in the future.[4] If this band is broad, default fears cannot be dismissed as irrational, since after all, the debt level could rise significantly. A stronger FRF narrows this band and leads to lower indicator values, reflecting the confidence by market participants that the government will take the actions necessary to restore financial stability. Conversely, more volatile interest and growth rates widen this band as they lead to more uncertainty.

Given the long-term character of a fiscal response, we collected annual historic data on GDP and government finances for nine OECD countries spanning over a century: United States, United Kingdom, Netherlands, Belgium, Germany, Italy, Spain, Portugal and Iceland.[5] In this paper we focus on the post Second World War period, using the pre-war data as a robustness check. We find that until the 1980s, public debt was reduced by real growth and relatively low -and at times negative- real interest rates. In practical terms, this means that in this period it was not necessary to implement fiscal austerity plans to substantially reduce public debt. With financial liberalization from the 1980s onward, however, real interest rates have risen substantially, which increased the importance of fiscal policy for debt sustainability. We find that for the United States, the United Kingdom, the Netherlands, Belgium and Germany the fiscal response to increases in the debt-to-GDP ratio has been robust and positive for the whole sample as well as the post-war period. On the other hand, Spain, Portugal and Iceland have non-significant fiscal responses in the post-war period, which creates doubts about their capacity to reduce debt by fiscal austerity. Italy has a positive and significant fiscal response coefficient, but its current debt level is so large that its debt sustainability is susceptible to even small fluctuations in interest and growth rates. For instance, the simulated future debt paths for Italy, Spain, Portugal and Iceland show that their larger interest and growth rate variance requires a relatively large fiscal response to prevent debt levels from becoming unsustainable.

Finally, we show that our sustainability indicator performs well as an early-warning indicator. When we use only data until 2007 (i.e. prior to the fiscal problems in the EU), we find that our indicator is highly correlated with sovereign risk premia between 2009 and 2012. Our indicator thus clearly identifies those countries that

[4]Note that we explicitly avoid using any upper limit or critical debt level as part of our sustainability indicator. Several papers find that debt above a certain level has negative consequences for economic growth (Reinhart and Rogoff, 2010; Cecchetti *et al.*, 2011; Checherita-Westphal and Rother, 2012; Égert, 2012; Baum *et al.*, 2013). However, the causality between debt and growth is difficult to establish and critical debt levels are generally country-specific (e.g. Japan has debt levels way above any critical level mentioned in the literature, while other countries had debt crisis well below these critical levels), which makes the cross-country results from these studies not informative in an indicator when applied to different countries.

[5]We do not present our panel analysis results. First, we find them non-informative due to large cross-country heterogeneity in fiscal policy, business cycles and temporary expenditure spells. Second, the results yield a country-average assessment of debt sustainability, while debt sustainability concerns are mainly country-specific.

were later hit by the debt crisis: Portugal, Iceland, Italy and Spain. Moreover, it outperforms market based indicators prior to the crisis such as CDS rates in 2007.

An assessment of debt sustainability using this approach can complement existing indicators. Currently, the EU monitoring framework is based on the evolution of momentary indicators in the short term: public debt levels, deficit levels and EMU balances (European Commission, 2012b), and a sustainability analysis taking the effects of an ageing population on public finances into account in the long term (European Commission, 2012a). In this monitoring framework, fiscal policy is essentially exogenous and economic uncertainty is not explicitly taken into account. Our approach explicitly models the effect of economic uncertainty on medium / long term debt sustainability, but it is important to note that it is not suitable for short-term analysis. In particular, our analysis is based on ex-post data that already accounts for any endogenous behavior between fiscal policy, financial markets and the real economy –but these mechanisms are still at play in any short-term debt sustainability assessment and thus, they are beyond the scope of our analysis. Finally, to estimate the expected response of fiscal policy we require long time series on public finances, which may not be readily available for all countries.

The paper is organized as follows. Section 2 presents the theoretical background on debt sustainability. Section 3 describes the data. In Section 4 we elaborate on our empirical strategy and present country-specific econometric results. Section 5 describes our stochastic analysis and Section 6 explains how we construct our debt sustainability indicator and how to apply it as an early-warning indicator. Section 7 summarizes our main results.

2 Fiscal reaction functions and debt sustainability

We follow the methodological approach developed by Bohn (1998, 2008) to analyse debt sustainability using historical information. In essence, Bohn's approach is to equate fiscal sustainability with the stationarity of the debt-to-GDP time series –i.e. when the debt-to-GDP time series is stationary over time without a trend, the debt-to-GDP ratio is sustainable. Most of the literature uses unit root or cointegration tests often to test whether the intertemporal budget constraint (IBC) holds.[6] This is not a sufficient condition for debt stationarity though, as it is possible to satisfy the IBC while simultaneously having a mildly explosive path of the debt-to-GDP ratio (Bohn, 2007). Furthermore, these approaches have low power in distinguishing unit root from near unit root alternatives. Hall (2013) shows that if the distribution of the simulated debt paths is properly defined (i.e. bounded) in the long run, the debt-to-GDP ratio is stationary and follows a near unit root instead of a unit root process. Finally, Bohn's approach is preferable as it distinguishes the channels through which governments can achieve sustainable debt paths.

The main object of interest in Bohn's methodology is the debt-to-GDP ratio. To analyse its behaviour he uses two equations, the accounting equation for government

[6]See Afonso (2005) for a survey of these type of studies.

debt and a behavioural equation for the government's primary surplus, both with an error-correction type specification for the primary surplus-to-GDP ratio and the debt-to-GDP ratio. The starting point of his analysis is the accounting equation:

$$D_{t+1} = (1 + r_t)(D_t - S_t),$$

which says that government debt D at the beginning of period $t+1$ equals debt at the beginning of period t minus the primary surplus S (i.e. the budget surplus excluding interest payments) over period t times the gross interest factor $(1 + r_t)$.[7] Following the standard approach in the literature Bohn uses the debt-to-GDP ratio as the main analytical variable. Therefore, we re-write the budget accounting equation in GDP-ratio form:

$$\frac{D_{t+1}}{Y_{t+1}} = (1 + r_t)\frac{Y_t}{Y_{t+1}}\left(\frac{D_t}{Y_t} - \frac{S_t}{Y_t}\right),$$

which can be simplified to:

$$d_{t+1} = \frac{1 + r_t}{1 + y_t}(d_t - s_t), \tag{1}$$

where r is the real interest rate and y is the real GDP growth rate.[8] Note that for the derivation of equation (1), we can also use nominal interest and GDP growth rates since expected inflation cancels out in the difference.

The second part of Bohn's methodology consists in estimating the fiscal reaction function (FRF), which indicates whether the government increases its primary surplus as a reaction to changes in the debt-to-GDP ratio. We estimate the error-correction behavioural specification $(d - s)$ with the following regression:

$$s_t = \alpha + \rho d_t + \beta \mathbf{Z}_t + \varepsilon_t, \tag{2}$$

where ρ is the fiscal reaction parameter, $\mathbf{Z}$ is a set of other primary surplus determinants and ε_t is an error term. This specification tells us how governments react to debt accumulation given a structure of shocks occurring in the background.

The use of $\mathbf{Z}$ is crucial to account for shocks and it consists of two variables: GVAR is a measure of temporary government spending (e.g. military expenditure during war periods) and YVAR measures cyclical fluctuations in output (e.g. business cycles). The presence of these shocks makes it difficult to detect if d is stationary. Including these variables, hence, is crucial for the results in Bohn (1998).[9]

[7] Throughout this paper upper case variables denote nominal values and their lower case counterparts denote the nominal value as a share of GDP. Stock variables are defined at the beginning of the period, whereas flow variables are defined over the period. Interest rates refer to effective interest rates defined as the proportion of interest payments to the overall government debt level.

[8] Nominal interest rate is given by $R \approx (1+r)(1+\pi) - 1$, where π is inflation estimated from the GDP deflator. We approximate $r \approx R - \pi$. These interest rates incorporate a risk premium, but as we are using historical (or ex-post) data only, we do not need to explicitly calculate and factor out these risk premia.

[9] Bohn (1998) uses Barro (1979)'s classical tax-smoothing theory to underpin the use of these variables as temporary government expenses and the effects of business cycle slow-downs should be financed by a higher budgetary deficit.

Substituting equation (2) in (1) yields an expression for the evolution of the debt level:

$$d_{t+1} = \gamma_t(1-\rho)d_t - \gamma_t\left(\alpha + \beta \mathbf{Z}_t + \varepsilon_t\right), \tag{3}$$

where γ summarizes the relationship between interest rates, growth rates and inflation:

$$\gamma_t = \frac{1+r_t}{1+y_t} \approx 1 + r_t - y_t. \tag{4}$$

As $E(\mathbf{Z}_t) = 0$, debt sustainability becomes a function of γ and ρ. When we use average values for interest and growth rates ($\bar{r}$ and $\bar{y}$, respectively), we can summarize this information using the parameter δ, such that:

$$\delta = \bar{\gamma}(1-\rho) = (1+\bar{r}-\bar{y})(1-\rho) \tag{5}$$

We distinguish three cases:

- $\delta < 1$ implies stationary debt-to-GDP ratios.[10] [11]
- $\delta > 1$ but with $0 < \rho < \bar{r} - \bar{y}$ implies mildly explosive paths for debt-to-GDP ratios (but growing slowly enough to be consistent with IBC).[12]
- $\delta > 1$ with $\rho < 0$ and $\bar{r} - \bar{y} > 0$ characterizes exponentially growing debt. It also suggests that the government disregards debt when setting fiscal policy and thus, raises doubts about its creditworthiness.

Most of the literature concludes that if there is any corrective action ($\rho > 0$), debt –and fiscal policy in general– is sustainable as it satisfies the intertemporal budget constraint. This methodology is also referred to as examining fiscal reaction functions or model-based sustainability. However, we follow Bohn (1998, 2008) in placing the border between sustainable and unsustainable fiscal policy at the requirement to have a stationary process for debt-to-GDP ratio. Because if mildly explosive debt paths are allowed, the fiscal response required (ρ times d) to keep on this mildly explosive debt path is increasing in the debt level d and thus mildly explosive as well. Requiring $\delta < 1$ excludes the second case and is used by Ghosh *et al.* (2013) in their study on fiscal space.

[10] Bohn (1998) uses this condition to argue that the coefficient estimates of equation (2) are unbiased. He assumes that $\bar{\gamma}$ and $\alpha + \beta\mathbf{Z}_t + \varepsilon_t$ are stationary and states that if $(1-\rho)\bar{\gamma} < 1$ then d should be stationary. If d is stationary, the debt-to-GDP ratio follows a auto-regressive process with near unit root behaviour and OLS coefficient estimates are unbiased. If $(1-\rho)\bar{\gamma} > 1$, coefficients and specifically ρ may be biased towards zero, which makes debt look even more non-stationary.

[11] The deterministic steady state debt level can be obtained by writing equation (3) in first differences:

$$\Delta d_{t+1} = -\left(1-\delta\right)d_t - \bar{\gamma}\alpha,$$

where we use that, by construction, $E(\mathbf{Z_t}) = E(\varepsilon_t) = 0$. Then the deterministic steady state yields $\overline{d} = \left(1-\delta\right)^{-1}\bar{\gamma}\alpha$. For $\alpha > 0$ this debt level is negative (i.e. in the long run the government will have more assets than liabilities) and for $\alpha < 0$ this debt level is positive.

[12] See Bohn (2007) for a formal proof. At the boundary between the first and the second case, where $\rho = \frac{\bar{\gamma}-1}{\bar{\gamma}}$, it implies a difference-stationary debt that satisfies the unit root and cointegration conditions (this is the most studied scenario in the unit root literature).

Requiring a stricter condition on sustainability makes the features of the FRF test not less convenient. Mendoza and Ostry (2008) describe in detail the benefits and limitations of the FRF analysis. First, it does not require knowledge of the specific set of government policies on debt, taxes and expenditures. The FRF test determines whether the outcome of a given set of policies implicit in the past primary balance and debt data is in line with fiscal solvency, without knowing the specifics of those policies. Second, since asset pricing applies to all kinds of financial assets, the analysis does not require particular assumptions about debt management, or the composition of debt in terms of maturity or denomination structure. Third, it relies entirely on ex-post realizations of all our variables. This means that it already contains the outcomes of the endogenous process that interacts governmental policies, financial market assessments and the response of the real economy in the short-run. For example, it will contain the realization of any possible causal relation between public debt and growth. For our medium term analysis, this is property is very convenient, as our estimate of the fiscal response is one where the short-run dynamics of the government will be already included in the data. However, it should be highlighted that this also means that the FRF is not suitable for short-term assessments of debt sustainability. A drawback from the dependence of the FRF on historical data, is that it cannot reproduce events which have not been recorded in the past. Thus a FRF test provides limited guidance when severe and fully unanticipated shocks occur, such as an unexpected shut down in the credit markets or a surprisingly large negative shock to the primary balance following from a war or natural disaster.[13]

A final remark concerning the FRF analysis is that a time-invariant conditional response of the primary balance to the debt level (ρ) alone is a sufficient but not a necessary condition for debt sustainability. A non-linear and/or time varying response can also generate fiscal solvency as long as the response is strictly positive above a certain debt-to-GDP threshold ratio. This implies that countries without a positive ρ and with $\gamma > 0$ not necessarily have unsustainable government finances. Theoretically, they could have a response that kicks in at some higher, not yet reached, debt level or specific set of government policies that will likely improve primary surplus in the future. In practical terms, this refers to non-linear relationships in equation (2), for which we test in our empirical analysis.

What does this analysis tell us about the channels through which the debt level is controlled? From equation (5) we see that the evolution of the debt ratio is driven by three contributing channels:[14]

1. Fiscal reactions. These are captured by the estimated coefficient (ρ) of the FRF and provide information on the historical fiscal reaction of governments (i.e. changes in primary surpluses) to changes in the debt-to-GDP ratio. A

[13] Also, the required ρ must be larger in case of non-complete markets of state-contingent claims with precautionary saving behaviour to accommodate for the tighter debt limits imposed by these precautionary savings (Dehejia and Rowe, 1995).

[14] We depart slightly from Bohn's classification, who defined "growth dividend" as the difference between real interest rates on government debt and real GDP growth rates.

positive and significant FRF coefficient denotes a country that has been historically committed to reduce or maintain steady debt-to-GDP ratios conditional on short-term economic fluctuations and temporary government expenditures (e.g. military expenditures during wars).[15] The estimated FRF coefficient is a long-term country-specific institutional indicator that provides information on the fiscal behaviour of that country's government and its attitudes towards fiscal sustainability. A reason why these fiscal reactions are likely to be persistent is that in advanced economies fiscally responsible politicians at the national level have larger re-election probabilities (Brender and Drazen, 2005, 2008).

2. Real growth dividend. This term has a beneficial effect on the debt-to-GDP ratios when real GDP growth (y) is positive and sustained over time. Therefore, this term groups governmental policies –such as structural reforms– and external factors –such as foreign demand– on the real economy that have a medium- to long-term effect on real growth rates.

3. Real effective interest rates (r) on government debt. This is the difference between the nominal interest and the inflation rate. Thus, this category groups the monetary and financial policy instruments available to governments to reduce debt levels. It is also linked to the term "financial repression" coined by Reinhart and Sbrancia (2011), who define it as policies that depress real interest rates –while the extreme case of periods with negative real interest rates is defined as "liquidation years".

If economic growth y is larger than the effective interest rate r, then $\gamma < 1$ and the debt-to-GDP ratio decreases over time and a positive fiscal response is not needed to assure debt sustainability. This is also known as the "Aaron condition" (Aaron, 1966). If this condition is not met (e.g. $y < r$ and thus, $\gamma > 1$) then debt sustainability depends on the fiscal reaction function: primary surpluses should be sufficiently responsive to the debt-to-GDP ratio to arrive at a stationary debt-to-GDP level.

3 Data

We have chosen countries with historical time series of at least 100 years.[16] The countries in our sample are: the United States (USA), the United Kingdom (GBR), the Netherlands (NLD), Belgium (BEL), Germany (DEU), Italy (ITA), Spain (ESP), Portugal (PRT) and Iceland (ISL). Long-time series are necessary to have reliable estimates of the FRF that can span several business cycles and contain periods of increasing and decreasing debt levels in order to achieve reliable results.

[15]For instance, in terms of the recent Euro crisis, a positive FRF coefficient can be interpreted as a government that engages in fiscal austerity to reduce debt levels even when fiscal policy is not pro-cyclical, or markets are not specifically concerned about those debt levels, nor is there international pressure (e.g. EU institutions) to reduce them.

[16]With the exception of Germany, for which we only had data from 1970 onwards.

Our selection of the post-war sample is based on the fact that all countries were directly or indirectly affected by the Second World War and this sample provides more than 60 years of observations.[17] In this period, our estimations can be directly related to the current institutional settings in our nine OECD countries. Thus, using the post-war sample we can abstract from considering other institutional settings that may have been present if we used the full historical sample –that includes data as far back as 1691 for the United Kingdom. Therefore, we only use the full-sample results as a robustness check for the post-war sub-sample.

We use the following time series: nominal GDP, real GDP, GDP deflator, gross debt, primary surplus, interest payments on gross debt and government expenditures. YVAR is obtained from the real GDP series and GVAR from government expenditures as a percentage of GDP. The sources and the assumptions we made while preprocessing the data are fully described in the Appendix A.[18]

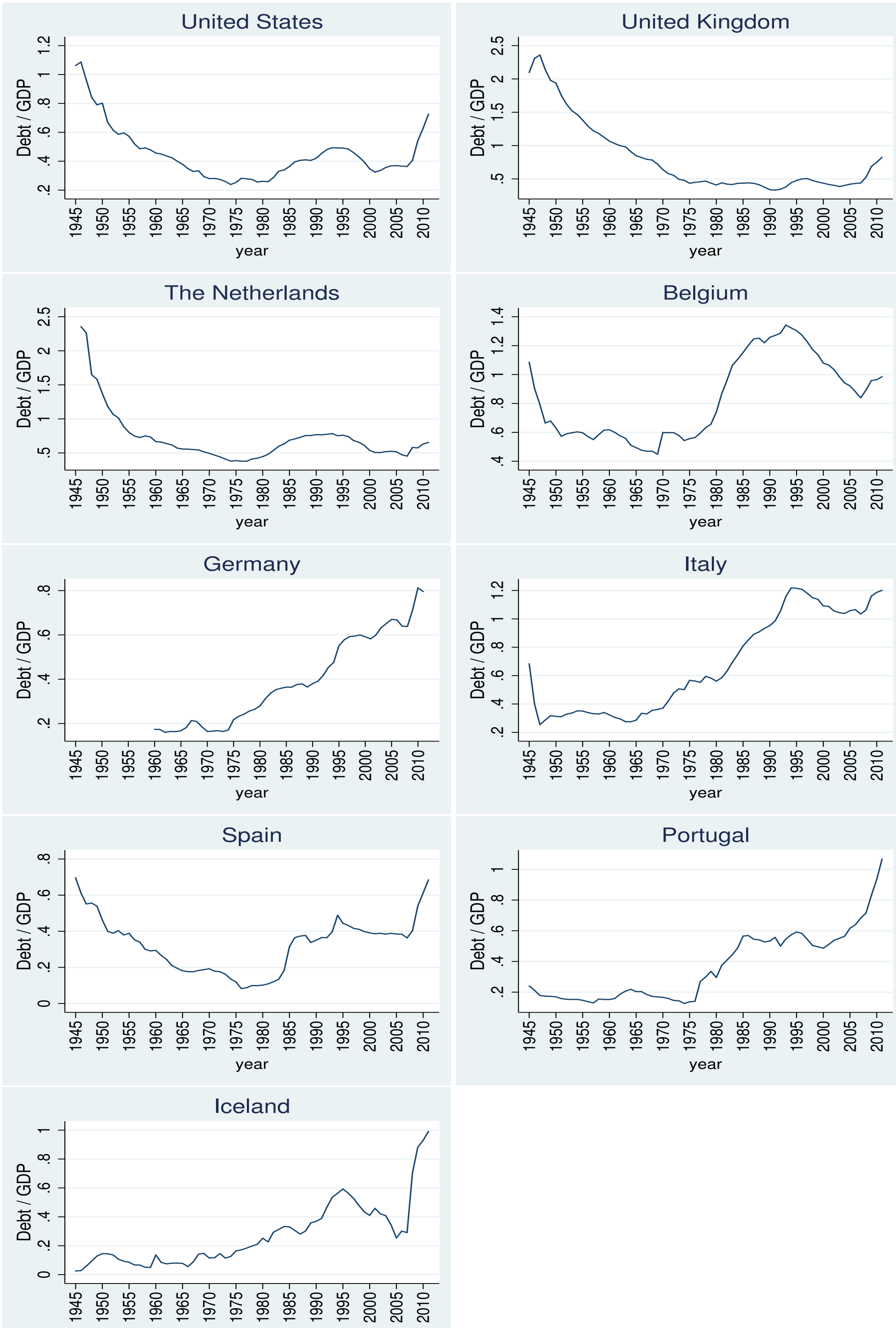
In Figure 1 we show the debt levels for all countries in the post-war sample. Four countries: the United States, the United Kingdom, the Netherlands, Belgium and Spain begin with high debt levels after the Second World War. These levels declined sharply afterwards, but have increased again in the later period -especially in the last decade. The other countries: Germany, Italy, Portugal and Iceland began the period with relatively low debt levels and have experienced steady debt increases.

In Figure 6 in the Appendix we include the real growth rates in the post-war period for all countries. We observe that real growth rates have been on steady decline for all countries in this period. Italy, Spain and Portugal, however, experienced a growth boom in the 1960s, while Iceland did so in the late 1970s. To identify periods with financial repression, in Figure 7 in the Appendix we plot nominal interest rates against inflation (estimated from the GDP deflator). Financial repression was present in all countries until at least the 1980s. This is consistent with the findings in Reinhart and Sbrancia (2011). Financial repression was largest for Iceland, Italy, Portugal and Spain.

[17] The exact initial year of the sample varies between countries because of particular data limitations.

[18] Table 3 presents a short summary of the available data. Furthermore, we performed unit root tests on all the time series for all countries used in the analysis and were able to reject the presence of a unit root for all variables, except the debt-to-GDP ratio series. However, the use of this variable in our regressions has already been justified in Section 2.

Figure 1: Debt-to-GDP ratios in the post-war period.



Finally, in Figures 8 to 12 in the Appendix we show the time series for the debt-to-GDP ratio, the primary surplus to GDP ratio, and the γ-parameter for all seven countries for the post-war sample. In addition, we show the military expenditure for the US and the UK.

4 Estimating fiscal reaction functions

Our empirical strategy is straightforward and consists of two main components. First, we estimate equation (2) for all countries to obtain the fiscal response parameter ρ. Second, we calculate the average values for the real interest rates and real growth rates and use these values to obtain γ from equation (4). With both sets of information we then estimate δ as defined in equation (5). Using this information we analyse if government finances have been sustainable and if it was due to prudent fiscal policy, financial repression or the growth dividend.

4.1 Linear regressions

We estimate all the regressions using both OLS and autocorrelation and heteroskedasticity consistent estimators. We estimate the post-war period and we only present the multivariate regressions including the indicator of fluctuations in income growth (YVAR) and fluctuations in government expenditures (GVAR) as control variables, given the importance of these two variables for the analysis (cf. Bohn, 1998).[19]

Analogous to Bohn (2008), we use an HP-filter ($\lambda = 100$) to extract the trend component of log real GDP and define YVAR as the gap between the actual value and this trend in percentage points of GDP.[20] We extract GVAR analogously to YVAR by using the cyclical component of government spending.[21] These linear multivariate regressions are the core of our empirical analysis, but we check the robustness of these results in Section 4.4.

4.2 Non-linear regressions

We also examine whether the response of the primary surplus to an increase in the debt-to-GDP ratio is non-linear. There are different ways, however, to interpret these non-linearities. On one hand, non-linearities may arise because –above a certain debt-to-GDP ratio– the incentives for policy makers to increase the primary surplus are missing, causing a debt overhang problem. This concept was introduced by Krugman (1988) and confirmed empirically by Callen *et al.* (2003) and Mendoza

[19]The univariate regressions are available upon request.

[20]A more structural way is to estimate potential GDP first and then define the difference between actual and potential GDP as the output gap. This is used by the OECD (2005). We do not apply this method due to data limitations and because the structure of the economy likely changed over time. The HP-filter generates an output gap comparable to those in OECD (2005), so the potential measurement error is small.

[21]Mendoza and Ostry (2008) also use this approach. Their results are robust to different specifications for extracting the cyclical component.

and Ostry (2008) for emerging market economies. On the other hand, non-linearities may arise because policy makers get increasingly nervous about the possibility of losing access to capital markets. In that case there are larger fiscal responses at higher debt levels. For instance, high debt levels can raise financing difficulties for the government. Bohn (1998, 2008) finds that for the United States the conditional response of primary surplus to debt is stronger when the debt-to-GDP ratio is high by historical standards. In their recent study on fiscal space Ghosh *et al.* (2013) combine both views.

We test for non-linearities using two approaches. First, we add the quadratic terms: d_t^2 and $(d_t - \overline{d})^2$ as explanatory variables in equation (2), where $\overline{d}$ is the mean value of d. Second, we examine if the fiscal response is different above a certain level of debt-to-GDP. To test this we create three dummy variables: $dmax$, $dmax_{40}$ and $dmax_{60}$, where $dmax = 1$ if the debt-to-GDP ratio is above the historical debt average, and otherwise $dmax = 0$. Accordingly, $dmax_{40}$ and $dmax_{60}$ are equal to one if d is above 40% and 60%, respectively. We add each additional variable separately in equation (2), but only present the results for d_t^2 and $dmax_{60}$.

4.3 Econometric results

The country specific results are summarized in Table 1. The first part of this table presents a summary of the multivariate regression results –while Tables 4 to 12 in the Appendix show the full econometric results. For the United States (USA), the United Kingdom (GBR), the Netherlands (NLD), Belgium (BEL), Germany (DEU) and Italy (ITA) the fiscal reaction coefficient (ρ) is positive and significant for the post-war sample.[22] This shows that these governments have significant and strong fiscal responses to increases in the debt-to-GDP ratio, once we control for the business cycle (YVAR) and temporary expenditures (GVAR). On the other hand, Spain (ESP), Portugal (PRT) and Iceland (ISL) do not have significant fiscal reaction coefficients, and it is even negative for the last two countries.

Furthermore, when using non-linear specifications we find that for the United States, the Netherlands, Belgium and Italy the response to square debt is positive and significant. This means that these governments have a stronger fiscal response when debt-to-GDP ratios are above their historical average. Contrastingly, the non-linear coefficients are negative and significant for Germany, Spain, Portugal and Iceland, which creates additional concerns about the debt sustainability of these countries.

In the lower part of Table 1 we summarize the information on real interest and growth rates by presenting average values for three different time periods. We also

[22] The policy response of the United States decreases significantly when 2010 and 2011 are included in our estimations. This may imply a structural break in the fiscal stance towards debt or it may indicate that the US government has a delayed response to debt increase which has yet to materialize. However, we present only the 1948-2009 period for the United States (see columns 1 to 5 in Table 4 in the Appendix, which provides a FRF consistent with the historical time series from 1792-2011 (column 6). Since the ρ coefficient is a long-term institutional indicator we prefer long-term fiscal responses to short-term variations at the end of the sample.

Table 1: Summary of estimated coefficients, average values and the δ-parameter, for all countries in the post-war and sub-samples.

	USA 1948-2009	GBR 1946-2011	NLD 1948-2011	BEL 1955-2011	DEU 1970-2011	ITA 1948-2011	ESP 1946-2011	PRT 1945-2011	ISL 1946-2011
Estimated coefficients from linear regressions:									
ρ	0.078 ***	0.045 ***	0.078 ***	0.038 ***	0.026 *	0.073 ***	0.005	-0.005	-0.020
yvar	0.284 **	0.410	0.451 ***	0.612 ***	0.227 **	0.360	0.298 **	0.053	0.203 **
gvar	-0.083	-0.143	0.046	-0.101 ***	-0.260 ***	-0.039	-0.039	-0.048	-0.093 ***
Estimated coefficients from non-linear regressions:									
d^2	0.238 *	0.021	0.158 ***	0.153 *	-0.196 ***	0.370 ***	-0.820 ***	-0.106	-0.114 **
$dmax_{60}$	0.044 *	-0.037	-0.019	-0.038	-0.022 **	0.125 ***	-0.120 ***	-0.050 ***	-0.050 ***
Average values, full post-war sample (varies by country):									
$\bar{r}$	0.022	0.015	0.022	0.042	0.036	-0.004	0.001	-0.036	-0.054
$\bar{y}$	0.032	0.022	0.035	0.028	0.021	0.036	0.042	0.037	0.053
$\bar{\gamma}$	0.991	0.993	0.988	1.014	1.014	0.961	0.960	0.929	0.898
δ	0.913	0.948	0.911	0.975	0.988	0.891	0.960	0.929	0.898
Average values, post-war until 1986:									
$\bar{r}$	0.013	0.003	0.008	0.043	0.025	-0.026	-0.044	-0.056	-0.093
$\bar{y}$	0.036	0.023	0.042	0.034	0.025	0.051	0.049	0.045	0.069
$\bar{\gamma}$	0.978	0.980	0.968	1.009	1.000	0.927	0.912	0.903	0.848
Average values, 1987-2011:									
$\bar{r}$	0.037	0.035	0.044	0.042	0.042	0.033	0.073	-0.001	0.010
$\bar{y}$	0.025	0.021	0.024	0.021	0.019	0.013	0.031	0.025	0.028
$\bar{\gamma}$	1.011	1.013	1.020	1.021	1.024	1.020	1.041	0.974	0.983

Notes: The full regression results are shown in Tables 4 to 12. Significance levels: *** p<0.01, ** p<0.05, * p<0.1 (computed using heteroskedasticity- and autocorrelation-consistent standard errors with Newey-West lag window of size 1). Dependent variable is the primary surplus to GDP ratio. Explanatory variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend for all countries but the USA and GBR (for which "gvar" equals military expenditure). Both trends are extracted using an HP-filter (lambda=100). d^2 is the square of the debt-to-GDP ratio (d) and $dmax_{60}$ is a dummy variable equal to one if d is above 60%. Average values for the real interest ($\bar{r}$) and real growth rates ($\bar{y}$) come from Figure 7.

provide the information on our summary parameters: γ (i.e. the Aaron condition) and δ (which combines the Aaron condition with fiscal reaction coefficient ρ). The Aaron condition ($\gamma < 1$) can be satisfied with financial repression, the growth dividend, or a combination of both. Recall that government finances are deemed sustainable if $\delta = \gamma(1 - \rho) < 1$.

We observe that the Aaron condition ($\gamma < 1$) is satisfied –on average– for all countries in the post-war period, except for Belgium and Germany. However, since Belgium and Germany have a positive and significant fiscal reaction coefficient (ρ), their δ parameter is below one. This applies to all other countries on average in the post-war period. However, there is a structural change for most countries at the end of the 1980s. We chose the year 1987 as the break point of the series for illustrative purposes, these results do not depend on the exact year that is chosen. For Italy and Spain we find financial repression (i.e. negative real interest rates) until the late 1980s. After 1987 interest rates have large average values well above their average growth rates and thus, both countries do not satisfy the Aaron condition any longer. This combination of higher interest rates and lower growth after 1987 also applies to the United States, the United Kingdom, the Netherlands, Belgium and Germany. On the other hand, Portugal and Iceland still had relatively low real interest rates well into the 2000s. This means that for both countries debt in the post-war period was made sustainable as a result of financial repression (for part or the whole period), and thus, strong fiscal responses were not required.

The post-war results do not ensure that debt is or will be sustainable in the near future for these countries. In particular, since financial repression ended, the importance of fiscal responses has greatly increased. Countries that lack a significant fiscal response may then have difficulties to maintain debt-to-GDP ratios at sustainable levels. Moreover, the absence of a linear fiscal response, in conjunction with a negative non-linear response for Germany, Spain, Portugal and Iceland rises the concern that debt may not be sustainable in these countries. Our stochastic analysis in Section 6 will use this information combined with a VAR analysis to provide our early-warning sustainability indicator.

4.4 Sensitivity analysis

We tested the robustness of our econometric results and present the test results per country in Tables 4 to 12 in the Appendix.

First, we included the real interest rate as an explanatory variable in equation (2). The intuition is that the primary surplus can also react to changes in real interest rates. For instance, when financial repression is strong enough to generate negative real interest rates the government faces less pressure to reduce the debt with fiscal responses. Another link is that high real interest rates can also force the government to apply fiscal austerity, even when the debt-to-GDP level is not that high. For the United States, Bohn (1998) found that interest rates are not a significant control variable. We find similar results for all countries but Germany, where the inclusion of the real interest rates yields a non-significant ρ parameter.

Second, we included inflation as an additional variable but it was generally not significant, and when it was significant it did not change the main results. Third, we employed alternative definitions of GVAR for some countries. For the United States and the United Kingdom –where military spending has been historically a big driver of temporary government spending– we define GVAR as the military spending-to-GDP ratio.[23] For the Netherlands we used gas revenue as a proxy for GVAR following Wierts and Schotten (2008).

[23] This is comparable to Bohn (2008) who defines GVAR as the gap between a permanent component of military outlays to GDP from an estimated AR(2) process and the actual values. Our

Finally, we assessed the stability of the fiscal response over time. The most obvious choice is to include the full historical sample available and not only the post-war period. These results are presented in the last column of the country-specific results in the Appendix B –except for Germany for which we do not have data before 1970. It is remarkable that for the United States and the United Kingdom, the full-sample historical fiscal reaction coefficient is significant and very close in value to the post-war estimation. For the Netherlands and Belgium the full-sample coefficient is significant but has a lower value. While for Italy and Spain the significance of the coefficients for both samples is reversed. In Italy ρ is significant in the post-war sample, but it is not significant in the full-sample, while in Spain only the historical ρ coefficient is significant. Portugal and Iceland have non-significant coefficients for both samples. We also used Bai and Perron (1998, 2003) tests for endogenous structural breaks after the Second World War. The resulting sub-samples, however, are usually too short to estimate a stable fiscal reaction function that can assess the long-run institutional stance towards fiscal sustainability. In that case ρ picks up short-term policy fluctuations and is not suitable for our analysis. The same logic applies when dropping an arbitrary post-war sub-sample. In developing our early-warming indicator we estimated our FRF for a post-war sample that ends in 2007, instead of 2011. This specification intends to eliminate the possible effects of the current financial crisis on our estimations. We find that for only three countries the significance of the ρ parameter is changed: Spain has now a highly significant ($p < 0.01$) coefficient of 0.07, Portugal has a lower significant ($p < 0.1$) coefficient of 0.02, and the significance level of the German coefficient is increased.

5 Stochastic debt sustainability simulations

The results presented in the previous section are based on the assumption that interest and growth rates are equal to their long-run average. However, interest and growth rates fluctuate over time and this has a significant impact on the evolution of the debt level. Higher interest and growth rate volatility increases the distribution of future debt levels and requires a larger fiscal response to keep government debt under control. To assess this relationship we extend the results from the previous section by simulating future interest and growth rate values, which in turn provide a probability distribution for future debt-to-GDP levels.

Specifically, we insert simulated interest and growth rates, represented by $\hat{\gamma}_t$, into equation (3):

$$d_{t+1} = \hat{\gamma}_t(1-\rho)d_t - \hat{\gamma}_t\alpha, \tag{6}$$

where by construction $E(\mathbf{Z_t}) = E(\varepsilon_t) = 0$. To obtain a path for the public debt level d_t, we simulate equation (6).[24] In each step interest and growth rates are obtained

approach probably overestimates temporary military spending by a constant term, which likely has no impact on ρ.

[24]For time $t+1$ we use d_t at the start of period t and the interest and growth rates in period t with equation (6) to get d_{t+1}, the debt level at the start of period $t+1$. Then, for time $t+2$ we

from a simple two variable VAR model following Budina and van Wijnbergen (2008). This VAR model captures the historic volatility of interest and growth rates:

$$\begin{aligned} \begin{pmatrix} r_t \\ y_t \end{pmatrix} &= \alpha_0 + \sum_{j=1}^{\infty} A_j \begin{pmatrix} r_{t-j} \\ y_{t-j} \end{pmatrix} + \eta_t, \\ \text{var}\left(\eta_t\right) &= \mathbf{V} \end{aligned} \tag{7}$$

In this set-up shocks to real interest and growth rates are not correlated over time but are correlated within the same time period. The interest and growth rate themselves are correlated over time and within the same time period due to the auto-regressive specification.

The above procedure describes how to generate a single path for the future debt level. We then run this procedure ten thousand times to obtain a distribution of future debt paths. The set of debt levels from the simulation at time t is then the distribution of expected future debt levels at time t. This distribution is informative on fiscal sustainability as a narrower distribution indicates greater certainty on future debt levels and characterizes a country that is more in control of its finances. Higher volatility in interest and growth rates is characterized by larger values in the covariance matrix $\mathbf{V}$ or in other words, lower autocorrelation in interest and growth rates in equation (7)– broadens the distribution of future debt levels, whereas a larger fiscal response narrows the distribution. This happens because fiscal policy for $\rho > 0$ responds to a deviation in the debt level from $\overline{d}$, which is its historical average. Shocks in interest and growth rates that drive the debt level away from $\overline{d}$ are thus in the subsequent period countered by a fiscal response. Similarly, shocks that drive the debt level towards $\overline{d}$ are mitigated by a smaller fiscal response. This effect is stronger if ρ is bigger relative to the interest and growth rate volatility. Furthermore, if the fiscal response is too weak or if the volatility of interest and growth rates is too strong, the width of the distribution may grow without bound over time. In this case the distribution is not properly defined (i.e. bounded) and, as Hall (2013) shows, the debt level is not stationary.

Our simulations generate a debt path for the period: 2012-2031. We run the simulation ten thousand times for two scenarios: one with the econometrically estimated ρ values (cf. Table 1) and another with no fiscal response: $\rho = 0$.[25] Equation (7) is estimated per country using 1987 as a starting point. We thus do not use the historically low real interest rates during financial repression in the preceding decades. In addition, we set the number of lags equal to two in equation (7) following the Akaike information criterion.

The simulated distributions of expected future debt levels are shown in Figures 2 to 4. Note that the left-hand figure always includes a fiscal response ($\rho > 0$).

use the just obtained d_{t+1} and the interest and growth rates in period $t+1$ as inputs for d_{t+2} and so on and so forth.

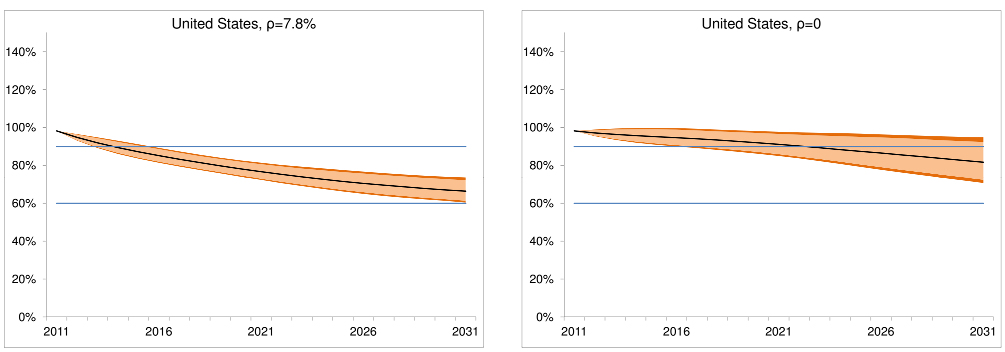
[25]For $\rho = 0$ we use $\alpha = \overline{\text{ps}}$ in equation (6), with $\overline{\text{ps}}$ the average primary surplus. This correction prevents a change in the average fiscal stance while making fiscal policy irresponsive to the debt level.

This implies that for those countries with $\rho = 0$: Spain, Portugal and Iceland, we artificially set $\rho = 7\%$ for illustrative purposes. All right-hand figures have $\rho = 0$. The debt level is on the vertical axes and time is on the horizontal axis. The black line indicates the median debt level from the simulated distribution, the light orange area contains 90% of the simulation results and the dark orange area contains the next 5%. The two blue lines are visual aids at the 60% and the 90% debt level.

The debt levels have relatively small 90% and 95% confidence bands for the first set of countries: the United States, United Kingdom, Netherlands, Belgium and Germany. For the second group –Italy, Spain, Portugal and Iceland– these confidence bands are larger. From Figures 3 and 4 it is clear that the imposed value of $\rho = 7\%$ is not sufficiently large for Italy, Spain nor Portugal to reduce the 95% bandwidth of the simulations.

For $\rho = 0$ we see that the width of the distribution increases vis-à-vis $\rho > 0$ for all countries. Furthermore, for all countries except for the Netherlands the median debt levels increase without a fiscal response.[26]

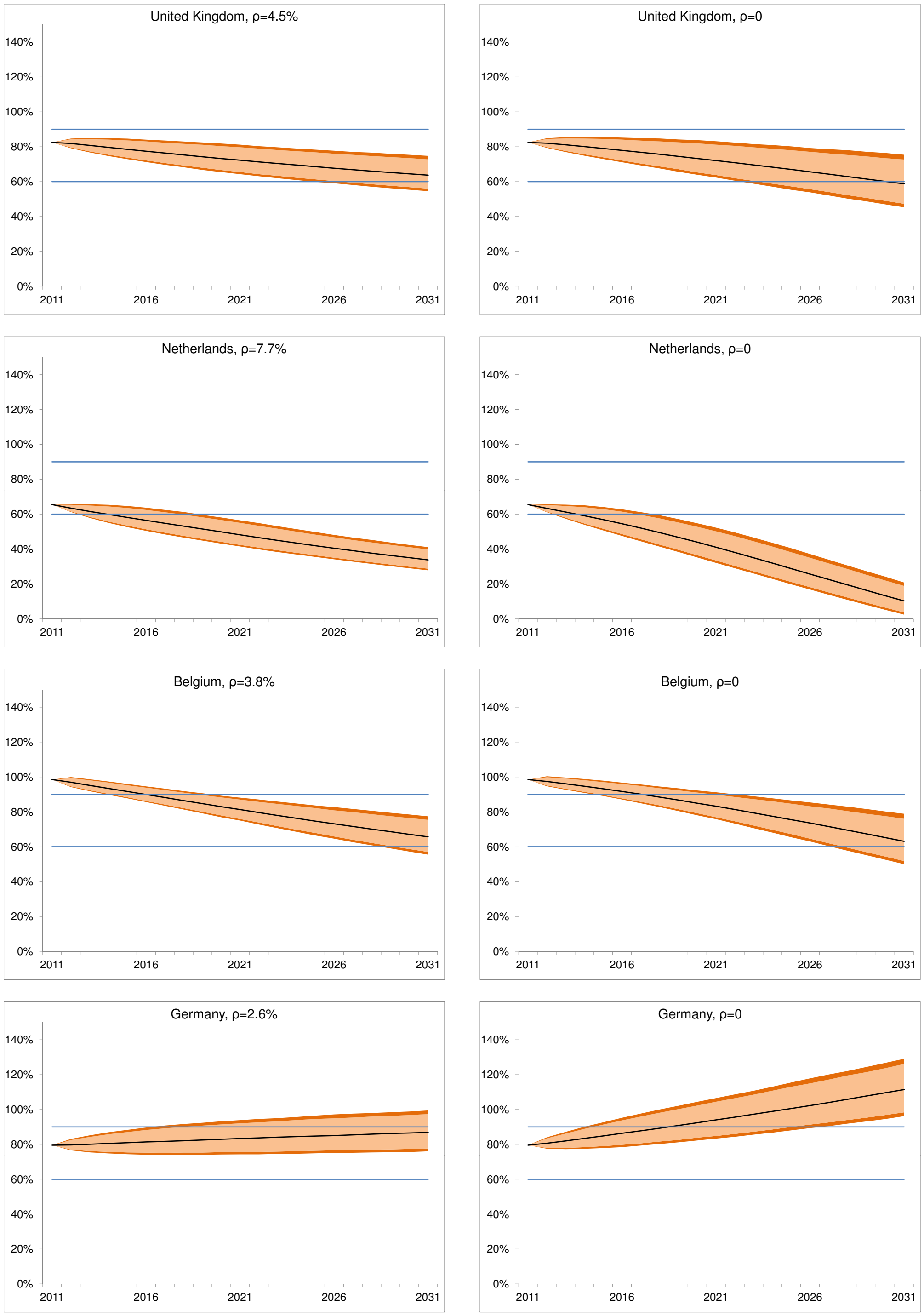
Figure 2: Stochastic debt projections for the United States from 2011 to 2031. ρ denotes the fiscal response.



Notes: The dark orange band with p<0.95 encompasses also the light orange band with p<0.90.

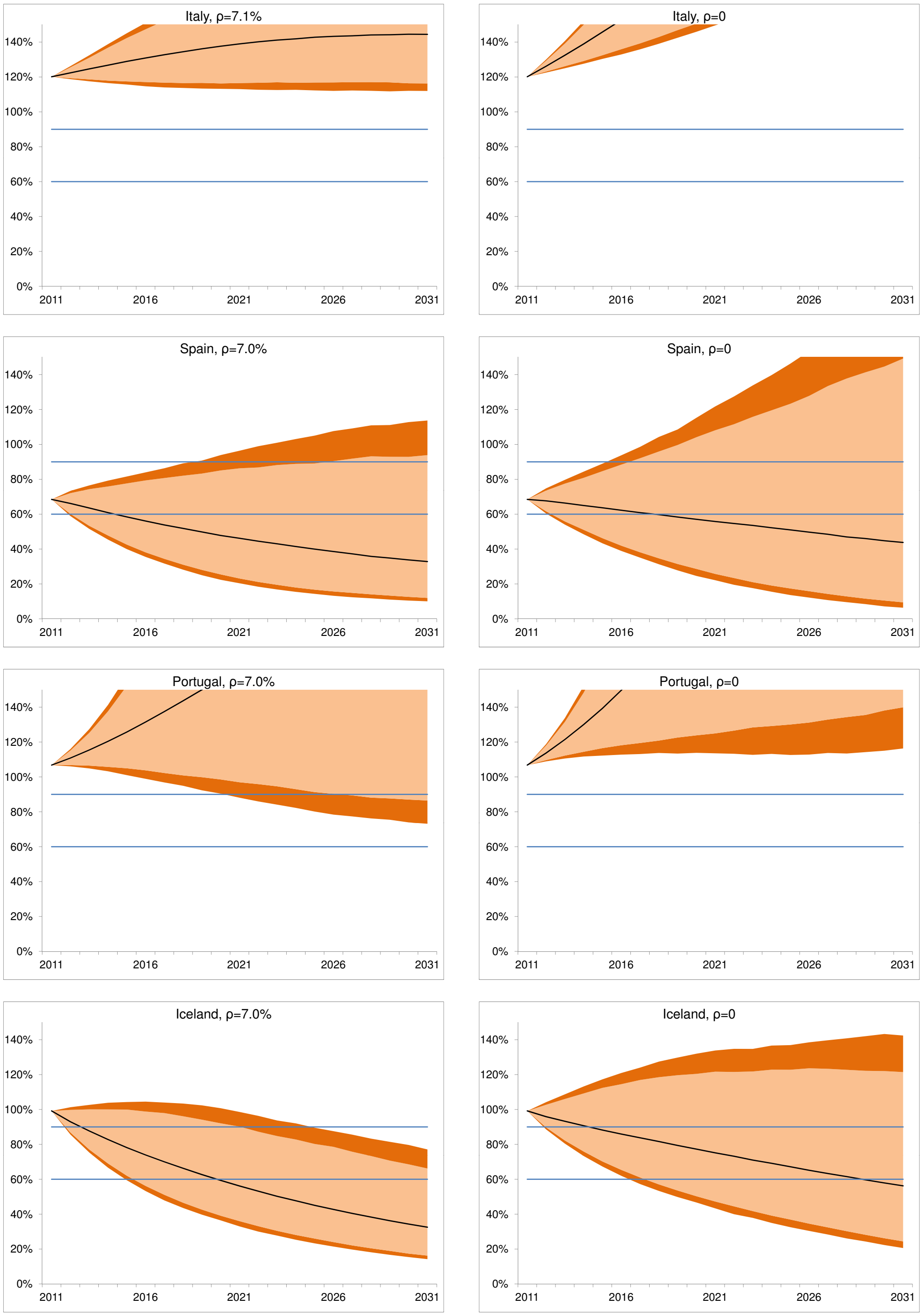
[26]This happens because the Netherlands is characterized by a strong fiscal response and a relatively stable debt level (see Figure 1). The fiscal response reacts to both increases *and* reductions of the debt level from its long-run average, but in the Dutch case the latter applies. Hence, the inclusion of the fiscal response "stabilizes" its debt level around its post-war average of 60%, and this results in the projected median debt level being lower with $\rho = 0$ than with $\rho > 0$.

Figure 3: Stochastic debt projections for the United Kingdom, Netherlands, Belgium and Germany from 2011 to 2031. ρ denotes the fiscal response.



Notes: The dark orange band with p<0.95 encompasses also the light orange band with p<0.90.

Figure 4: Stochastic debt projections for Italy, Spain, Portugal and Iceland from 2011 to 2031. ρ denotes the fiscal response.



Notes: The dark orange band with p<0.95 encompasses also the light orange band with p<0.90.

6 Debt sustainability indicator

A positive and significant fiscal response has two main effects on our debt projections. First, it directly contributes to a more sustainable fiscal policy by lowering expected future debt levels, which in turn reduces the width of the distribution of debt levels –as the value of the product of deviations in the interest and growth rates with the debt levels is reduced. Second, a stronger fiscal response reduces the impact of the effect deviations in interest and growth rates have on the debt level over time by forcing the debt level towards its steady state. Both effects lead to a narrower distribution of expected future debt levels. We capture this in a debt sustainability indicator $X_{95,10}$ which is defined as follows:

$$X_{95,10} = d_{t+10}^{97.5\%} - d_{t+10}^{2.5\%} \tag{8}$$

where $d_{t+10}^{97.5\%}$ is the 97.5th percentile of the simulated distribution of d_{t+10} and $d_{t+10}^{2.5\%}$ is defined analogously. This $X_{95,10}$ indicator, therefore, is the width of the band that contains 95% of the simulated debt levels 10 years into the future.

Table 2 presents the 2011 debt level, the median 2021 debt level and the value of the $X_{95,10}$ indicator both with and without a fiscal response. For the first group of countries: United States, United Kingdom, Netherlands, Belgium and Germany, we find that $X_{95,10}$ has low values. In the case of Germany, the fiscal response is crucial in reducing debt. For all five countries, recent low volatility in interest and growth rates yield relatively narrow confidence bands. The fiscal response (ρ) is instrumental in reducing the $X_{95,10}$ indicator. Without a positive fiscal response (i.e. ρ=0) our indicator increases substantially, in particular for the last four countries in Table 2.

Table 2: Summary of simulation outcomes for 2021

	2011	2021		2021, $\rho = 0$	
	debt level	debt level	$X_{95,10}$	debt level	$X_{95,10}$
United States	102	82	9	95	12
United Kingdom	82	73	16	73	20
Netherlands	65	50	15	44	19
Belgium	99	83	12	85	14
Germany	80	83	19	92	22
Italy	120	137	57	182	84
Spain	68	42	60	56	91
Portugal	107	132	165	199	253
Iceland	99	50	52	78	86

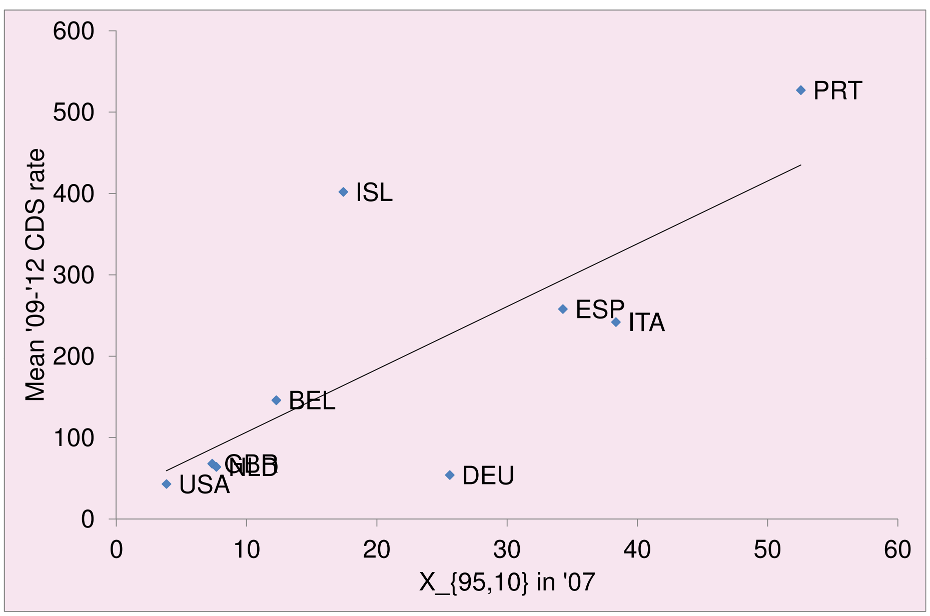
Notes: Units are percent of GDP. $X_{95,10}$ is the width of the band that contains 95% of the simulated debt levels in 2021. We take the ρ values from Table 1, except for Spain, Portugal and Iceland for which we set $\rho = 7\%$.

For the second group of countries: Italy, Spain, Portugal and Iceland we find that our $X_{95,10}$ has high levels, with values ranging between 84% for Italy and a staggering

253% for Portugal. This is caused by high volatility in growth and interest rates and a lack of fiscal response. For these four countries the value of the $X_{95,10}$ indicator drops significantly when $\rho = 7\%$ is assumed. In other words, in the case that these countries have a positive and significant fiscal response, the sustainability indicator is significantly reduced. Yet, it still does not reach the low levels of the other set of countries.

To check whether $X_{95,10}$ can be used as an early-warning indicator, we estimate using only data until 2007 and then run the simulations from 2007 to 2017. Figure 5 plots the $X_{95,10}$ –based on data until 2007– against the average sovereign credit default swaps (CDS) rate between January 2009 and November 2012. We find a high correlation of 0.75 indicating a strong predictive power of our sustainability indicator. On the contrary, in 2007 there was hardly any variation in sovereign CDS spreads. Hence, our $X_{95,10}$ indicator in 2007 was, with the benefit of hindsight, more informative than the market based data on the fiscal stress that occurred after 2008. It clearly shows that Portugal, Iceland, Italy and Spain had high $X_{95,10}$ values in 2007 that could predict the debt sustainability problems that occurred later on.

Figure 5: CDS rates in the crisis versus $X_{95,10}$ prior to the crisis



Source: Rates in basis points for 5 year CDS from Datastream, observation window Jan-09 to Nov-12, indicators in percentages.

7 Summary and conclusions

We develop an indicator for debt sustainability which combines the effect of economic uncertainty –represented by stochastic simulations of interest and growth rates– with the expected response of the government budget to the debt level. We show that this indicator can be useful as an early-warning indicator for debt sustainability, as it would have provided valuable information back in 2007 regarding the European sovereign debt crisis. For the stochastic simulation we follow Budina and van Wijnbergen (2008) and for the expected response of the government we use the fiscal reaction function (FRF) methodology developed by Bohn (1998, 2008). The FRF in this framework reflects country-specific institutions and attitudes towards debt sustainability and is essential to understand the prospects to reduce debt levels in the medium and long-term.

We find that five countries: the United States, United Kingdom, Netherlands, Belgium and Germany have persistently positive and significant fiscal reaction coefficients, conditional on temporary government spending (e.g. war expenditure) and cyclical economic fluctuations. These strong fiscal responses are found in both the full sample and also in the post-war period. In conjunction with average moderate real growth rates (and high rates in some countries for some periods) and some periods of financial repression, these countries have experienced sustainable debt-to-GDP ratios over time. In particular, from the end of the Second World War to around the mid-1970s, the effective use of financial repression mechanisms in these countries allowed for drastic reductions in the debt-to-GDP ratio, which had reached significantly high levels due to military expenditures during the Second World War (specially in the United States and United Kingdom). After the mid-1970s with the end of financial repression and lower real growth rates, these countries relied increasingly on fiscal responsibility (i.e. moderate primary surplus to GDP ratios) to maintain debt at sustainable levels. Debt sustainability in these countries is clearly reflected in the extremely low values of our $X_{95,10}$ indicator.

On the other hand, for Spain, Portugal and Iceland, we do not find significant fiscal reaction coefficients. With the end of financial repression, this means that these countries are less prepared to maintain sustainable debt levels in the future. Our debt sustainability indicator $X_{95,10}$ clearly identifies this weakness in the debt dynamics of these countries by providing large positive probabilities. Finally, Italy has a large positive and significant fiscal response coefficient. Its current debt level, however, is so large that its debt sustainability is susceptible to even small fluctuations in interest and growth rates. Therefore its debt sustainability indicator is large, which justifies the doubts on Italy's debt sustainability.

For medium to long-term fiscal policy assessments, indicators based on stochastic analysis and expected fiscal responses have several advantages over the currently available indicators which are static and do not capture volatility in the economy and the government's ability to control public finances. Alternative indicators currently in use, such as structural balances or cyclically adjusted budget balances (CABB), are often plagued by measurement issues. They depend on projections of future growth, which are known to have an upward bias (Larch and Salto, 2005), and their estimates are vulnerable to endogeneity problems: from the debate on the size of the fiscal multiplier it becomes clear that it is nontrivial to disentangle the effects of expected growth on the CABB from the effects the CABB has on growth. Our indicator do not suffer from these shortcomings, since it incorporates economic volatility and the government's expected policy response from ex-post realizations only.

Our analysis, however, has some caveats. First, our estimated fiscal response (ρ) is an institutional variable that measures how - over medium and long-time periods – the government of a particular country deals with medium/long term changes in debt levels. This means that we require long time series to estimate ρ, and furthermore, our approach is not suitable to analyse short-term debt sustainability. For instance, it cannot provide information on whether -for example - Spain will be able to roll over its debt in the coming months. Second, the shocks in our simulations depend on the historic volatility of interest and growth rates. That means they do not contain unexpected exogenous events (e.g. war, natural disasters). The results of our simulation exercise are not informative on debt sustainability under such conditions, as concerns other than debt sustainability will receive priority under such circumstances. Third, our indicator is not informative on the specific policy change which solves debt sustainability issues. Our framework merely states that countries with a higher and more significant historical fiscal policy response are more likely to solve such issues, should they arise. In contrast, aging studies (European Commission, 2012a) analyse the impact that aging has on public finances given constant policy arrangements. If unsustainability is found, the policy arrangements most impacted by aging should change until the unsustainability is alleviated.

Our indicator is not meant to replace current indicators, but rather to complement the short term indicators (i.e. government yearly deficits and debt-to-GDP levels) and the indicators from aging studies. They could, for instance, provide guidance on whether it is reasonable for a country to join a monetary union. In such a union, the use of financial and monetary policies is limited for individual countries, making it unlikely for them to achieve debt reductions through policies that yield very low or negative real interest rates. In this case there is an increased dependence on fiscal policy to tackle debt sustainability. It is precisely this medium- to long-term institutional relation between fiscal policy and debt sustainability that is captured by our indicator.

References

Aaron, H. (1966). "The Social Insurance Paradox", *Canadian Journal of Economics and Political Science* 32(3): 371–374.

Afonso, A. (2005). "Fiscal Sustainability: The Unpleasant European Case", *FinanzArchiv* 61(1): 19–44.

Arezki, R., Candelon, B. and Sy, A. (2011). "Sovereign Rating News and Financial Markets Spillovers: Evidence from the European Debt Crisis", IMF Working Paper 11/69.

Baffigi, A. (2011). "Italian National Accounts, 1861-2011,", Economic History Working Papers 18, Banca d'Italia.

Bai, J. and Perron, P. (1998). "Estimating and Testing Linear Models with Multiple Structural Changes", *Econometrica* 66(1): 47–78.

Bai, J. and Perron, P. (2003). "Computation and Analysis of Multiple Structural Change Models", *Journal of Applied Econometrics* 18(1): 1–22.

Barro, R. J. (1979). "On the Determination of Public Debt", *Journal of Political Economy* 87(5): 940–971.

Bartoletto, S., Chiarini, B. and Marzano, E. (2012). "The Sustainability of Fiscal Policy in Italy: A Long-Term Perspective", CESifo Working Paper Series 3812.

Baum, A., Checherita-Westphal, C. and Rother, P. (2013). "Debt and growth: New Evidence for the Euro Area", *Journal of International Money and Finance* 32: 809–821.

Beetsma, R. and Giuliodori, M. (2010). "The Macroeconomic Costs and Benefits of the EMU and Other Monetary Unions: An Overview of Recent Research", *Journal of Economic Literature* 48(3): 603–641.

Berti, K. (2013). "Stochastic public debt projections using the historical variance-covariance matrix approach for EU countries", European Economy - Economic Papers 480, Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission.

Bohn, H. (1991). "Budget balance through revenue or spending adjustments?: Some historicalevidence for the United States", *Journal of Monetary Economics* 27(3): 333–359.

Bohn, H. (1998). "The Behavior of U.S. Public Debt and Deficits", *Quarterly Journal of Economics* 113(3): 949–963.

Bohn, H. (2007). "Are Stationary and Cointegration Restrictions Really Necessary for the IntertemporalBudget Constraing?", *Journal of Monetary Economics* 54(7): 1837–1847.

Bohn, H. (2008). "The Sustainability of Fiscal Policy in the United States", *in* R. Neck and J. Sturm (eds.), *Sustainability of Public Debt*, MIT Press, pp. 15–49.

Bos, F. (2007). "The Dutch fiscal framework; history, current practice and the role of theCPB", *CPB Discussion Paper* 150.

Brender, A. and Drazen, A. (2005). "Political budget cycles in new versus established democracies", *Journal of Monetary Economics* 52(7): 1271–1295.

Brender, A. and Drazen, A. (2008). "How Do Budget Deficits and Economic Growth Affect Reelection Prospects? Evidence from a Large Panel of Countries", *American Economic Review* 98(5): 2203–20.

Budina, N. and van Wijnbergen, S. (2008). "Quantitative Approaches to Fiscal Sustainability Analysis: A Case Study of Turkey since the Crisis of 2001", *World Bank Economic Review* 23(1): 119–140.

Callen, T., Terrones, M., Debrun, X., Daniel, J. and CelineAllard (2003). "Public Debt in Emerging Markets: Is is Too High?", *World Economic Outlook*, International Monetary Fund, chapter III.

CBS (1959). "Zestig jaren statistiek in tijdreeksen (1899-1959)".

CBS (1994). "Vijfennegentig jaren statistiek in tijdreeksen (1899-1994)".

CBS (2001). "Tweehonderd jaar statistiek in tijdreeksen (1800-1999)".

Cecchetti, S., Mohanty, M. and Zampolli, F. (2011). "The real effects of debt", BIS Working Papers 352, Bank for International Settlements.

Checherita-Westphal, C. and Rother, P. (2012). "The Impact of High Government Debt on Economic Growth and its Channels: An Empirical Investigation for the Euro Area", *European Economic Review* 56(7): 1392–1405.

Comín, F. and Díaz, D. (2005). "Sector Público Administrativo y Estado del Bienestar", *in* A. Carreras and X. Tafunell (eds.), *Estadísticas Históricas de España: Siglos XIX y XX*, 2nd edition edn, Fundación BBVA, Bilbao, Spain, pp. 873–964.

Dehejia, V. H. and Rowe, N. (1995). "The Laffer Curve and Precautionary Taxation: A Rationale for Paying Down the National Debt", *Australian Economic Papers* 34(65): 322–31.

Égert, B. (2012). "Public Debt, Economic Growth and Nonlinear Effects: Myth or Reality?", OECD Economics Department Working Papers 993, OECD Publishing.

European Commission (2012a). “Fiscal Sustainability Report 2012”, European Economy 8/2012, Directorate-General for Economic and Financial Affairs, Brussels.

European Commission (2012b). “Report on Public Finances in the EMU”, European Economy 4/2012, Directorate-General for Economic and Financial Affairs, Brussels.

Ghosh, A. R., Kim, J. I., Mendoza, E. G., Ostry, J. D. and Qureshi, M. S. (2013). “Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies”, *The Economic Journal* 123(566): F4–F30.

Hall, R. E. (2013). “Fiscal Stability of High-Debt Nations under Volatile Economic Conditions”, Working Paper 18797, National Bureau of Economic Research.

Höppner, F. and Kastrop, C. (2004). “Fiscal Institutions and Sustainability of Public Debt in Germany”, *Banca d'Italia Workshop on Public Debt*, Banca d'Italia, Perugia, Italia, pp. 575–594.

Krugman, P. R. (1988). “Financing vs. Forgiving a Debt Overhang”, *Journal of Development Economics* 29(3): 253–268.

Larch, M. and Salto, M. (2005). “Fiscal rules, inertia and discretionary fiscal policy”, *Applied Economics* 37(10): 1135–1146.

Lejour, A., Lukkezen, J. and Veenendaal, P. (2011). “Sustainability of Government Debt in the EMU”, *in* W. Meeuwen (ed.), *The Economic Crisis and European Integration*, Edward Elgar, Cheltenham, UK, pp. 35–54.

Maddison, A. (2003). *The World Economy, historical satistics*, OECD, Development Centre Studies, Paris, France.

Marinheiro, C. F. (2006). “The Sustainability of Portuguese Fiscal Policy from a Historical Perspective”, *Empirica* 33(2-3): 155–179.

Medeiros, J. (2012). “Stochastic debt simulation using VAR models and a panel fiscal reaction function - results for a selected number of countries”, European Economy - Economic Papers 459, Directorate General Economic and Monetary Affairs (DG ECFIN), European Commission.

Mendoza, E. G. and Ostry, J. D. (2008). “International evidence on fiscal solvency: Is fiscal policy "responsible"?”, *Journal of Monetary Economics* 55(6): 1081–1093.

Michell, B. R. (1988). *British historical statistics*, Cambridge University Press, Cambridge, UK.

OECD (2005). “Measuring Cyclically-adjusted Budget Balances for OECD Countries”, *OECD Working Paper* 434.

Peacock, A. T. and Wiseman, J. (1961). *The Growth of Public Expenditure in the United Kingdom*, NBER Books, National Bureau of Economic Research.

Pirard, J. (1999). *L'extension du rôle de l'Etat en Belgique aux XIXe et XXe siècles*, publisher, Brussel, Belgium.

Prados de la Escosura, L. (2003). *El Progreso Económico de España (1850-2000)*, Fundación BBVA, Bilbao, Spain.

Reinhart, C. M. and Rogoff, K. S. (2010). "Growth in a Time of Debt", Working Paper 15639, National Bureau of Economic Research.

Reinhart, C. M. and Sbrancia, M. B. (2011). "The Liquidation of Government Debt", NBER Working Paper 16893, National Bureau for Economic Research.

van Zanden, J. L. (1996). "The development of Government Finances in a chaotic period, 1807-1850", *Economic and Social History in the Netherlands* 7: 53–71.

Wierts, P. and Schotten, G. (2008). "De Nederlandse gasbaten en het begrotingsbeleid: theorie versus praktijk", *DNB Occasional Studies* 6(5).

Appendix

A Data sources and country graphs

Under the subsequent country headings we describe our data sources, elaborate on the definitions used (general/central government) and whether breaks in the data are present. Table 3 presents a summary of the available data for each country.

Table 3: Available data per country

Country	Samples			Observations
USA	1792-2011			218
United Kingdom	1691-2011			321
Netherlands	1816-1939	1948-2011		188
Belgium	1830-1913	1919-1939	1955-2011	157
Germany	1970-2011			42
Italy	1861-2011			150
Spain	1850-1935	1940-2011		156
Portugal	1852-2011			160
Iceland	1908-2011			104

In Figures 8-12 we plot the debt-to-GDP ratio, the primary surplus to GDP ratio, a smoothed series of interest minus growth rate and for the United Kingdom and the US the GVAR indicator.

As the effective interest rate we use the simple formula: $i_t^e = ip_t/d_{t-1}$ where i^e is the effective interest rate and ip is interest payment on debt d, both in nominal terms.[27]

United States

We use the data from Bohn (2008) from 1792-2009.[28] We updated this data until 2011. This is a continuous dataset on nominal and real GDP, government gross debt, government primary surplus, government interest expenditure and government military expenditure. A detailed description of the data used there can be found in Bohn (1991) and Bohn (2008), and references therein.

Note that the data by Bohn provides gross federal debt levels for the United States that are in the hands of the public. It does not include other governmental holdings of federal debt. In our analysis we assume the gross federal debt not in hands of the public remains constant as a share of GDP.

[27] This simple formula is almost perfectly correlated (0.9975) with a more precise specification given by: $i_t^e = \frac{ip_t}{2}\left(\frac{d_t}{\sqrt{\pi_t}} + \frac{d_{t-1}}{\sqrt{\pi_t}}\right)^{-1}$ where p is the GDP deflator and $\pi_t = p_t/P_{t-1}$.

[28] Henning Bohn kindly provided us with an updated database which runs until 2009. He used the 2011 Budget of the United States for this update.

Government military expenditure is used as an indicator for temporary government spending GVAR as government expenditure in wartime is significantly different from government spending in peacetime. Using dummy variables for the war years[29] instead of actual military expenditure does not change the results significantly.

United Kingdom

The sample from the United Kingdom is from 1692-2011 with a break in 1800 and in 1946. There is a shift in the reporting year in 1800, prior to 1946 we use central government data and after 1946 we use general government data. We use military expenditure as a proxy for temporary government spending and there is a break in this series in 1980 probably due to a different specification. Prior to 1946 military expenditure and interest expenditure are the only two large items on the central governments budget. Non-military non-interest spending never exceeds 2% of GDP.[30] Data is obtained from six sources:

- Data on central government expenditure (1692-1945), interest expenditure (1692-1945), military expenditure (1692-1980), revenue (1692-1946), public debt (1692-1979), CPI as a proxy for the GDP deflator (1691-1792), a GDP deflator (1830-1946) and nominal GDP (1830-1946) has been obtained from Michell (1988).
- From Peacock and Wiseman (1961) real and nominal GDP is obtained for the years 1792, 1800, 1814, 1822, 1831. Constant growth rates in real and nominal GDP are assumed between these dates.
- Real GDP for 1700 and 1800 is obtained from Maddison (2003) world historic tables. Constant growth rates in real GDP is assumed between 1700 and 1792 and in the period 1691-1700.
- General government expenditure, interest expenditure, revenue and real and nominal GDP from 1946 onwards is obtained from the Office of National Statistics
- Public debt is obtained from the United Kingdom Ministry of Finance for 1980-2011
- Military expenditure for 1980-2011 is obtained from the OECD.

A GDP deflator and real GDP from 1692-2011 is obtained by coupling the various partial GDP series with each other. Furthermore GDP has been adjusted such that Ireland is excluded prior to 1920 as it is after 1920. Temporary government expenditure GVAR is military spending. As an alternative we use gap between the log of government expenditure and its trend.

[29] 1812-1815 War of 1812, 1846-1849 (Mexican-American War), 1861-1865 (American Civil War), 1917-1920 (World War I), 1940-1945 (World War II)

[30] Except for 1836 and 1837. In these years the government compensated slave owners for outlawing slavery.

Netherlands

Data is obtained from Bos (2007) and contains nominal GDP, a piecewise continuous GDP deflator, gross government debt and a decomposition of government revenue and government expenditure in their main components. It runs from 1815 till 2009 with gaps in the inter-war years. Bos acquires data on general government finances in the period 1815-1900 from the work of van Zanden (1996) and from 1900 onwards from Statistics Netherlands. In the period 1850-1900 only data on central government finances are available. Furthermore data on local government interest expenditure is missing until 1947.

We correct for that by using two assumptions. First, we assume that the interest rate on non-central government debt equals the interest rate on central government debt. In the Netherlands, the central government steps in and assumes full liability when local governments are in financial distress. Therefore local government default risk is equal to central government default risk. Second we assume that local government finances have run a balanced budget, as they are required by law, and we interpolate non-central government debt between 1850 and 1900 linearly. This seems a reasonable first assumption as in 1850 non-central government debt is 26.0% of general government debt, in 1900 it is 20.2%. If these assumptions underestimate local interest expenditure, primary surplus and implied interest rates prior to 1947 would be lower than their actual value. The effect on debt sustainability will be absent, as primary surplus and the implied interest rates have opposite signs in the accounting equation. This has been tested by using central instead of general government finances. As none of the regression coefficients except for the constant to changed by more than one standard deviation, we deem these assumptions reasonable.

Bos (2007) provides GDP deflators from 1815-1913, 1921-1939 and 1948-2009, Statistics Netherlands (CBS, 1959, 1994, 2001) provides a consumer price index from 1900-2009. A continuous GDP deflator is constructed by using consumer price indices to bridge the gap between the broken piece-wise continuous GDP deflators. We approximate the GDP deflator from 1913-1921 and from 1939-1948 by the consumer price index. The consumer price index is highly correlated with the GDP deflator in the period 1900-1913, 1922-1939 and 1949-2009: correlation is 0.998 on level and 0.949 on first differences.

Temporary government expenditure is defined as the residual of government expenditure after its HP-filtered mean has been removed. Note that Bohn (2008) uses military expenditure as an alternative measure of temporary government expenditure. Unlike the United States, where military spending drives government spending prior to 1948, the only notable Dutch event is the Belgian war of independence in 1830. Wierts and Schotten (2008) argue gas revenue should be used from 1970s onwards as it had considerable impact in budgetary policy. Both alternative specifications are used in robustness checks and do not provide significant changes.

Belgium

The sample from Belgium is from 1830-1913, 1920-1939 and 1955-2011 with a break in 1970. Prior to 1970 central government data is used, after 1970 general government data. Data for Belgium has been obtained from 4 sources:

- A dataset on Belgium's central government finances from the independence of the state in 1830 until the first world war (1913) was created by Joseph Pirard and published in Pirard (1999). He reports central government revenue, expenditure, gross government debt, interest expenditure nominal GDP from 1830 onwards. In this book Pirard also publishes this data for the Inter bellum (1920-1939) and the years after the Second World War (1945-1995), which he obtains from other sources. finance data after 1945 as the increase in debt is always smaller than the difference between government revenue and government expenditure and much to persistent to be due to stock-flow adjustments. This might be due to the fact that some government bond redemptions are classified as government expenditure.

- Real GDP for the period 1830-1960 is obtained from Maddison (2003) world historic tables. For the years 1831-1839 no data is available here and thus a constant increase between 1830 and 1840 is assumed.

- Data on central government finances (revenue from 1955 until 1970 is obtained from the annual reports of the NBB, the Belgian central bank.

- From 1970 onwards data on general government finances and GDP is available from the AMECO database of the European Commission.

The nominal GDP estimates of Pirard are in the period 1970-1995 approximately 13% lower than the AMECO data. We correct for this by increasing every data-point in the nominal GDP series of Pirard by 13%. Temporary government expenditure GVAR is determined as the gap between the log of government expenditure and its trend.

Germany

We use the data from Höppner and Kastrop (2004), which was updated for 2011 by Elke Baumann. The have GDP and debt series starting in 1960, but only fiscal data starting from 1970. From 1980 to 2011 we use the IMF World Economic Outlook (April 2012) data on General Government expenditure, revenue, and budgetary surplus.

Italy

We use the data from Bartoletto *et al.* (2012) from 1861-2009 for consolidated debt of the General Government, total expenditure and interest payments. This is complemented with data from Baffigi (2011) for time series on nominal and real GDP.

Furthermore, we update these data until 2011 using the IMF World Economic Outlook (April 2012) to obtain data on real and nominal GDP, General Government gross debt, expenditure, revenue and primary surplus.

Since there was no data in Bartoletto *et al.* (2012) on governmental revenue, the budgetary surplus was estimated using the difference in gross debt with respect to its previous year. From 1988 onwards we substitute the historical data with the IMF WEO data on primary surplus. The correlation between both primary surplus sources is 0.86 and our econometric results are qualitatively the same.

Spain

The sample is from 1850 to 2011 with a gap for the Spanish Civil War (1936-1939). Data prior to 1995 is taken from Prados de la Escosura (2003) for GDP data in real and nominal terms and Comín and Díaz (2005) for the public sector data and concerns national and provincial government finances. After 1995 data is obtained from the AMECO database of the European Commission.

Temporary government expenditure GVAR is determined as the gap between the log of government expenditure and its trend.

Portugal

The sample is from 1850 to 2011. Data comes from Marinheiro (2006) for the period 1852-1995. In the statistical appendix to that paper Marinheiro describes the sources from which he obtains his data. This is a continuous dataset on nominal and real GDP, government gross debt, government primary surplus and government interest expenditure. The government finances are on cash basis. After 1995 data is obtained from the AMECO database of the European Commission.

Marinheiro (2006) also constructed interest rates for Portugal which are about three percentage points higher on average than the ones estimated using actual interest payments. However, our results did not change qualitatively using this alternative.

Temporary government expenditure GVAR is determined as the gap between the log of government expenditure and its trend. Portugal defaulted on its government debt in 1892, which was ultimately resolved in 1902.

Iceland

The sample for Iceland is from 1908-2011. Data for Iceland has been obtained from two sources. A since 2000 defunct Icelandic organization "Þjóðhagsstofnun" published general government revenue, expenditure, gross government debt, interest expenditure and real and nominal GDP from 1908 until 1999. Iceland statistics publishes data on general government revenue, expenditure, gross government debt, interest expenditure and real and nominal GDP from 1945 onwards. The data between 1945 and 1999 is identical to the data on the "Þjóðhagsstofnun" website.

Icelandic data concerns general government and contains long periods of high inflation. Temporary government expenditure GVAR is determined as the gap between the log of government expenditure and its trend. Iceland sought and received assistance from the IMF and the Scandinavian countries after the 2008 banking crisis turned into a sovereign debt crisis for Iceland. This is considered as a public debt default.

Figure 6: Real growth rates (smoothed series) in the post-war period.

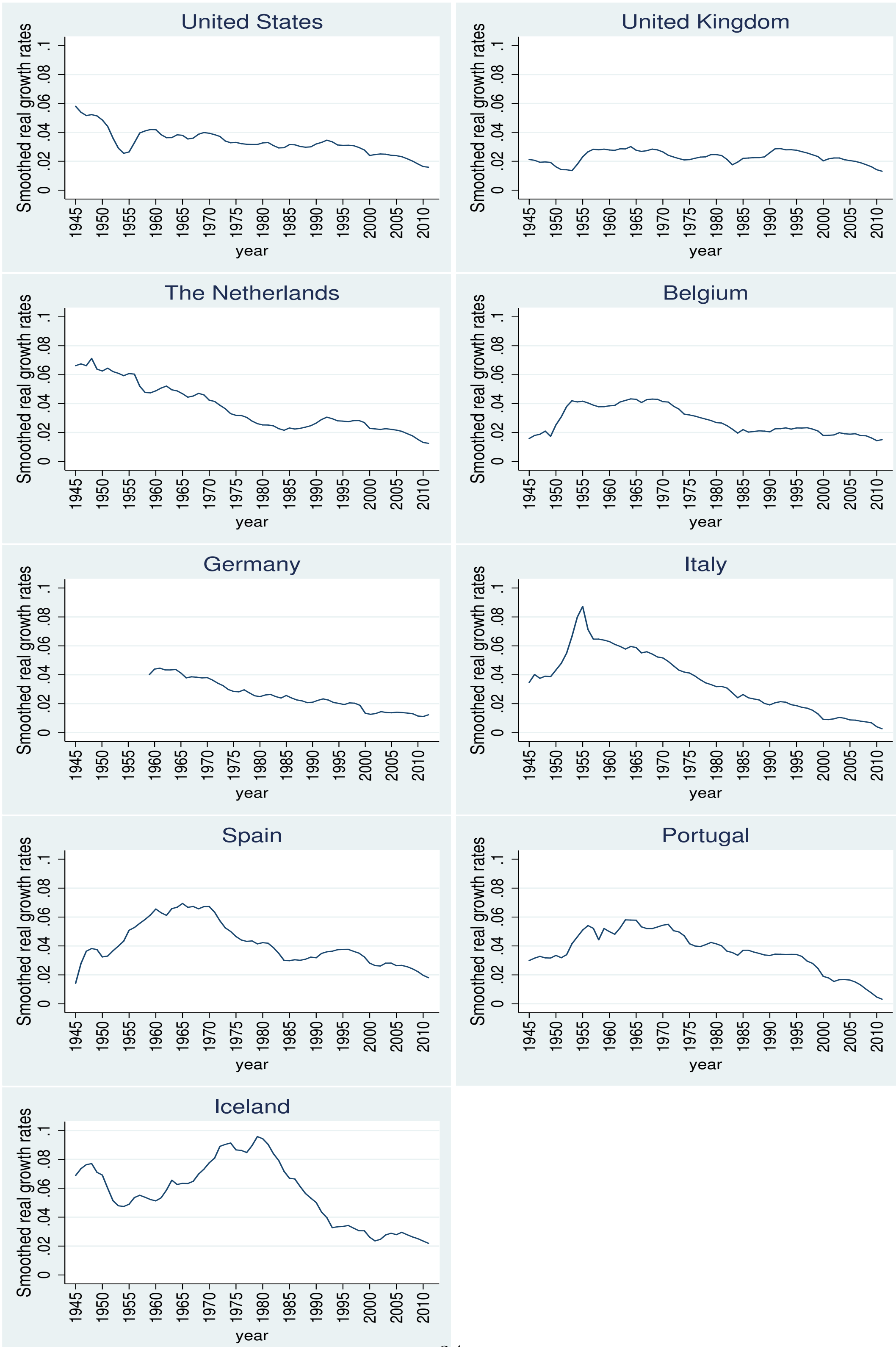


Figure 7: Financial repression: interest rates (straight line) and inflation (dotted line) in the post-war period (the vertical scale can be different across countries).

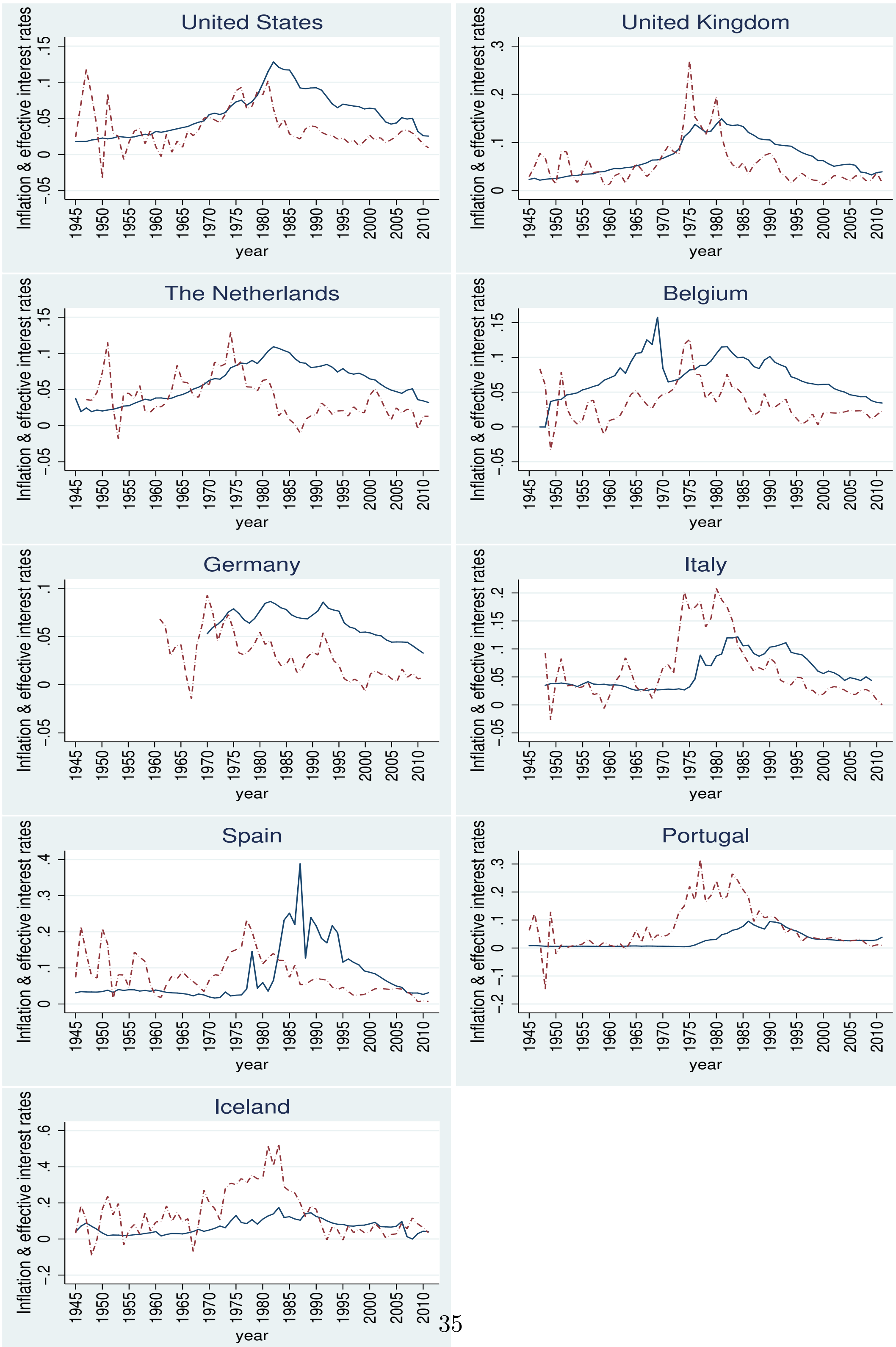
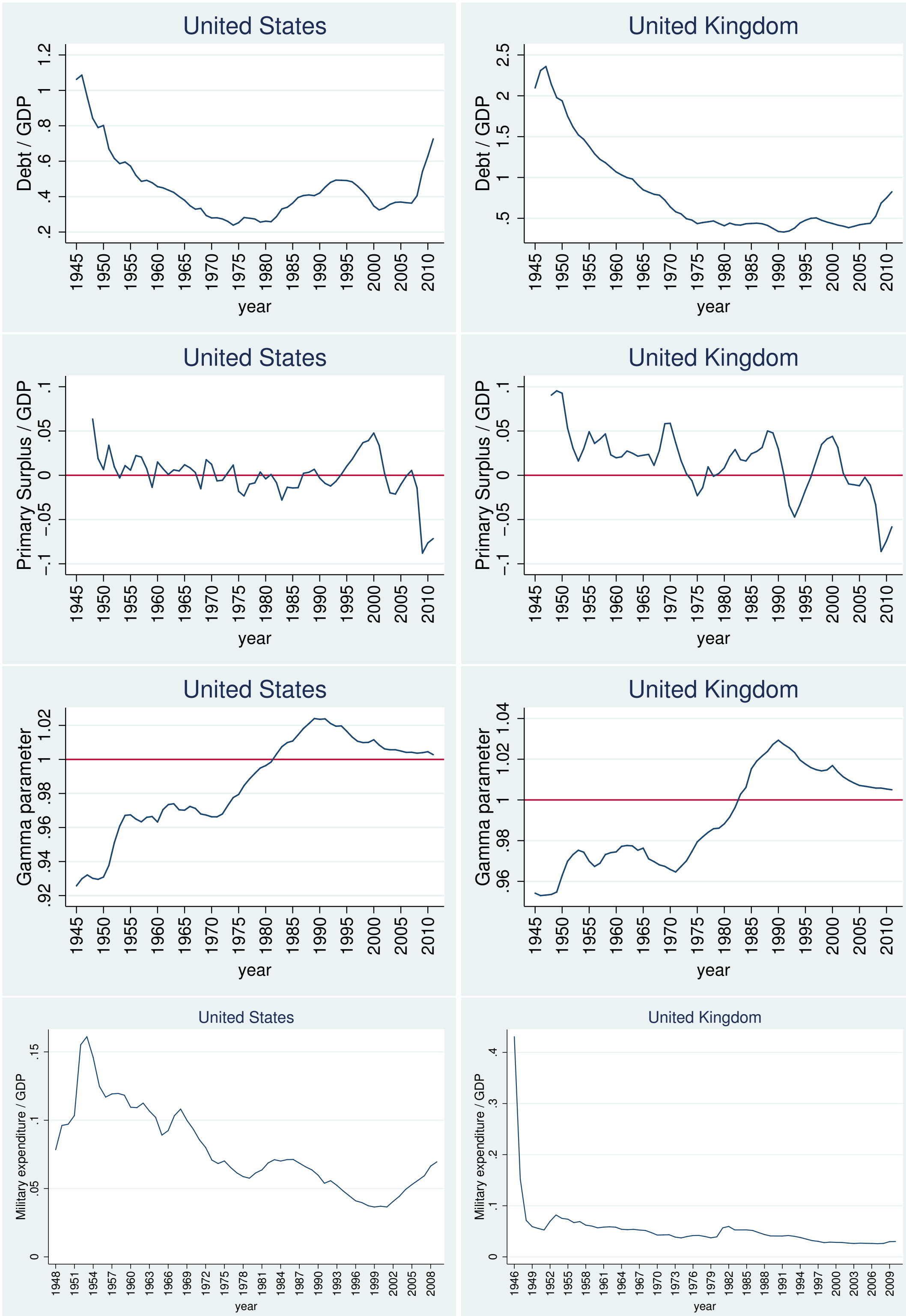
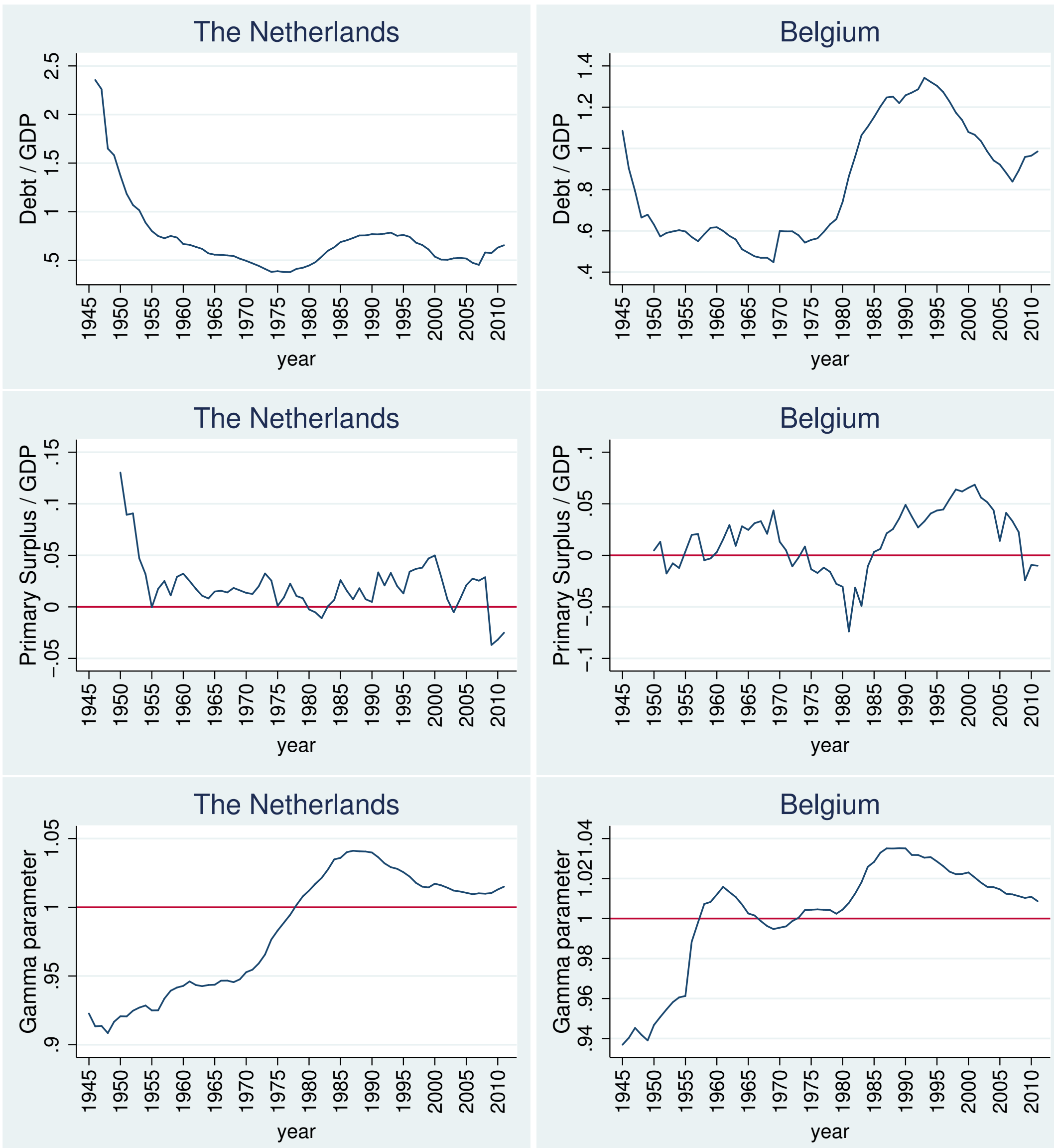


Figure 8: United States and United Kingdom: Debt, primary surplus and military expenditure ratios to GDP, and gamma parameter, post-war samples



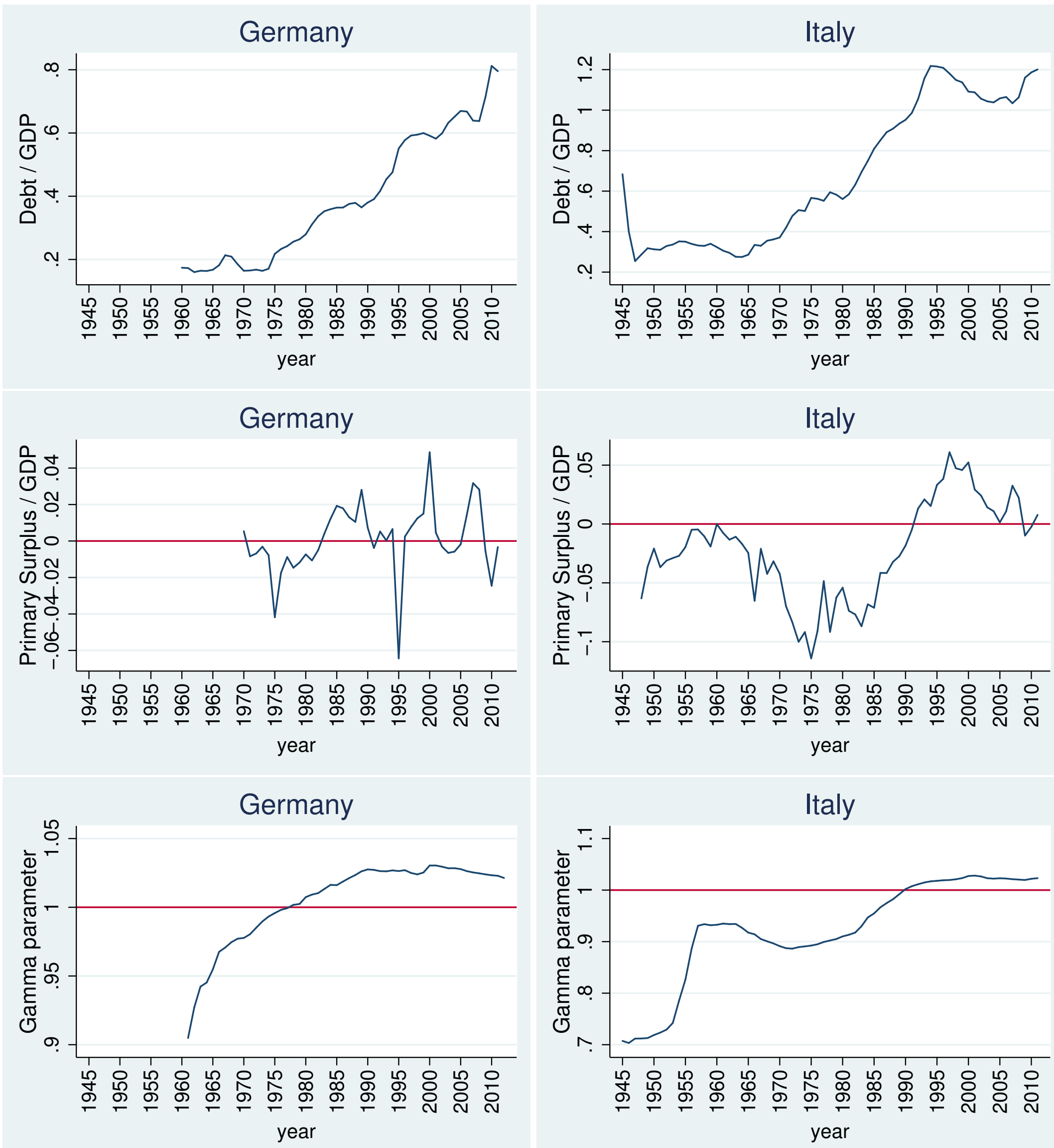
Notes: The *gamma* parameter is one plus the smoothed series of the effective interest rates minus the smoothed series of nominal GDP growth rates.

Figure 9: Netherlands and Belgium: Debt and primary surplus ratios to GDP, and gamma parameter, post-war samples



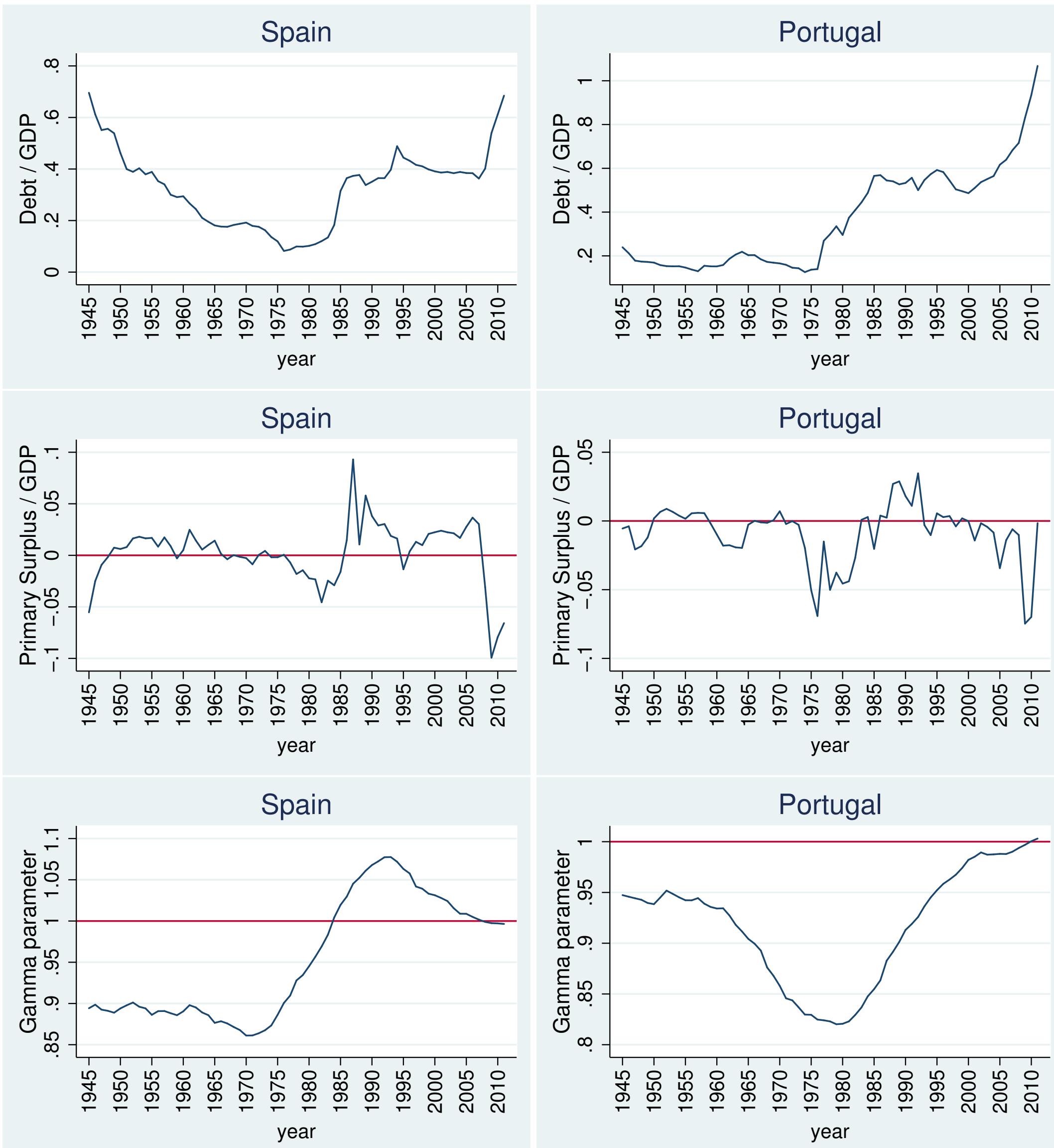
Notes: The *gamma* parameter is one plus the smoothed series of the effective interest rates minus the smoothed series of nominal GDP growth rates.

Figure 10: Germany and Italy: Debt and primary surplus ratios to GDP, and gamma parameter, post-war samples



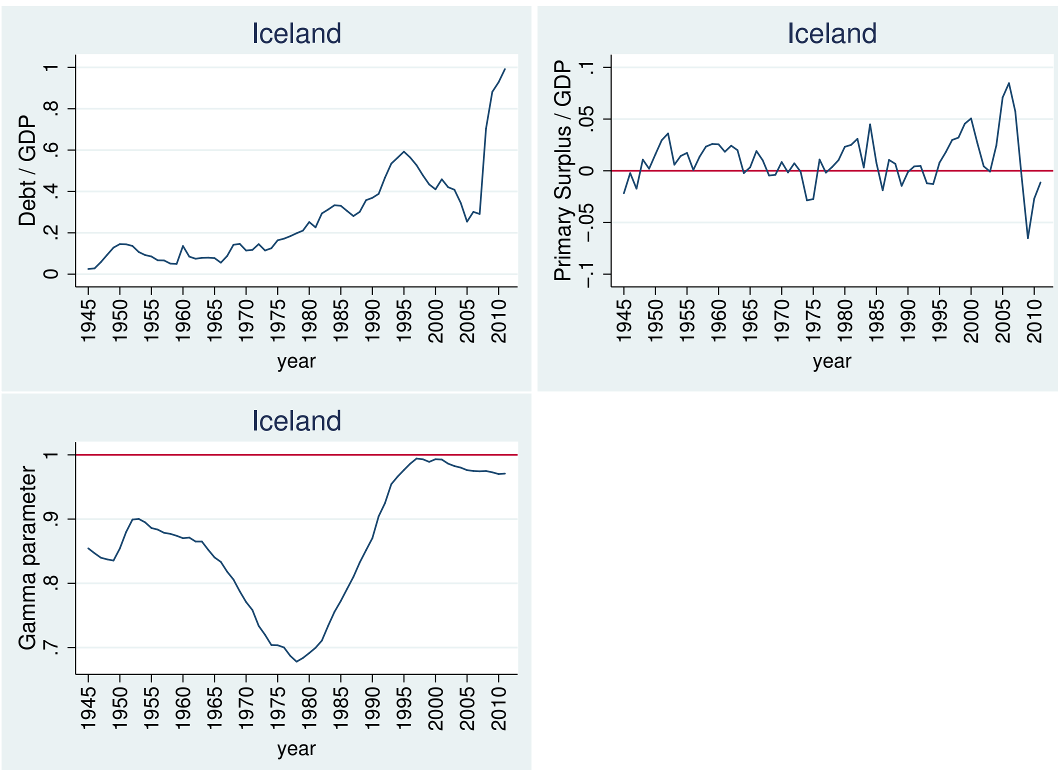
Notes: The *gamma* parameter is one plus the smoothed series of the effective interest rates minus the smoothed series of nominal GDP growth rates.

Figure 11: Spain and Portugal: Debt and primary surplus ratios to GDP, and gamma parameter, post-war samples



Notes: The *gamma* parameter is one plus the smoothed series of the effective interest rates minus the smoothed series of nominal GDP growth rates.

Figure 12: Iceland: Debt and primary surplus ratios to GDP, and gamma parameter, post-war sample



Notes: The *gamma* parameter is one plus the smoothed series of the effective interest rates minus the smoothed series of nominal GDP growth rates.

B Regression results

Table 4: United States, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1) 1948-2009	(2) 1948-2009	(3) 1948-2009	(4) 1948-2009	(5) 1948-2009	(6) 1792-2011
debt	0.078 ***	0.067 ***	0.077 ***	-0.150	0.034	0.079 ***
	[0.026]	[0.021]	[0.023]	[0.127]	[0.035]	[0.024]
	(0.024)	(0.024)	(0.024)	(0.099)	(0.029)	(0.013)
yvar	0.284 **	0.267 **	0.285 **	0.391 **	0.358 **	0.072
	[0.120]	[0.123]	[0.124]	[0.150]	[0.142]	[0.065]
	(0.125)	(0.124)	(0.124)	(0.129)	(0.124)	(0.049)
mil	-0.083		-0.136	-0.081	-0.078	-0.707 ***
	[0.117]		[0.131]	[0.121]	[0.122]	[0.115]
	(0.092)		(0.097)	(0.088)	(0.088)	(0.051)
gvar		0.075				
		[0.227]				
		(0.291)				
rir			-0.140			
			[0.091]			
			(0.093)			
debt_sq				0.238 *		
				[0.128]		
				(0.101)		
debt_60					0.044 *	
					[0.025]	
					(0.018)	
Constant	-0.023 **	0.019	-0.015	0.027	-0.007	0.006 **
	[0.009]	[0.131]	[0.009]	[0.025]	[0.010]	[0.003]
	(0.010)	(0.172)	(0.011)	(0.023)	(0.011)	(0.003)
Observations	62	62	62	62	62	218
R-squared	0.160	0.149	0.192	0.235	0.236	0.559

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "mil" is military expenditure as a percentage of nominal GDP, "rir" is the effective real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 5: United Kingdom, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1) 1946-2011	(2) 1946-2011	(3) 1946-2011	(4) 1946-2011	(5) 1946-2011	(6) 1792-2011
debt	0.045 ***	0.023 **	0.052 ***	-0.001	0.096 *	0.049 ***
	[0.010]	[0.009]	[0.011]	[0.035]	[0.055]	[0.002]
	(0.009)	(0.008)	(0.010)	(0.032)	(0.035)	(0.002)
yvar	0.410	0.354	0.446 *	0.474	0.413	-0.223 *
	[0.266]	[0.246]	[0.266]	[0.291]	[0.255]	[0.122]
	(0.172)	(0.161)	(0.171)	(0.176)	(0.170)	(0.052)
mil	-0.143		-0.151 *	-0.200 ***	-0.171 **	-0.706 ***
	[0.092]		[0.087]	[0.075]	[0.078]	[0.059]
	(0.091)		(0.090)	(0.098)	(0.092)	(0.019)
gvar		-0.211 ***				
		[0.064]				
		(0.063)				
rir			0.180 *			
			[0.099]			
			(0.113)			
debt_sq				0.021		
				[0.017]		
				(0.015)		
debt_60					-0.037	
					[0.039]	
					(0.025)	
Constant	-0.011	-0.004	-0.019 *	0.009	-0.030	0.009 ***
	[0.010]	[0.009]	[0.011]	[0.015]	[0.025]	[0.003]
	(0.007)	(0.007)	(0.008)	(0.015)	(0.014)	(0.002)
Observations	66	66	66	66	66	321
R-squared	0.307	0.391	0.335	0.331	0.332	0.860

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "mil" is military expenditure as a percentage of nominal GDP, "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 6: Netherlands, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1) 1948-2011	(2) 1948-2011	(3) 1948-2011	(4) 1948-2011	(5) 1816-2011
debt	0.078 ***	0.079 ***	-0.157 **	0.118 ***	0.031 ***
	[0.014]	[0.013]	[0.072]	[0.034]	[0.004]
	(0.013)	(0.013)	(0.055)	(0.033)	(0.003)
yvar	0.451 ***	0.400 ***	0.265 *	0.443 ***	0.173 *
	[0.162]	[0.172]	[0.145]	[0.160]	[0.103]
	(0.114)	(0.122)	(0.109)	(0.114)	(0.064)
gvar	0.046	0.041	-0.128 *	0.013	0.025
	[0.050]	[0.051]	[0.067]	[0.055]	[0.048]
	(0.038)	(0.038)	(0.052)	(0.045)	(0.020)
rir		-0.082 *			
		[0.061]			
		(0.069)			
debt_sq			0.158 ***		
			[0.049]		
			(0.037)		
debt_60				-0.019	
				[0.015]	
				(0.014)	
Constant	-0.030 ***	-0.029 ***	0.050 **	-0.048 ***	-0.011 **
	[0.008]	[0.008]	[0.024]	[0.017]	[0.005]
	(0.009)	(0.009)	(0.020)	(0.016)	(0.004)
Observations	64	64	64	64	188
R-squared	0.748	0.754	0.809	0.756	0.386

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 7: Belgium, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1) 1955-2011	(2) 1955-2011	(3) 1955-2011	(4) 1955-2011	(5) 1830-2011
debt	0.038 ***	0.036 ***	-0.232	0.103 *	0.025 ***
	[0.013]	[0.011]	[0.151]	[0.052]	[0.008]
	(0.013)	(0.012)	(0.103)	(0.039)	(0.006)
yvar	0.612 ***	0.670 ***	0.590 ***	0.563 **	0.322 ***
	[0.224]	[0.219]	[0.215]	[0.236]	[0.083]
	(0.212)	(0.211)	(0.201)	(0.210)	(0.083)
gvar	-0.101 ***	-0.099 ***	-0.091 ***	-0.103 ***	-0.030 **
	[0.036]	[0.031]	[0.033]	[0.038]	[0.012]
	(0.034)	(0.034)	(0.033)	(0.034)	(0.009)
rir		0.228 *			
		[0.119]			
		(0.135)			
debt_sq			0.153 *		
			[0.084]		
			(0.058)		
debt_60				-0.038	
				[0.030]	
				(0.022)	
Constant	-0.016	-0.025 ***	0.090	-0.046	-0.009 **
	[0.012]	[0.009]	[0.058]	[0.028]	[0.004]
	(0.012)	(0.012)	(0.042)	(0.021)	(0.004)
Observations	57	57	57	57	157
R-squared	0.250	0.289	0.338	0.291	0.261

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 8: Germany, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1)	(2)	(3)	(4)
	1970-2011	1970-2011	1970-2011	1970-2011
debt	0.026 *	0.012	0.205 ***	0.050 ***
	[0.013]	[0.010]	[0.045]	[0.014]
	(0.010)	(0.010)	(0.048)	(0.014)
yvar	0.227 **	0.345 ***	0.261 ***	0.233 **
	[0.107]	[0.095]	[0.086]	[0.105]
	(0.106)	(0.098)	(0.091)	(0.100)
gvar	-0.260 ***	-0.275 ***	-0.261 ***	-0.269 ***
	[0.047]	[0.051]	[0.053]	[0.051]
	(0.035)	(0.031)	(0.030)	(0.033)
rir		0.361 ***		
		[0.088]		
		(0.100)		
debt_sq			-0.196 ***	
			[0.046]	
			(0.051)	
debt_60				-0.022 **
				[0.010]
				(0.009)
Constant	-0.011 *	-0.018 ***	-0.045 ***	-0.018 ***
	[0.006]	[0.004]	[0.009]	[0.005]
	(0.005)	(0.005)	(0.010)	(0.006)
Observations	42	42	42	42
R-squared	0.611	0.713	0.722	0.664

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 9: Italy, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1)	(2)	(3)	(4)	(5)
	1948-2011	1948-2011	1948-2011	1948-2011	1861-2011
debt	0.073 ***	0.033 ***	-0.468 ***	-0.112 *	0.026
	[0.017]	[0.012]	[0.069]	[0.059]	[0.038]
	(0.013)	(0.011)	(0.058)	(0.047)	(0.025)
yvar	0.360	0.387 **	0.356 **	0.435 *	0.338
	[0.229]	[0.157]	[0.148]	[0.233]	[0.276]
	(0.210)	(0.156)	(0.135)	(0.188)	(0.151)
gvar	-0.039	-0.077 *	-0.056	-0.104 *	-0.133
	[0.071]	[0.042]	[0.043]	[0.061]	[0.104]
	(0.049)	(0.037)	(0.032)	(0.047)	(0.056)
rir		0.549 ***			
		[0.068]			
		(0.078)			
debt_sq			0.370 ***		
			[0.050]		
			(0.040)		
debt_60				0.125 ***	
				[0.039]	
				(0.031)	
Constant	-0.075 ***	-0.046 ***	-0.079 ***	-0.007	-0.045 *
	[0.015]	[0.010]	[0.020]	[0.025]	[0.026]
	(0.010)	(0.009)	(0.018)	(0.019)	(0.023)
Observations	64	64	64	64	150
R-squared	0.346	0.646	0.736	0.489	0.075

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 10: Spain, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1) 1946-2011	(2) 1946-2011	(3) 1946-2011	(4) 1946-2011	(5) 1850-2011
debt	0.005	-0.026	0.540 ***	0.053	0.025 ***
	[0.041]	[0.040]	[0.058]	[0.035]	[0.005]
	(0.024)	(0.022)	(0.066)	(0.023)	(0.005)
yvar	0.298 **	0.342 **	0.150 *	0.232 **	0.077
	[0.141]	[0.147]	[0.079]	[0.105]	[0.063]
	(0.127)	(0.111)	(0.089)	(0.110)	(0.051)
gvar	-0.039	-0.075 **	-0.101 ***	-0.058	-0.085 ***
	[0.041]	[0.037]	[0.030]	[0.035]	[0.029]
	(0.043)	(0.038)	(0.030)	(0.037)	(0.019)
rir		0.166 ***			
		[0.048]			
		(0.036)			
debt_sq			-0.820 ***		
			[0.104]		
			(0.098)		
debt_60				-0.120 ***	
				[0.025]	
				(0.025)	
Constant	0.002	0.011	-0.069 ***	-0.011	-0.005
	[0.012]	[0.011]	[0.007]	[0.009]	[0.004]
	(0.008)	(0.008)	(0.010)	(0.008)	(0.004)
Observations	66	66	66	64	156
R-squared	0.109	0.338	0.587	0.756	0.243

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 11: Portugal, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1)	(2)	(3)	(4)	(5)
	1945-2011	1945-2011	1945-2011	1945-2011	1852-2011
debt	-0.005	-0.012	0.092 *	0.032 **	0.018
	[0.019]	[0.018]	[0.053]	[0.014]	[0.014]
	(0.012)	(0.012)	(0.042)	(0.014)	(0.008)
yvar	0.053	0.011	0.038	0.052	0.075 **
	[0.087]	[0.078]	[0.082]	[0.082]	[0.030]
	(0.071)	(0.071)	(0.069)	(0.064)	(0.024)
gvar	-0.048	-0.050	-0.066	-0.060	-0.044 **
	[0.047]	[0.040]	[0.046]	[0.045]	[0.018]
	(0.029)	(0.028)	(0.029)	(0.027)	(0.014)
rir		0.100 *			
		[0.050]			
		(0.042)			
debt_sq			-0.106		
			[0.066]		
			(0.044)		
debt_60				-0.050 ***	
				[0.016]	
				(0.013)	
Constant	-0.007	-0.001	-0.023 ***	-0.016 ***	-0.008
	[0.007]	[0.005]	[0.009]	[0.006]	[0.006]
	(0.005)	(0.005)	(0.008)	(0.005)	(0.004)
Observations	67	67	67	67	160
R-squared	0.058 _	0.138	0.140	0.242	0.132

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 12: Iceland, fiscal reaction function estimations and robustness tests, dependent variable is the primary surplus to GDP ratio.

	(1) 1946-2011	(2) 1946-2011	(3) 1946-2011	(4) 1946-2011	(5) 1908-2011
debt	-0.020	-0.025	0.075 *	0.014	0.008
	[0.016]	[0.016]	[0.037]	[0.018]	[0.019]
	(0.012)	(0.013)	(0.035)	(0.016)	(0.014)
yvar	0.203 **	0.184 *	0.161 *	0.188 *	0.111
	[0.095]	[0.093]	[0.086]	[0.094]	[0.072]
	(0.068)	(0.070)	(0.066)	(0.065)	(0.049)
gvar	-0.093 ***	-0.091 ***	-0.081 ***	-0.083 ***	-0.059 **
	[0.030]	[0.029]	[0.028]	[0.029]	[0.026]
	(0.026)	(0.026)	(0.025)	(0.025)	(0.017)
rir		0.035			
		[0.034]			
		(0.031)			
debt_sq			-0.114 **		
			[0.045]		
			(0.039)		
debt_60				-0.050 ***	
				[0.018]	
				(0.017)	
Constant	0.017 ***	0.020 ***	0.005	0.010 **	-0.004
	[0.005]	[0.006]	[0.006]	[0.005]	[0.005]
	(0.004)	(0.005)	(0.006)	(0.005)	(0.004)
Observations	66	66	66	66	104
R-squared	0.227	0.243	0.319	0.326	0.114

Notes: Variable "yvar" is the gap between log real GDP and its trend and "gvar" is the gap between log of government expenditures and its trend (trends are extracted using an HP-filter with lambda=100), "rir" is the real interest rate, "debt_sq" is the debt-to-GDP ratio squared, "debt_60" is the debt-to-GDP ratio if it is larger than 60% and zero otherwise; [ ] = Heteroskedasticity- and autocorrelation-consistent standard errors (computed with Newey-West lag window of size 1); ( ) = ordinary standard errors. Significance levels: *** p<0.01, ** p<0.05, * p<0.1