Sapienza Università di Roma

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria

Francesca Di Iorio   Stefano Fachin

Fiscal reaction functions for the advanced economies revisited

DSS Empirical Economics and Econometrics Working Papers Series

DSS-E3 WP 2019/1

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria
“Sapienza” Università di Roma
P.le A. Moro 5 – 00185 Roma - Italia

http://www.dss.uniroma1.it

Fiscal reaction functions for the advanced economies revisited

Francesca Di Iorio[a]   Stefano Fachin[b]

Abstract

We revisit the evidence on the relationship between the Primary Balances/GDP and Debt/GDP ratios (Fiscal Reaction Function, FRF), in the advanced economies, showing that taking carefully into account the stochastic properties of the data leads to question the validity of the current consensus. More precisely, we find that before the 2008 financial crisis long-run FRF's existed only in a small number of advanced economies, and that they were more likely in countries characterized by higher sovereign spreads. Finally, we also find limited evidence of non-linearities leading to fiscal fatigue.

Keywords: Public Debt, Fiscal Reaction Function, Fiscal sustainability, Panel cointegration, Cointegrating Polynomial Regression, OECD, Spread.

JEL codes: C23, C32, E62, H62, H63.

([a]) Dip. di Scienze Politiche, Università di Napoli Federico II

([b]) Dept. of Statistics, “Sapienza” Università di Roma

1 Introduction[1]

The question if and how governments react to increases in public debt adjusting the primary surplus has attracted considerable attention since the publication of the seminal paper by Bohn (1998), in which the existence of a positive relationship between the Primary Balance/GDP and the Debt/GDP ratios was shown to be a sufficient condition for sustainability of the debt. Many empirical tests of debt sustainability have been carried out estimating this relathionship, known as Fiscal Reaction Function (FRF), either for individual economies or for panels of countries. The prevailing conclusion is that there is good support for the hypothesis of the existence of FRF's and thus sustainability (see, e.g., the recent and comprehensive review by Checherita-Westphal and Žďárek, 2017).

However, there are good reasons for claiming that a large part of this evidence is unreliable, so that the question of the existence of a debt-primary balances relationship is still essentially open. The problem is that while the early studies (Bohn, 1998, and before that, Trehan and Walsh, 1991) carefully took into account the stochastic properties of the data, this is often not true for the more recent contributions, especially those taking a panel approach. It is well known that non stationarity is a particularly critical point of any modeling exercise, and two of its implications are especially important for applied FRF analysis[2]. First, a stable long-run relationship between two variables may exist

[1]Research supported by "Sapienza" University of Rome, grant n. B82F16004935 001. Many thanks to A. Nobili for helping with the data collection and to M. Wagner for discussions and sharing computer programs. The usual disclaimers apply.

[2]Technical details on the following arguments can be found, e.g., in the textbook by Hamilton (1994), pp. 557-561.

if and only if they have the same stochastic structure: either both stationary or both non stationary. Regressing a stationary variable on a trending one is an attempt to estimate a relationship which cannot exist, and thus bound to produce seriously misleading results. Second, under non stationarity the usual asymptotic Gaussian significance tests are strongly oversized, and thus biased towards finding a significant relationship. In spite of the importance of the issue, Berti, Colesnic, Desponts, Pamies and Sail (2016) comment in their thorough review that "Surprisingly, [...], relatively few papers deal with stationarity problems [...]" (p. 9). In some cases this attitude may be explained by the admittedly stationary appearance of both the debt and primary balances series over the very long run, which makes formal tests appear as unnecessary complications[3]. However, in most circumstances the assumption of stationarity is instead questionable, especially for debt over shorter periods such as the second half of the 20th century (see, e.g., Bohn, 1998: "Empirically, it is difficult to reject a unit root in real debt and in the debt-GDP ratio").

Given this background, our aim is to reach reliable conclusions on the FRF for the advanced economies. To this end we will first of all carry out a careful study the time series properties of the Primary Balances/GDP and Debt/GDP series, in order to establish in which of these countries they are such that FRF's may exist. In these cases we shall then estimate country-specific FRF's using adequate techniques, which for the non stationary case differ for linear and non-linear specifications, a point totally ignored in the literature. Following this approach we shall find that, contrary to the commonly accepted

[3]For instance, this is certainly the case for the data spanning from the mid-19th century to the late 2000's plotted in Fig. 1 in Mauro *et al.* (2015).

conclusions, before the 2008 financial crisis long-run FRF's existed only in a small number of advanced economies. Rather unexpectedly, we shall see that these tend to be countries with a poor fiscal reputation.

The paper is organised as follows: in Section 2 we define the set-up, in Section 3 we describe that data and carry out an univariate preliminary analysis, in Section 4 estimate the long-run FRF's, and finally in Section 5 draw some conclusions.

2 Fiscal reaction functions

2.1 Set-up

The FRF literature originates essentially from Trehan and Walsh (1991) and Bohn (1998), and is summarized, for instance, in Bohn (2005). The central relationship is the intertemporal budget constraint (IBC), which states that at any moment in time debt at the start of the current period, d_t^*, must be backed by the expectation of the present value of all future primary surpluses (s):

$$d_t^* = \sum_{i=0}^{\infty} E_t(u_{t,i} s_{t+i}) \tag{1}$$

where E_t is the conditional expectation operator and u a pricing kernel. Bohn (1998) showed that a sufficient condition for the IBC to hold is that the ratio between Primary Balances (PB) and GDP (hereafter pb) is an increasing function of the lagged ratio between Debt (D) and GDP (hereafter d) and a bounded residual. The simplest case is the linear FRF:

$$pb_t = \rho d_{t-1} + \mu_t \tag{2}$$

where $\rho > 0$ and the term μ_t is bounded as a share of GDP (whose present value is assumed to be finite). Given the typical dimensions of d and pb the FRF coefficient ρ is expected to be quite small, most likely below 0.10.

Model-based sustainability analysis can be carried out estimating either directly equation (2) or an augmented version including a set of stationary variables Z capturing cyclical conditions, for instance the output gap[4]. In either case, it is important to keep in mind that the existence of a stable and significant relationship between pb and d is a sufficient, but by no means necessary, condition for sustainability. As remarked by Bohn (2005), the point is the degree of confidence markets have that a country will actually implement all future policies necessary to satisfy the budget constraint. Thus, empirical violations of the sustainable fiscal policy rule defined by (2) are possible if markets expect future policy shifts ensuring that the IBC is nevertheless respected.

An important development in the literature is due to Gosh *et al.* (2013), who pointed out that the FRF should be generalized to account for the increasing difficulty governments may find to increase primary balances as debt grows, or "fiscal fatigue". In practice, they estimated a homogeneous panel model using a cubic specification of the type:

[4] In this context an important, but generally overlooked point, concerns the different implications of using revised or real time data, see Golinelli, Mammi and Musolesi (2018).

$$pb_t = \theta + \rho_1 d_{t-1} + \rho_2 d_{t-1}^2 + \rho_3 d_{t-1}^3 + \beta Z_t + e_t \tag{3}$$

finding significant non-linear effects of the expected shape for their panel of 23 advanced economies for the period 1970-2007. Interestingly, Bohn (1998) had already investigated the issue for a single country and a much longer period (USA, 1916-1995) obtaining opposite results (effect increasing with debt). In fact, Everaert and Jansen (2018) recently showed that non-linear effects in homogeneous panel models may be simply the consequence of heterogeneity ignored by the homogeneous panel models.

2.2 Conditions for the existence of the FRF

A careful scrutiny of the possibility to use the FRF in its various forms as a tool for empirical model-based sustainability analysis can be based on the concept of "balanced equation". From Ermini and Granger (1993): "An equation is balanced if the dominant property of the means of the dependent and explanatory variables are identical" (p. 370). For the linear long-run FRF (2) this means that pb and d must have (i) the same order of integration, and, (ii) the same deterministic kernel, that is, depend on linear trends of the same order. If one of these two conditions is not satisfied the residual μ of equation (2) will be unbounded, and the sustainability condition will not be satisfied either. Empirically, it is thus crucial that the balancing conditions (i) and (ii) are tested before estimation is carried out, either directly through unit root tests or indirectly through cointegration tests.

These conditions can be extended to the non-linear specification (3) using a result again by Ermini and Granger (1993): although a power of an integrated variable is *not* difference stationary[5], it can be shown that a polynomial transformation like (3) will have a polynomial time trend and autocorrelations that decay slowly, similar to those of an $I(1)$ process. The trend is present even when the variable has no drift and its order depends on that of the polynomial. The implication is that with non-stationary d the right-hand side of (3) will behave like a $I(1)$ series, and estimation of the FRF will make sense only if the variable on the left-hand side, pb, also does, exactly as it happens for the basic linear specification. There are instead no a priori constraints on the deterministic kernels, as the precise form of the trend in the right-hand side will depend on that of the polynomial transformation.

To recapitulate:

(i) estimation of *linear* FRF's makes sense only when the Primary Balances/GDP and Debt/GDP ratios have the same order of integration and the same deterministic structure;

(ii) estimation of *non-linear* (quadratic or cubic) FRF's requires only the same order of integration;

(iii) finally, under non stationarity standard asymptotic inference does not apply; suitable estimators and inference must be used (as we shall see, these will be different for linear and non-linear models).

[5] The simple case of the square of a pure random walk is illustrated in Example 2 of Granger (1995), p. 270.

In spite of the fact that these points mostly follow from well-established results, none of them has received proper attention in recent studies. Even when mentioned, the issue is often treated in a rather cavalier way. For instance, Plödt and Reicher (2015), after commenting that d is likely to be $I(1)$ while the integration order of pb is not clear, simply proceed to estimate by OLS e 2SLS regressions both in levels and first differences; the fact that in case of non-stationarity standard asymptotic inference would not apply is not considered. In Checherita-Westphal and Žďárek (2017) non-stationarity is described in a footnote as a "potential problem" not further investigated. Even the few exceptions, as for instance Berti *et al.* (2016) and Lamé, Lequien and Pionnier (2014), fail to extend their assessment of the properties of the two variables to the deterministic part of the DGPs. The problem is particularly serious for multi-country studies. While the stationarity assumption may not be particularly restrictive for panels with very short time samples, this is typically not the case for FRF panel studies, which use datasets with at least 30 or 40 time observations (Mendoza and Ostry, 2008, Fournier and Fall, 2017, Everaert and Jansen, 2018) and in some cases up to 60 (D'Erasmo, Mendoza, Zhang, 2016). In these circumstances, if pb and d are not stationary usual asymptotic inference applied to panel estimators will be unreliable and misleading, exactly as it happens with pure time series models[6].

The situation is even worse if the stochastic and deterministic structure of pb and d are such that the FRF is not balanced. Although the estimates

[6]See Entorf (1997) and Banerjee, Eberhardt and Reade (2010). Considering that panel estimators suitable for non-stationary data have been made available more than a decade ago (Mark and Sul, 2003, Pedroni, 2001), the use of standard methods suitable only for stationary data is actually a choice difficult to understand.

of the ρ coefficient will asymptotically converge to zero, in moderate samples they will be small, but greater than zero: precisely the dimension expected on a priori grounds. The estimated coefficients from these unbalanced models will thus misleadingly conform to a priori expectations. Standard significance tests as those typically reported in the literature will not help either, as the usual asymptotic inference will not apply. The implication is that quite a substantial part of the evidence allegedly found in the literature in support of the existence of FRF's might be entirely spurious. In particular, this casts some serious doubts on the panel models estimated pooling together "FRF" and "Non-FRF" countries.

On the basis of this picture, our aim can be more precisely stated as to re-examine the evidence on the existence of FRF's in the 22 advanced economies present in the dataset assembled by Mauro, Romeu, Binder and Zaman (2015), which are essentially the countries analyzed also by Gosh *et al.* (2013) and Mendoza and Ostry (2008). More precisely, we will first of all carry out a careful analysis of the time series properties of the data and their implications for the estimation of long-run linear and non linear FRF's, then, when possible, proceed to estimate them using adequate methods. We shall thus hopefully reach reliable and robust conclusions on the debt-primary balances relationship.

3 Primary Balances/GDP and Debt/GDP paths in the advanced economies

3.1 Data

Our empirical study will be based on the dataset assembled by Mauro *et al.* (2015) updated using IMF's "World Economic Outlook Database" (WEO). The starting year is 1961, the earliest available for all advanced countries in the Mauro *et al.* (2015) data, while the final year is 2013, the latest available in WEO. Our panel will include the G7 (Canada, France, Germany, Italy, Japan, UK and USA), all the other western and southern European economies[7] (Austria, Belgium, Denmark, Finland, Greece, Iceland, Ireland, Netherlands, Norway, Portugal, Spain, Sweden and Switzerland), Australia and New Zealand. The time series plots of mean, median, first and third quartiles of the distribution of the Primary Balances/GDP and Debt/GDP ratios across countries are in Fig. 1, country plots in Figs 2-4, and some averages in Table 1.

A first comment on the summary plots in Fig. 1 is that the distributions of both variables are rather narrow, sign of strong comovements across countries. The existence of significant cross-country dependence is confirmed by the CD test of Pesaran (2015). The CD test statistics for pb and d are respectively 19.7 and 21.7, which both reject with p-values close to zero the null hypothesis of weak cross-sectional dependence in favor of the alternative hypothesis of strong dependence.

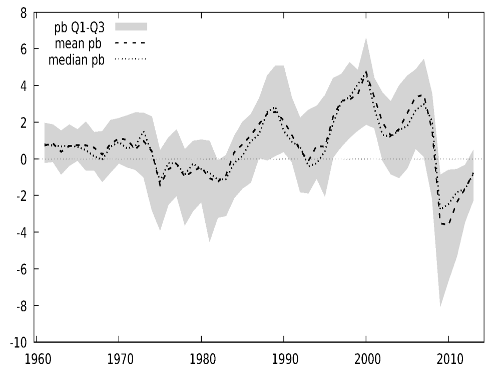
The second, and crucial, remark suggested by both the summary and the

[7] Except the very small ones, Luxembourg, Malta and Cyprus.

country plots is that until the early 2000's most Debt/GDP series appeared to have an upward trend not shared by Primary balances/GDP ratios. The following years, the aftermath of the 2008 crisis, are obviously extremely unsettled. Between 2008 and 2010 the combination of increases in expenditure and decline in revenues produced generalized, drastic falls in primary balances and significant growth of debt relatively to GDP (in mean, respectively from 1.5% to about -3.5% and from 65% to 80%), while after 2010 the two variables appear to move somehow in the same direction, suggesting that fiscal consolidation efforts were generally carried out[8]. This is consistent with the finding by D'Erasmo, Mendoza and Zhang (2016) of breaks in the primary balances-debt relationship both for the USA and a large multicountry panel. Considering that the period after the crisis is relatively short and extremely troubled, rather than using the full sample and testing and estimation methods robust to breaks we preferred to follow Gosh *et al.* (2013) and truncate the study period at 2007. Thus, our exercise can be described as an assessment of model-based of debt sustainability analysis in the advanced economies at the eve of the 2008 financial crisis. Country-level statistics (minimum, median and maximum) for this period are reported in Table A.1 in the Appendix. The average of the median ratios (that is, the average across countries of the median d for each country over the sample period) is 46%, close to the grand average across countries and years reported in Table 1 above, 51.5%. The average of its maximums (that is, the average of the maximum of each country over time) is 86%.

[8] In some cases (Greece, Ireland and Portugal) these have been explicitly required to access extraordinary financing by EU and IMF (see IMF-IEO, 2016).

Figure 1: Mean, median, 1st and 3rd quartiles of Primary Balances/GDP (pb, left panel) and Debt/GDP (d, right panel) in 22 advances economies, 1961-2013. Countries: *Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, USA.*



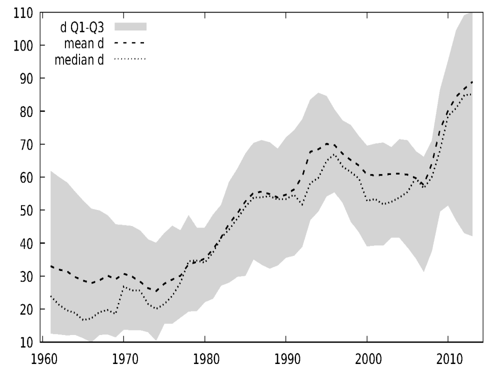


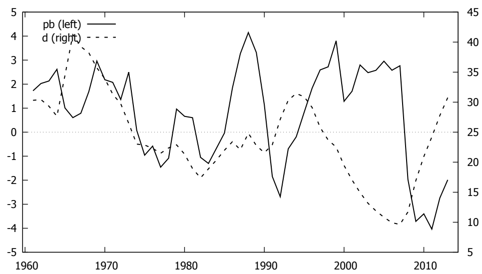
Table 1: Debt and Primary Balances, averages 1961-2013

	1961–74	1975-96	1997-2007	2008-13	1961-2007	*1961-2013*
pb	0.7	0.4	2.9	-1.7	1.1	*0.8*
d	31.7	50.8	62.7	80.1	47.9	*51.5*

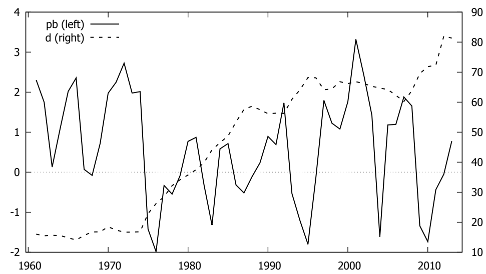
Values ×100. *Countries*: see legend of Fig. 1.

Figure 2: Debt/GDP (d) and Primary Balances/GDP (pb), 1961-2013.

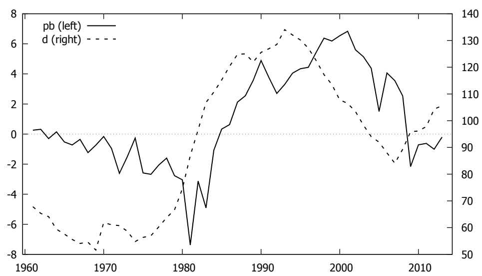
Australia



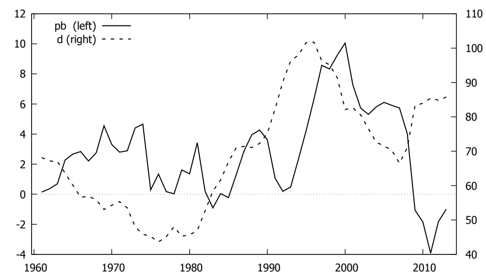
Austria



Belgium

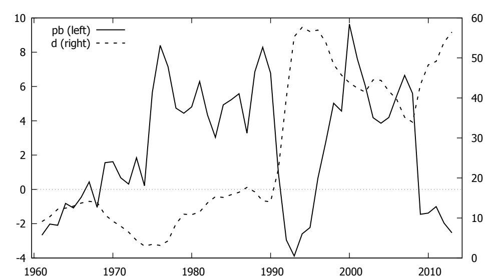


Canada

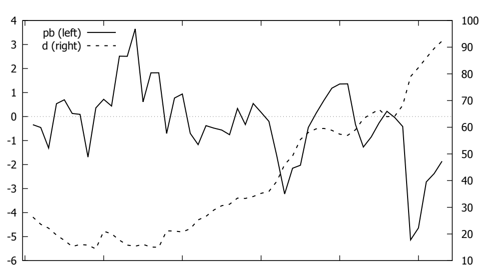


Denmark

Finland



France



Germany

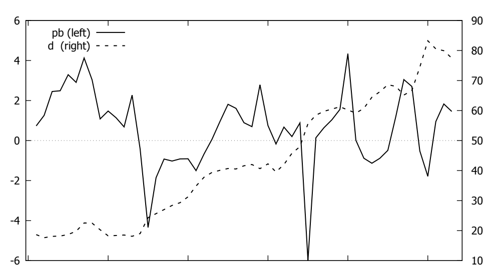


Figure 3: Debt/GDP (d) and Primary Balances/GDP (pb), 1961-2013.

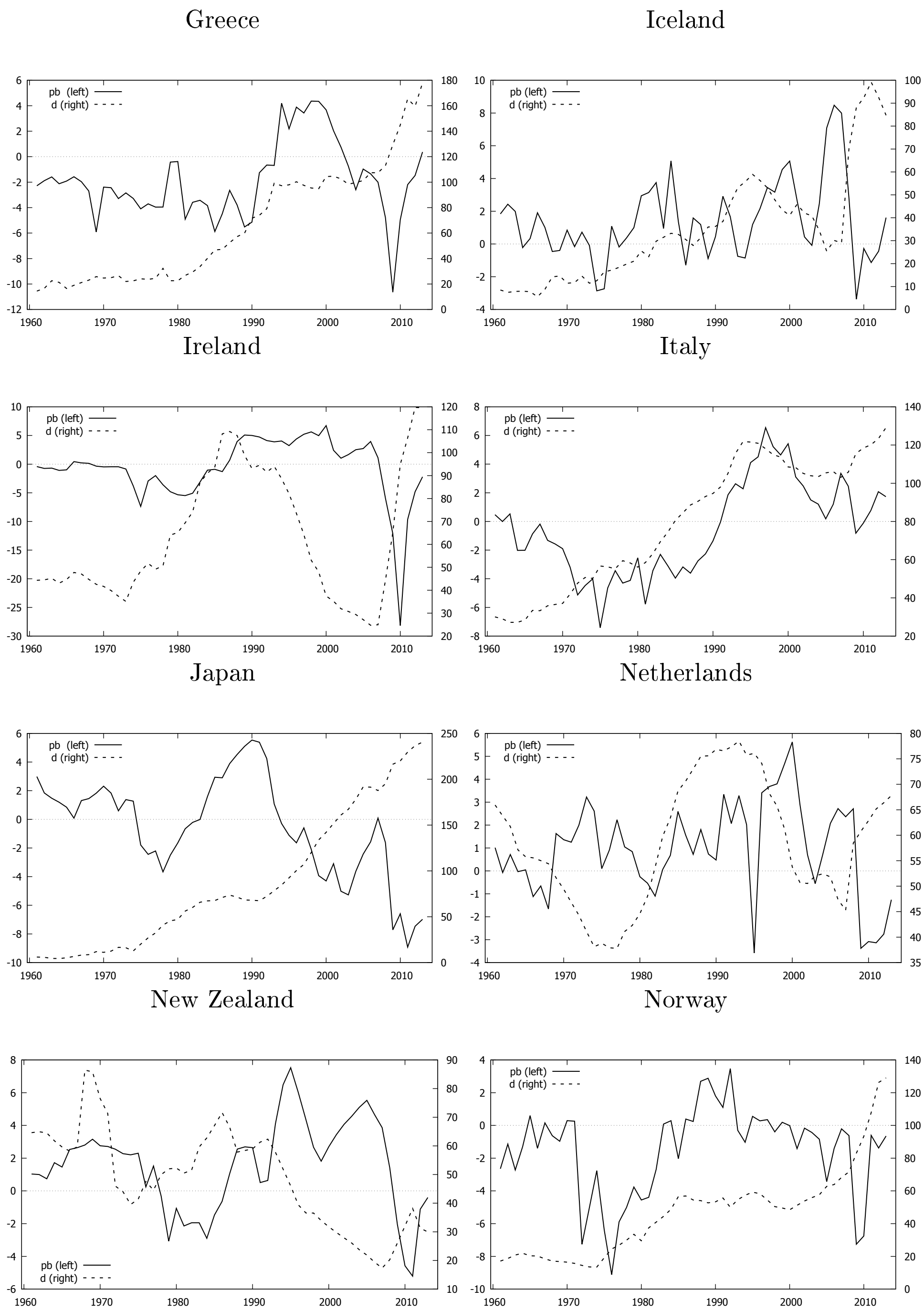
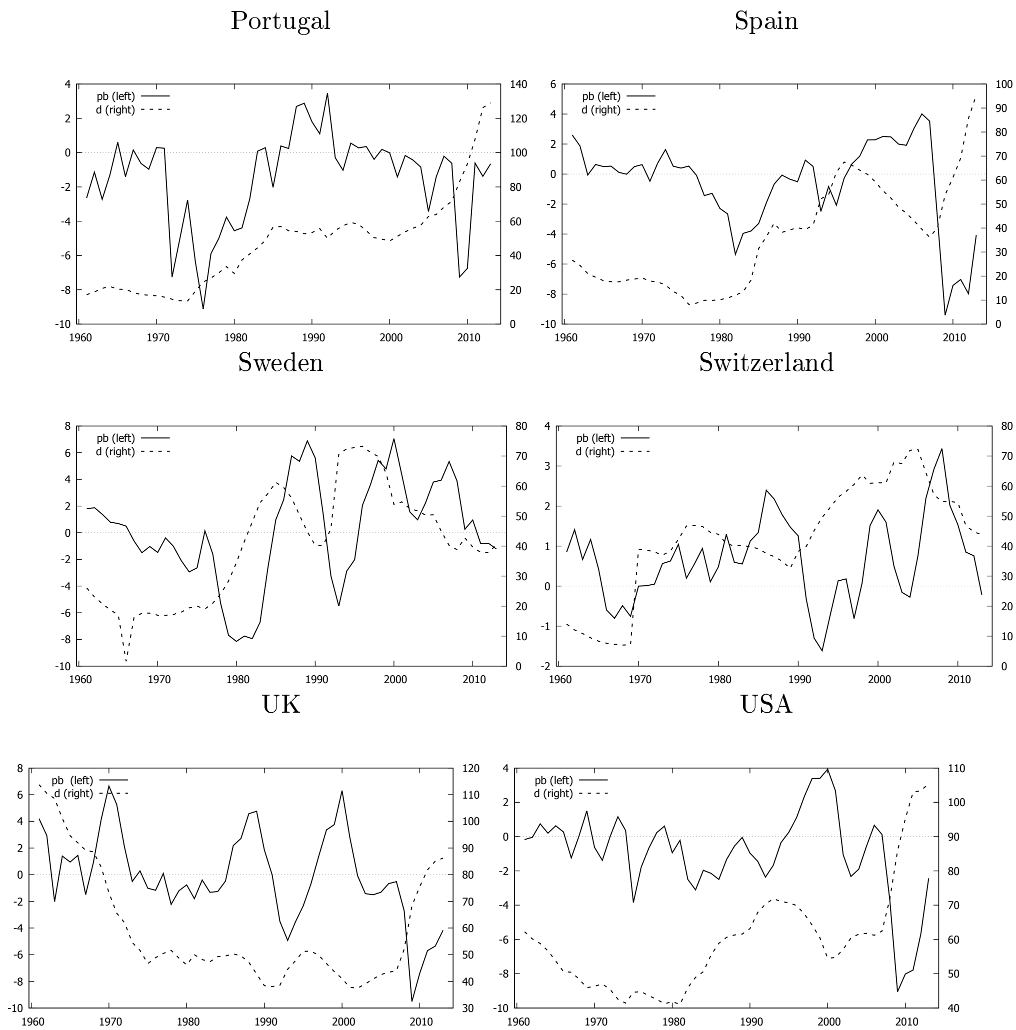


Figure 4: Debt/GDP (d) and Primary Balances/GDP (pb), 1961-2013.



3.2 Univariate preliminary analysis

As anticipate above, the first step of our study will be to assess the stochastic (order of integration) and deterministic (linear trends) time series properties of d and pb for the period 1961-2007 through a careful univariate analysis. A summary of the conclusions reached on the basis the customary ADF tests is in Table 2, with all details in Tables A.2-A.6 in the Appendix. These results, broadly consistent with those reported by Berti *et al.* (2016) for 1950-2008, have been confirmed by two robustness checks, KPSS tests for the null hypothesis of stationarity (Table A.3 in the Appendix) and more powerful ADF-GLS tests for the null hypothesis of non stationarity (not reported but available on request).

Table 2: Summary of conclusions of ADF unit root tests,1961-2007

Australia	***Austria***	*Belgium*	*Canada*	***Denmark***	*Finland*
0,1	0,1	1,1	1,1	0,1	1,1
France	***Germany***	*Greece*	*Iceland*	*Ireland*	*Italy*
0,1	0,1	$1{,}1^a$	1,1	1,1	1,1
Japan	***NL***	*NZ*	*Norway*	*Portugal*	***Sweden***
0,1	0,1	$1{,}1^b$	1,1	$1{,}1^a$	0,1
Spain	***CH***	*UK*	***USA***		
1,1	0,1	0,0	0,1		

For each country p,q is respectively the order of integration of pb and d;
^{a}: with a negative drift; ^{b}: with a positive drift;
Bold case: countries where $p \neq q$ and the FRF is not balanced.

A first, striking conclusion is that in ten countries out of 22 (USA, Germany, Japan, France, Australia, Austria, Denmark, Netherlands, Sweden and Switzerland) pb is stationary while d is not, so that neither linear nor non-linear FRF's are possible. Note that this groups includes four of the G7 economies.

In one country, UK, pb and d are both stationary, so that the FRF is definitely balanced. In the remaining 11 countries we have two groups. In eight countries (Belgium, Canada, Finland, Iceland, Ireland, Italy, Norway, Spain) both ratios are non stationary and the deterministic structure compatible, as neither pb nor d have a drift. The FRF is then clearly balanced. In the remaining three countries d has a drift, negative in New Zealand and positive in Greece and Portugal. In these three countries a linear FRF may be balanced only if we regress pb (which has no drift) on the deviations of d from a linear trend. The interesting implication is that in these countries policy makers seem to have acted as if these linear trends were ineluctable; for Greece and Portugal this is entirely consistent with the 2010 sovereign debt crisis, for New Zealand it suggests possible systematic overaccumulation of assetts. We now proceed to estimate the long-run FRF's for these 11 countries.

4 Long-Run Fiscal Reaction Functions

4.1 Non-linear FRF's

4.1.1 UK: stationary Primary balances/GDP and Debt/GDP

Under stationarity testing for a primary balances-debt relationship is a simple task which can be carried out using standard tools. We chose the bounds test by Pesaran, Shin and Smith (2001), constructed as a joint test of significance for the level variables in an ARDL. Starting from the cubic specification:

$$\Delta pb_t = \theta + \sum_{i=1}^{3} \alpha_i \Delta pb_{t-i} + \sum_{i=1}^{3} \beta_i \Delta d_{t-i-1} + \gamma pb_{t-1} + \sum_{i=1}^{3} \delta_i d_{t-2}^i \qquad (4)$$

we obtained the preferred specification reported in Table 3, with model selection tests and diagnostics in Table A.7 in the Appendix. The restricted model appears to be well specified and clearly supporting a level relationship with linear effects only. The long-run coefficient $\rho = \delta_1/\gamma$ is estimated equal to 0.24, a rather high value which should be interpreted keeping in mind the particular trajectory followed by d in this country (see Fig. 4). Starting from the extremely high levels of the early '60's, legacy of the massive WWII expenditure[9], by the mid-70's in this country public debt had declined to less than 50% of GDP. Until the 2008 financial crisis d fluctuated around that level, with very small increases of debt tackled by large offsetting movements of the Primary balance; this is particularly evident in the early 1990's.

Table 3: A dynamic FRF for UK, 1961-2007 (p-values$\times$100 in brackets.)

	β_1	γ	δ_1	θ
coeff [s.e.]	0.279 [0.084]	-0.180 [0.099]	0.043 [0.013]	-1.86 [0.72]
t [$p \times 100$]	3.33 [0.2]	-1.82 [7.6]	3.36 [0.2]	-2.58 [1.4]

Model: $\Delta pb_t = \theta + \beta_1 \Delta d_{t-2} + \gamma pb_{t-1} + \delta_1 d_{t-2}$
H_0 : *"No non-linear effects"* $t(\delta_3 = 0) = 0.33$ [$p = 74.7$]; $t(\delta_2 = 0) = -1.29$ [$p = 20.4$];
H_0 : *"No level relationship"* $F(\gamma = \delta_1 = 0) = 7.04$ (1% critical value: 6.84)

[9]In 1947 UK debt was 238% of GDP (source: Mitchell, 2011).

4.1.2 Countries with non stationary Primary balances/GDP and Debt/GDP

With non stationary variables non-linear modeling is much more challenging. The problem is that, as discussed above, powers of integrated variables are not difference stationary, for any order of differencing. This implies that the asymptotic results for the usual cointegration tests and estimators for cointegrating regressions, based on the assumption of difference stationary variables, cannot be used for polynomial non stationary models. We need a completely different set of econometric tools, recently developed by Wagner and coauthors in a series of contributions (Wagner 2015, Wagner and Hong, 2016, and the references therein). The starting point is the concept of a stable polynomial relationship, or *Cointegrating Polynomial Regression* (CPR), that is, an equation including powers of non stationary variables and with stationary residuals[10]. Existence of a CPR may be tested using two direct, variance ratio, tests, labeled as *Shin* and P_u, and a LM specification test. We shall refer to these three tests as "CPR tests". The P_u test has as a null hypothesis "no cointegration", just like the Engle-Granger test, while the Shin and LM tests have the null hypothesis of cointegration; more details are provided in the Appendix. The availability of tests with opposite null hypothesis will allow us to reach rather reliable conclusions: to conclude in favour of the existence of a stable polynomial relationship we will require no rejections from the Shin and LM tests to be accompanied by a rejection from the P_u test.

[10]Note that the concept of cointegrating polynomial regression is totally unrelated to that of "polynomial cointegration" in the sense of a cointegrating relationship involving $I(2)$ variables and their first differences.

In our case the tests, summarized in Table 4 while all details are reported in Table A.8 in the Appendix, clearly point to the existence of a Non-linear equilibrium relationship between pb and d only in two cases, Greece and Italy. In all the other countries the tests appear to suffer from low power: the null hypothesis, be it cointegration (Shin and LM tests) or, on the opposite, no cointegration (P_u test), is never rejected.

For Greece and Italy we then proceeded to estimate the polynomial relationship using a suitably modified least squares estimator similar to FM-OLS developed in Wagner and Hong (2016), which we label FM-CPR. The estimates are reported in Table 5. Although cubic effects are significant in both cases, the coefficients are such that the shape of the function is deeply different in the two countries. This is evident from Fig. 5, where the FRF's for the two countries are plotted for the empirical 1961-2007 ranges. In Greece the function has a sinusoidal shape, initially concave with a minimum at about d=40%, a maximum at about d=85% (a level already abundantly exceeded in 1993, when the ratio, with a sharp acceleration from the 79% value of 1992, reached 99%), and a final section steeply negatively inclined. This shape is consistent with the presence of "fiscal fatigue", and is in fact remarkably similar to the curve fitted by Gosh *et al.* (2013) to the scatter plot of means and medians of pb (Gosh *et al.* 2013, Fig. 3). The situation is quite different in Italy, where the FRF is concave with a minimum at d=60% and practically linear for d greater then 80% (a level reached in the mid-80's), thus showing no signs of "fiscal fatigue".

Summing up, the results of our non-linear analysis are the following. First, non-linear effects appear to be clearly present in two countries, Greece and Italy, but only in one of these, Greece, they support the "fiscal fatigue" hypothesis. This heterogeneity is a further warning on the risks of estimating homogeneous panel models. Second, in all the other countries (Belgium, Canada, Finland, Iceland, Ireland, New Zealand, Norway, Portugal, Spain) no conclusion can be reached. Although to a certain extent this is certainly a consequence of lack of power of the available CPR tests, it also suggests that even allowing for a considerably flexible functional form the links between the two variables are not very strict. For these cases we shall attempt estimation of the simpler linear specification.

Table 4: Cointegrating Polynomial Regression tests, 1961-2007

	Belgium	*Canada*	*Finland*	*Greece*
P_u	No	No	No	**Yes**
Shin	Yes	Yes	Yes	**Yes**
LM	No	No	Yes	**Yes**
	Iceland	*Ireland*	*Italy*	*New Zealand*
P_u	No	No	**Yes**	No
Shin	Yes	Yes	**Yes**	Yes
LM	Yes	Yes	**Yes**	Yes
	Norway	*Portugal*	*Spain*	
P_u	No	No	No	
Shin	Yes	Yes	Yes	
LM	Yes	Yes	Yes	

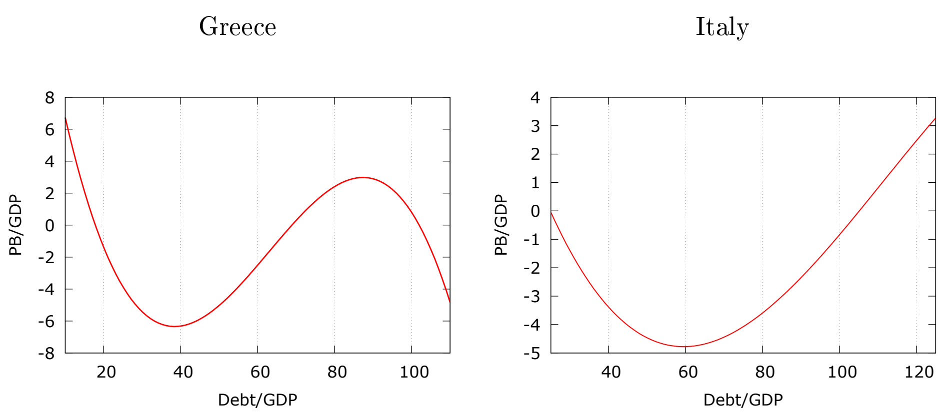
"Yes", the model $pb_t = \theta + \rho_1 d_{t-1} + \rho_2 d_{t-1}^2 + \rho_3 d_{t-1}^3 + e_t$ is a Cointegrating Polynomial Regression; "No", it is not.

Table 5: Long-run non-linear FRF's, FM-CPR estimates 1961-2007

	Greece	*Italy*
ρ_1[s.e]	-1.601 [0.337]	-0.610 [0.169]
ρ_2[s.e]	0.030 [0.006]	0.007 [0.002]
ρ_3[s.e]	-1.6E-04 [3.0E-05]	-2.1E-0.5 [1.1E-0.5]
constant	19.90 [5.03]	11.16 [3.50]

Model: $pb_t = \theta + \rho_1 d_{t-1} + \rho_2 d_{t-1}^2 + \rho_3 d_{t-1}^3 + e_t$
NB S.e.'s in brackets; all coefficients significant at 5%.

Figure 5: Estimated non-linear FRF's



4.2 Linear FRF's

The non-linear analysis described in the previous section was inconclusive in nine cases: Belgium, Canada, Finland, Iceland, Ireland, New Zealand, Norway, Portugal, and Spain. Accordingly, we proceed to a linear study, first testing for linear cointegration, and, if this is found to hold, estimating the linear FRF's.

On the basis of the univariate analysis, for Portugal and New Zealand d will be taken as deviations of from the respective linear deterministic trends.

Looking at the plots for our panel of countries (Figs. 2-4) as a whole the impression of some relationship is clear. Although in most countries pb and d obviously drifted away from each other in the 1960's and 1970's, it is also obvious that they did move together in the following decades, when in fact we know that debt adjustment plans were implemented in many economies at the same time (see, e.g., Mauro *et al.*, 2015, p. 56). However, standard individual Engle-Granger tests, reported in section A of Table 6, find only some rather weak evidence in favor of the existence of a long-run link between the debt and primary balance GDP ratios. Given the general impression of the existence of a relationship, we expect a panel test exploiting the joint evidence to be able to grant some power gains with respect to the individual tests. A crucial point is that, since our aim is to estimate country-specific FRF's, we need to test the null hypothesis "pb and d are cointegrated in *no* country" against the alternative hypothesis "pb and d are cointegrated in *all* countries". The popular panel cointegration tests based on the average of the individual statistics (for instance, the "Group mean" test by Pedroni, 1999) are not suitable: by the very definition of average, panels including only few but strongly cointegrating countries may yield significant tests of this type even with a non-negligible fraction of not cointegrating ones[11]. To ensure rejection of no cointegration if, and only if, *all* relations in the panel are cointegrating we need to summa-

[11] Hence, using such a procedure as pre-testing before the estimation of individual FRF's may lead to estimate spurious regressions, and things will be even worse for panel estimation, as in this case non-cointegrating and cointegrating countries will be pooled together.

rize the test statistics computed for all the countries in the panel with the one most favorable to the null of no cointegration. For instance, for Engle-Granger type tests the maximum of the individual statistics. Further, we need a test robust to the strong links between our countries. The "first generation" panel cointegration tests like Pedroni's "Group mean" tests mentioned above and applied for instance by Betty and Shiamptanis (2013) assume independence across units (here countries), a condition which was seen not to hold here (see the CD tests in Section 3.1), and are thus not valid.

The bootstrap test *Max(HEG)* by Di Iorio and Fachin (2014) appears instead fully suitable for our needs: it tests the null of no cointegration against the alternative that all units are cointegrated and it is robust to short- and long-run dependence across the units of the panel. Some more details are provided in the Appendix. Consistently with our expectations, the result of this panel cointegration test (Table 6, section B) is favorable to the hypothesis of a long-run link between d and pb than the individual tests: the p-value, 3.9%, is such that the null hypothesis of no cointegration can be rather safely rejected, and we can thus proceed to estimate separate FRF's for each country.

The estimates of the FRF coefficients, obtained by FM-OLS and reported in Table 7, are of the right sign in all cases but Finland and New Zealand, where they are negative and non significant[12]. We thus proceed to remove these two countries from the panel and recompute the *Max(HEG)* test for the panel including Belgium, Canada, Iceland, Ireland, Norway, Portugal and Spain. The p-value of the test increases marginally, to 6.2%. Given that the

[12]In both cases the KPSS test did not reject stationarity of the primary balances (see Table A.3), so that this is not surprising

cross-section sample size is at this point quite small (seven countries) this can still be considered quite favorable to the hypothesis that in these countries there is a stable long-run FRF. In fact, in five of these countries the FM-OLS estimates of the long-run FRF coefficient ρ are significant and in line with the literature[13]. Our FM-OLS estimates fall in the interval [0.04, 0.09], while Mauro *et al.* (2015) report for the same set of countries and a slightly longer period, 1950-2007, OLS estimates in the interval [0.05-0.08][14].

Table 6: Cointegration tests, 1961-2007

A. Engle-Granger Cointegration test, p-values ×100				
Belgium	*Canada*	*Finland*	*Iceland*	*Ireland*
80.6	13.0	20.0	9.5	75.6
New Zealand	*Norway*	*Portugal*	*Spain*	
44.1	8.9	50.4	63.7	
B. Panel cointegration *Max(HEG)* test, p-values ×100				
All countries (N=9)				3.9
All countries except Finland and New Zealand (N=7)				6.2

Model: $pb_t = \theta + \rho d_{t-1}$. *Portugal*: detrended d.
Max(HEG): bootstrap p-values, 5000 redrawings.

[13] Recall that the FRF coefficient for Portugal links pb to the detrended d, and is thus not comparable with those found in the literature.

[14] These OLS estimates are consistent under non-stationarity (but, however, not asymptotically Gaussian, so that the standard significance tests accompanying the estimates are not valid).

Table 7: Long-run linear FRF's, FM-OLS estimates 1961-2007

	Belgium	*Canada*	*Finland*
ρ[s.e]	**0.086** [0.007]	**0.092** [0.013]	-0.017 [0.042]
t [p ×100]	13.05 [0.0]	7.21 [0.0]	-0.3949 [69.3]
	Iceland	*Ireland*	*New Zealand*
ρ [s.e]	0.048 [0.042]	**0.038** [0.017]	-0.039 [0.078]
t [p ×100]	1.12 [26.1]	2.22 [2.7]	-0.45 [61.8]
	Norway	*Portugal*	*Spain*
ρ [s.e]	0.265 [0.205]	**0.227** [0.055]	**0.039** [0.015]
t [p ×100]	1.23 [19.5]	4.146 [0.0]	2.65 [0.8]

Model: $pb_t = \theta + \rho d_{t-1}$. Coefficients significant at 5% in bold case.
Portugal: detrended d.

4.3 Who and why needed a FRF?

Looking at the results of our testing and estimation exercise, summarized in Table 8, two questions naturally arise: first, what have in common the countries for which a FRF could be estimated ("FRF countries", in bold face)? Second, what makes them different from "Non-FRF" ones? In short, who, and why, needed a FRF? Considering FRF countries first, from the country summary statistics for the Debt/GDP ratio in Table A.1 and, visually, from the plots in Figs. 2-4 we can identify, with the only exception of UK, two possible common features. First, in Belgium, Canada, Ireland, Greece and Italy medians and maximums of d are much higher than their grand average over all countries and years (respectively, 70% against 48% for the median and

Table 8: FRF's: summary of results, 1961-2007

Australia	*Austria*	***Belgium***	***Canada***	*Denmark*	*Finland*
No	No	Yes/L	Yes/L	No	?
France	*Germany*	***Greece***	*Iceland*	***Ireland***	***Italy***
No	No	Yes/NL	?	Yes/L	Yes/NL
Japan	*Netherlands*	*New Zealand*	*Norway*	***Portugal***	***Spain***
No	No	?	?	Yes*/L	Yes/L
Sweden	*Switzerland*	***UK***	*USA*		
No	No	Yes/NL	No		

Bold face: countries for which the FRF could be estimated;
No: FRF not possible (Debt/GDP is $I(1)$, Primary balances/GDP is $I(0)$);
Yes/NL or */L*: FRF is a Non-linear (NL) or Linear (L) cointegrating relationship with significant coefficient(s);
*Yes**: FRF cointegrating with detrended Debt/GDP;
?: FRF possible but cointegration status not clear and/or coefficient not significant.

100% against 86% for the maximum). In the two other countries, Portugal and Spain, d was not globally as high as in the first group (medians between 30% and 50%, maximums around 60-70%), but it grew regularly during all the period or rapidly in part of it. The maximum value is about five times the minimum one in Portugal and eight in Spain, with the lowest values reached in the mid-1970's and the highest ones respectively in 2007 and 1996. The corresponding ratio of the averages across countries (average minimum/average maximum) is about 4.5. Thus, the answer to our first question seems to be rather clear: with the only expection of UK, in FRF countries d was either high on the average or somehow upward trending[15].

If we turn to Non-FRF countries we instead have some surprises. While in some of them (e.g., Australia and Finland) Debt/GDP ratios are definitely low on the average or negatively trending, and thus fundamentals are clearly different from those of the FRF group, in other countries they are not. In Japan the Debt/GDP ratio reached by far the highest level of all countries in our panel (192%, in 2006) and was the third highest in median (67%; that is, for half of the period the ratio was higher than 67%). In France and Germany Debt reached in the mid-2000's, after decades of almost uninterrupted growth, values close to 70% of GDP, for instance as high as in Portugal and much higher than in Spain. Thus, while it is true that all countries which have a FRF have a Debt/GDP ratio either high or trending, the inverse is not: not all the countries with high or trending Debt/GDP have a FRF.

What is, then, the key factor? The answer must be the point stressed by Bohn (2005) that lenders expectations on future policies are a crucial part of the picture. Since these are summarized by the price of credit, a first check of this conjecture may be carried out comparing sovereign yields of FRF and Non-FRF countries. From Table 9 we can see that the cost of debt was indeed generally higher for FRF than Non-FRF countries: the median over time of the median government bond yield for the former, 8.13%, was more than 100 basis points higher than that for the latter, 7.10%. In fact, the lowest yield

[15] Note that the first of these conditions was explicitly forbidden by the 1992 Maastricht treaty on the Economic and Monetary Union which required a Debt/GDP ratio below 60% or at least following a significant negative trend. Thus, in the last two decades of our period some EU economies had to implement FRF's in order to comply with EU rules.

across all countries and years was paid by Japan, followed by Switzerland[16], both Non-FRF countries.

More formalized support to this finding is provided by a probit regression on the average sovereign spread[17] of a binary variable taking the values 1 for FRF countries and 0 for Non-FRF countries. The model (Table 10) has an excellent fit, correctly predicting 19 cases out of 22, and confirms that the probability that a country had a FRF increases with its sovereign spread, thus it declines as its fiscal reputation grows.

We thus have a seeming paradox: although the FRF ensures respect of the IBC, countries which base their fiscal policy on a FRF pay a higher price for their debt. The explanation is that causality runs from the spread, proxy of the market assessment of the fiscal reliability of a country, to the presence or absence of the FRF. Countries which pay a higher price for debt are those widely held to be prone to "profligacy" (in short, with a "bad reputation"), thus forced to show that their primary balances react to debt increases in order to have credit. "Bad reputation" countries not implementing a FRF are bound to be short-lived exceptions. The countries paying a lower price for credit are instead those which had earned the reputation of being "fiscally prudent", or perhaps considered economically strong enough to afford some profligacy (with a "good reputation"). These countries do not necessarily have to implement a

[16] Respectively, 1.01%, in 2003, and 2.07%, in 2005, see Table A.9 in the Appendix.

[17] On medium-long term government bonds, reference country Germany, source IMF. The average is generally computed over 1961-2007, except for the following countries for which data on sovereign yield start at later years: Austria (1970), Finland (1988), Greece (1993), Iceland (1992), Japan (1967), Spain (1979). Data for Portugal for 1974 and 1975 do not exist because of the troubled transition from dictatorship to democracy (the "Carnation revolution").

FRF in order to convince lenders they will respect the IBC (Bohn, 2011: "U.S. fiscal policy has relied crucially on the *perceived* safety of Treasury debt."; p. 288, italics added).

Thus, the answer to the question "who, and why, needed a FRF?" is rather clear. In our panel of advanced countries between 1961 and 2007, essentially countries with a "bad reputation", which needed to reassure markets to have credit. Two remarks: first, our conclusion is consistent with Gosh *et al.* (2013), who estimate the difference between current debt and the estimated maximum sustainable debt to be generally large for Non-FRF countries (Japan is the only clear exception) and small or altogether exhausted for FRF countries. Second, since De Grauwe (2012) showed that adverse self-fulfilling prophecies can act in a particularly powerful way in integrated financial markets, it seems that the conditions of the Maastricht treaty which constrain High Debt EU economies to implement FRF's formalize a necessity which would have been likely to exist anyway.

Table 9: Medium-Long term Government bond yields, 1961-2007

	min	*median*	*max*
Median of FRF countries	3.80	8.13	15.90
Median of Non-FRF countries	3.39	7.10	12.96
Spread	0.41	1.03	2.94

min, median, max: over the 1961-2007 period; values $\times 100$.
Source: authors' calculations on IMF data.

Table 10: A Probit model of the probability of fiscal reaction

Dependent variable:
Dummy = 1 in countries with a FRF

	constant	*average spread*
coeff [*s.e.*]	-1.83 [0.65]	0.94 [0.31]
z [$p \times 100$]	-2.84 [0.4]	3.01 [0.3]

spread on medium-long term government bonds, reference country Germany.
QML, robust s.e.'s; McFadden $R^2 = 0.41$.
Diagnostics [$p \times 100$]: Normality 6.05 [4.8]; Heteroskedasticity 0.23 [82.0].

	Predicted	
Actual	*No FRF*	*FRF*
No FRF	13	1[a]
FRF	2[b]	6

[a]: Denmark
[b]: Belgium and Canada.

5 Conclusions

Our aim was to carry out a careful assessment of model-based debt sustainability analysis for a group of 22 advanced economies from the early 1960's until the 2008 financial crisis. We obtained two main results, both of which suggest that many empirical results reported in the literature need to be revised.

First, in ten out of the 22 countries considered FRF's, both linear and non-linear, are simply out of question because the Primary balances/GDP ratios are stationary but the Debt/GDP ratios are not. This holds for four countries of the G7 (USA, Germany, Japan and France) plus six other smaller countries (Australia, Austria, Denmark, Netherlands, Sweden and Switzerland), and is in contrast for instance with Mauro *et al.* (2015), Plödt and Reicher (2015) and Everaert and Jansen (2018), who report significant FRF coefficients in some of these countries. However, these papers use invalid procedures biased in favor of the existence of a debt-primary balances relationship. As a result, the FRF appears as an essentially standard feature of the fiscal policies of advanced economies while it is not. Bohn (1998) made very clear that the FRF is a sufficient, but by no means necessary, condition for the IBC to be respected: the absence of a fiscal reaction is not a problem if there is a widespread belief that a country can readily implement the policy shifts required to ensure debt sustainability. Indeed, our probit model of the existence or not of the long-run FRF as a function of the sovereign spread (inverse proxy of reputation) confirms that countries with an average low spread (that is, a "good reputation") were not likely to adjust in the long-run primary balances in relation to debt. On the contrary, countries with a high average spread ("bad reputation") were very likely to do so. Thus, while model-based sustainability analysis is confirmed to be a useful approach for analyzing the fiscal policy of countries with low sovereign ratings, it appears to be much less useful for that of countries with high sovereign ratings. At any moment in time sustainability is a matter of reputation as much as of actually implemented fiscal policies.

Second, we found very limited evidence of "fiscal fatigue", only in Greece. Although partial, this evidence warns against the widespread practice of imposing untested homogeneity assumptions in non-linear panel FRF's.

To conclude, the lesson seems to be that the FRF is a deceivingly simple model: inadequate techniques may lead to particularly misleading conclusions.

References

[1] Banerjee, A., M. Eberhardt and J.J. Reade (2010) ”Panel Estimation for Worriers” *Economics Series Working Papers* 514, University of Oxford, Department of Economics.

[2] Berti, K., E. Colesnic, C. Desponts, S. Pamies and E. Sail (2016) “Fiscal Reaction Functions for European Union Countries” *European Economy Discussion Paper* 028, European Commission, Luxembourg.

[3] Betty, C.D., and C. Shiamptanis (2013) “Pushing the limit? Fiscal policy in the European Monetary Union” *Journal of Economic Dynamics and Control*, 37, 2307-2321.

[4] Bohn, H. (1998) ”The behavior of U.S.public debt and deficits” *Quarterly Journal of Economics*, 113, 949–963.

[5] Bohn, H. (2005) ”The sustainability of fiscal policy in the United States” *CESifo Working Paper* N. 1446.

[6] Bohn, H. (2011) ”The Economic Consequences of Rising U.S. Government Debt: Privileges at Risk” *FinanzArchiv/Public Finance Analysis*, 67, 282-302.

[7] Checherita-Westphal, C., and V. Žďárek (2017) “Fiscal reaction function and fiscal fatigue: evidence for the euro area” *ECB Working Paper* 2036

[8] D'Erasmo, P., E.G. Mendoza, J.Zhang (2016) "What is a Sustainable Public Debt?" *Handbook of Macroeconomics,* Vol. 2*,* pp. 2493-2597, North-Holland, Amsterdam (NL).

[9] De Grauwe, P. (2012) "The Governance of a Fragile Eurozone" *The Australian Economic Review*, 45, 255–268.

[10] Di Iorio, F., and S. Fachin (2014) "Savings and investments in the OECD: a panel cointegration study with a new bootstrap test" *Empirical Economics*, 46, 1271–1300, https://doi.org/10.1007/s00181-013-0722-5.

[11] Entorf, H. (1997) "Random walks with drifts: Nonsense regression and spurious fixed-effect estimation" *Journal of Econometrics*, 80, 287-296.

[12] Ermini, L., and C.W.J. Granger (1993) "Some Generalizations on the Algebra of I(1) Processes" *Journal of Econometrics,* 58, 369–384.

[13] Everaert, G., and S. Jansen (2018) "On the estimation of panel fiscal reaction functions: Heterogeneity or fiscal fatigue?" *Economic Modelling*, 70, 87–96.

[14] Fournier, J.-M., and F. Falilou (2017) "Limits to government debt sustainability in OECD countries", *Economic Modelling*, 66, 30–41.

[15] Ghosh, A.R., J.I. Kim, E.G. Mendoza, J.D. Ostry and M.S. Qureshi (2013) "Fiscal fatigue, fiscal space and debt sustainability in advanced economies" *Economic Journal*, 123, F4–F30.

[16] Golinelli, R., I. Mammi and A. Musolesi (2018) "Parameter heterogeneity, persistence and cross-sectional dependence: new insights on fiscal policy reaction functions for the Euro Area", *DSE WP* 1120.

[17] Hamilton, J.D. (1994) *Time Series Analysis* Princeton University Press, Princeton, NJ, USA.

[18] IMF-IEO (2016) *The IMF and the crises in Greece, Ireland, and Portugal* International Monetary Fund, Washington (USA).

[19] Lamé, G., M. Lequien and P.-A. Pionnier (2014) "Interpretation and limits of sustainability tests in public finance" *Applied Economics*, 46, 616-628. DOI: 10.1080/00036846.2013.861587

[20] Mark, N.C, and D. Sul (2003) "Cointegration Vector Estimation by Panel DOLS and Long-run Money Demand" *Oxford Bulletin of Economics and Statistics*, 65, 655-680.

[21] Mauro, P., R. Romeu, A. Binder and A. Zaman (2015) "A Modern History of Fiscal Prudence and Profligacy," *Journal of Monetary Economics*, 76, 55–70.

[22] Mendoza, E.G., and J.D. Ostry (2008) "International evidence on fiscal solvency: is fiscal policy 'responsible'?" *Journal of Monetary Economics*, 55, 1081–1093.

[23] Mitchell, B.R. (2011) *British Historical Statistics,* Cambridge University Press, Cambridge.

[24] Pedroni, P. (1999) “Critical values for cointegration tests in heterogeneous panels with multiple regressors” *Oxford Bulletin of Economics and Statistics*, 61, 653–670.

[25] Pedroni, P. (2001) “Fully modified OLS for heterogeneous cointegrated panels” in B. H. Baltagi, T. B. Fomby, R. Carter Hill *(Eds.) Nonstationary Panels, Panel Cointegration and Dynamic Panels*, *Advances in Econometrics*, vol. 15, 93-130.

[26] Pesaran, H.M. (2015) “Testing Weak Cross-Sectional Dependence in Large Panels” *Econometric Reviews*, 34, 1089-1117.

[27] Pesaran, M.H., Y. Shin and R.P. Smith (2001) “Bounds testing approaches to the analysis of level relationships” *Journal of Applied Econometrics*, 16, 289-326.

[28] Plödt, M., and C.A. Reicher (2015) “Estimating fiscal policy reaction functions: The role of model specification” *Journal of Macroeconomics*, 46, 113-128.

[29] Trehan, B., and C. Walsh (1991) “Testing intertemporal budget constraints: theory and applications to U.S. Federal budget and current account deficits” *Journal of Money, Credit and Banking*, 23, 210–223.

[30] Wagner, M. (2015) “The Environmental Kuznets Curve, Cointegration and Nonlinearity” *Journal of Applied Econometrics*, 30, 948-967.

[31] Wagner, M. and S.H. Hong (2016) "Cointegrating polynomial regressions: Fully Modified estimation and inference" *Econometric Theory*, 32, 1289–1315.

A Appendix

Data sources Countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK and USA. For 1961-2011: supplementary material of Mauro, Romeu, Binder and Zaman (2015). For 2012-2013: IMF, World Economic Outlook Database.

Tests for unit root tests and trends Since the FRF will be balanced if the ADF tests for pb and d have the same outcome (*i.e.*, both reject or both do not reject the null hypothesis of non stationarity) we want to test the joint null hypothesis "pb and d are $I(1)$" against the joint alternative hypothesis "pb and d are $I(0)$". We are thus in a multiple testing set-up. On the basis of the Bonferroni principle (see, e.g., Savin, 1980), separate unit root tests with a significance level α imply a joint test with an significance level 2α, so that to have an overall significance level of 5% and 10% we need to run individual tests respectively at 2.5% and 5%. The customary 10% level cannot be used for individual tests, as it will imply a joint test with an unacceptably high 20% level. In the cases where pb and d are found to be $I(1)$ we then checked if the deterministic structure of their DGPs is compatible, necessary condition for the linear FRF to be balanced. To these end we estimated $AR(p)$ models in differences, starting with a maximum number of lags $p = 4$ and the most general deterministic kernel, constant and trend (that is, a quadratic trend in the levels). Note that although in principle Primary balances cannot exceed DGP (unless defined as claims on future incomes, and even this is an extreme, purely notional case), in practice deterministic trends cannot be excluded in finite samples. For full generality we thus allowed for this possibility in the initial specification of the models.

Cointegrating polynomial regression (CPR) tests We applied three tests, Shin, P_u and LM:

(i) *Shin test*: H_0: "cointegration". It is based on the Shin (1994) variance ratio test, which, in turn, is the extension to linear cointegration of the KPSS test of the hypothesis of stationarity for random variables. Let $\hat{u}^+$ the FM-OLS residuals and $\hat{\Omega}_{u.v} = \hat{\Omega}_{uu} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{v.u}$ the conditional long-run variance of model errors with respect to the first differences of the explanatory variables. Then,

$$Shin = \frac{1}{T\hat{\Omega}_{u.v}} \sum_{t=1}^{T} \left(\frac{1}{\sqrt{T}} \sum_{j=1}^{t} \hat{u}_j^+ \right)^2$$

(ii) P_u *test*: H_0: "no cointegration". It is based on the Phillips and Ouliaris (1990) variance ratio test. Let $\hat{u}$ the CPR residuals and $\widetilde{\Omega}_{\Delta e.v}$ the estimated conditional long-run variance of the first differences of the CPR dependent variable. Then

$$P_u = \frac{\widetilde{\Omega}_{\Delta e.v}}{T^{-2}\sum_{t=1}^{T} \hat{u}_t^2}$$

(iii) *LM test*: H_0: "cointegration". This is one of the two possible versions of a specification test of the augmented-regression, RESET type, which essentially involves assessing the significance of some test variables in the polynomial model. In the Wald version these test variables (higher powers of the explanatory variable or of a linear trend) are directly added to the model and in the LM version they are used as explanatory variables in a auxiliary regression. The intuition is that when the model of interest is *not* a cointegrating regression the coefficients of the test variables will *always* have non-zero limits: if the augmented model is cointegrating thanks to consistency, and when it is not because of spurious regression. Hence, finding the test variables to be not significant is tantamount to not rejecting the hypothesis that the base model is cointegrating. Although the Wald and LM versions are asymptotically equivalent, simulation evidence in Wagner and Hong (2016) suggests that the LM version delivers overall better finite sample performances, and will thus be adopted in our empirical analysis.

The key point to be taken into account in the empirical implementation of the CPR tests is that simulation results in Wagner and Hong (2016) suggest that with our, rather small, sample size all of them may suffer from low power. Accordingly, we examined jointly all the three tests, so to consider as a null hypothesis both cointegration (Shin) and no cointegration (P_u, LM). The LM test has been constructed using the fourth power of debt as a test variable, as in Wagner (2015) for the cubic specification of the Environmental Kuznets Curve, formally identical the our FRF. Finally, given the multiple testing set-up we relied on 1% and 5% tests.

Panel cointegration test The panel cointegration *Max(HEG)* test by Di Iorio and Fachin (2014) is defined, similarly to the S_{max} test by Chang and Nguyen (2012), as the maximum over all units of the panel of the individual cointegration statistics. These are essentially Engle-Granger statistics with H_0 : "no cointegration" and H_1 : "cointegration". Since the rejection region of these statistics is the left tail of the distribution, their maximum is the most favourable to the null hypothesis of no cointegration. Hence, rejection of H_0 : "no cointegration" for the maximum implies its rejection for all other units, in which the evidence in favor of H_0 is by definition weaker. The p-value of the test is computed with a bootstrap procedure, shown analytically to be asymptotically valid under independence and by simulation to have good small sample size and power properties in panels with dependent units.

The full procedure (computation of the statistic and of the bootstrap p-value) can be carried out using the "DIF_panelcoint" add-on function for the free econometric program Gretl (http://gretl.sourceforge.net/).

References

[1] Chang, Y., and C.M. Nguyen (2012) "Residual based tests for cointegration in dependent panels" *Journal of Econometrics,* 167, 504-520.

[2] Mauro, P., R. Romeu, A. Binder and A. Zaman (2015) "A Modern History of Fiscal Prudence and Profligacy," IMF WP/13/5.

[3] Phillips, P.C.B., and S. Ouliaris (1990) "Asymptotic properties of residual based tests for cointegration" *Econometrica*, 58,165-193.

[4] Savin, N.E. (1980) "The Bonferroni and the Scheffé multiple comparison procedures" *Review of Economic Studies*, 47, 255-273.

[5] Sephton, P. S. (1995) "Response surface estimates of the KPSS stationarity test" *Economics Letters*, 47, 255-261.

[6] Shin, Y. (1994) "A Residual Based Test for The Null of Cointegration Against The Alternative of No Cointegration" *Econometric Theory,* 10, 91-115.

Table A.1: Primary balances and Debt: country summary statistics, 1961-2007

	min [yy]	median	max [yy]	min [yy]	median	max [yy]	min [yy]	median	max [yy]
		Australia			*Austria*			*Belgium*	
d	9.6 [61]	22.9	41.2 [66]	14.1 [66]	46.5	68.2 [95]	51.6 [69]	92.0	134.1 [93]
pb	-2.7	1.1	4.1	-2.0	0.7	3.3	-7.4	0.2	6.8
		Canada			*Denmark*			*Finland*	
d	43.6 [76]	67.0	101.7 [96]	4.3 [75]	37.8	80.1 [93]	3.0 [74]	14.3	57.6 [94]
pb	-0.9	2.7	10.1	-2.6	2.8	11.6	-3.9	3.0	9.6
		France			*Germany*			*Greece*	
d	14.4 [69]	29.0	66.7 [04]	17.6 [62]	39.5	68.6 [05]	14.4 [61]	40.1	107.4 [07]
pb	-3.2	-0.3	3.7	-6.0	0.8	4.3	-5.9	-2.2	4.4
		Iceland			*Ireland*			*Italy*	
d	5.6 [66]	29.1	58.9 [95]	24.7 [06]	48.3	109.2 [87]	27.2 [64]	74.4	121.8 [94]
pb	-2.8	1.1	8.5	-7.4	-0.5	6.7	-7.4	-0.8	6.6
		Japan			*Netherlands*			*New Zealand*	
d	4.4 [64]	67.3	191.6 [06]	37.8 [77]	55.0	78.5 [93]	17.4 [07]	52.0	86.5 [68]
pb	-5.3	-0.2	5.5	-3.6	0.9	5.6	-3.1	2.2	7.5
		Norway			*Portugal*			*Spain*	
d	22.4 [66]	33.0	60.5 [06]	13.5 [74]	48.5	68.3 [07]	8.2 [76]	26.7	67.5 [96]
pb	0.3	6.2	19.9	-9.1	-0.6	3.5	-5.4	0.0	4.0
		Sweden			*Switzerland*			*UK*	
d	1.6 [66]	41.5	73.3 [96]	7.1 [69]	40.2	72.2 [05]	37.5 [01]	49.3	113.8 [61]
pb	-8.1	0.5	7.1	-1.6	0.6	2.9	-4.9	-0.5	6.6
		USA							
d	41.1 [81]	56.7	71.9 [94]						
pb	-3.9	-0.6	3.9						

[yy]: year of minimum or maximum.

Table A.2: ADF tests, p-values $\times 100$, 1961-2007

	Australia	*Austria*	*Belgium*	*Canada*	*Denmark*	*Finland*	*France*	*Germany*
pb	**1.2**	**0.1**	77.2	34.3	**0.7**	6.9	**0.2**	**0.6**
d	38.9	65.9	38.2	29.7	36.3	40.6	23.4	21.2
	Greece	*Iceland*	*Ireland*	*Italy*	*Japan*	*NL*	*NZ*	*Norway*
pb	23.1	18.8	11.8	19.2	**0.4**	**0.3**	54.5	87.3
d	76.6	53.2	43.3	56.3	24.4	12.9	89.4	86.3
	Portugal	*Sweden*	*Spain*	*CH*	*UK*	*USA*		
pb	15.4	**0.6**	9.4	**1.4**	**0.2**	**0.0**		
d	43.4	56.1	43.9	60.5	**2.8**	37.1		

Lag selection: AIC, max lag=4; p-values $<$ 5% (non-stationarity rejected) in bold case.
CH: Switzerland; *NL*: Netherlands; *NZ*: New Zealand.

Table A.3: KPSS tests, 1961-2007

		Australia	*Austria*	*Belgium*	*Canada*	*Denmark*	*Finland*	*France*	*Germany*
pb	C	0.19	0.10	**0.81**	**0.60**	0.35	0.35	0.20	0.19
d	C	**0.66**	**1.20**	**0.83**	**0.70**	**0.81**	**0.93**	**1.13**	**1.22**
	CT	0.11	**0.18**	**0.20**	**0.16**	**0.19**	**0.16**	**0.27**	**0.16**
		Greece	*Iceland*	*Ireland*	*Italy*	*Japan*	*NL*	*NZ*	*Norway*
pb	C	**0.47**	**0.55**	**0.65**	**0.65**	0.27	0.45	0.41	**0.69**
d	C	**1.18**	**0.99**	0.27	**1.19**	**1.17**	0.28	**0.72**	**0.47**
	CT	**0.21**	**0.16**	**0.26**	**0.16**	**0.23**	**0.16**	**0.15**	0.13
		Portugal	*Sweden*	*Spain*	*CH*	*UK*	*USA*		
pb	C	0.33	0.34	0.31	0.12	0.10	0.16		
d	C	**1.13**	**0.93**	**0.88**	**1.00**	**0.91**	**0.57**		
	CT	0.14	**0.15**	**0.16**	0.10	**0.27**	**0.16**		

C: constant; CT: constant and trend; Window width: $4\sqrt[4]{\frac{T}{100}}$;
Critical values, 5% and 1% (computed as in Sephton, 1995): [C] 0.462, 0.723; [CT] 0.149, 0.213.
Statistics significant at 5% (stationarity rejected) in bold case.
CH: Switzerland; *NL*: Netherlands; *NZ*: New Zealand.

Table A.4: Primary Balance and Debt: AR models with constant and trend, 1961-2007

Belgium

	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e..	Coeff.	s.e.	Coeff.	s.e.
const	0.065	0.371	0.075	0.159	1.74	3.075	−0.198	1.849
time	0.006	0.013			−0.072	0.106		
ϕ_1	**−0.282**	0.145	**−0.309**	0.138	**0.415**	0.143	**0.431**	0.136
ϕ_2	0.123	0.153			**0.346**	0.156	**0.326**	0.137
ϕ_3	0.035	0.158			0.091	0.157		
ϕ_4	−0.049	0.154			−0.140	0.145		

Italy

	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e..	Coeff.	s.e.
const	−0.385	0.460	0.056	0.280	**2.690**	1.634	1.411	0.884
time	0.017	0.016			−0.050	0.058		
ϕ_1	−0.202	0.148			**0.50**5	0.146	**0.541**	0.125
ϕ_2	−0.074	0.140			0.026	0.165		
ϕ_3	**0.308**	0.145	**−0.320**	0.146	0.111	0.163		
ϕ_4	0.017	0.016			−0.050	0.058		

Norway

	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	0.035	0.504	0.352	0.210	−0.460	1.266	0.675	0.626
time	0.013	0.018			0.046	0.046		
ϕ_1	0.116	0.158			−0.007	0.145		
ϕ_2	**−0.364**	0.160	**−0.351**	0.140	−0.055	0.153		
ϕ_3	−0.202	0.157	**−0.238**	0.141	0.026	0.1523		
ϕ_4	0.005	0.163			−0.046	0.047	−0.118	0.151

Canada

	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	0.116	0.350	0.111	0.157	0.103	2.202	−0.340	1.369
time	−2e−4	0.013			−0.012	0.078		
ϕ_1	0.090	0.138			**0.619**	0.144	**0.524**	0.121
ϕ_2	−0.028	0.139			−0.127	0.164		
ϕ_3	−0.032	0.137			**0.357**	0.161	0.200	0.121
ϕ_4	**−0.291**	0.135	**−0.300**	0.134	−0.188	0.144		

Table A.5: Primary Balance and Debt: AR models with constant and trend, 1961-2007

				Finland				
	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	0.340	0.594	0.189	0.250	0.797	1.668	0.447	0.783
time	-0.006	0.021			-0.014	0.060		
ϕ_1	0.145	0.145			**1.268**	0.144	**1.268**	0.144
ϕ_2	-0.041	0.142			$\mathbf{-1.039}$	0.220	$\mathbf{-1.040}$	0.221
ϕ_3	-0.213	0.142	-0.220	0.140	**0.632**	0.218	**0.633**	0.218
ϕ_4	0.030	0.143			$\mathbf{-0.304}$	0.140	$\mathbf{-0.306}$	0.140
				Greece				
	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	0.059	0.450	0.007	0.220	1.250	1.629	**2.037**	0.870
time	-0.002	0.016			0.032	0.058		
ϕ_1	-0.188	0.145	-0.160	0.144	0.046	0.146		
ϕ_2	-0.155	0.144			-0.046	0.142		
ϕ_3	0.10	0.144			0.242	0.147	0.245	0.140
ϕ_4	0.138	0.142			-0.082	0.058		
				Iceland				
	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	-0.332	0.345	0.131	0.200	1.704	1.286	0.428	0.676
time	0.019	0.013			-0.051	0.046		
ϕ_1	-0.107	0.147			-0.008	0.159	0.244	0.142
ϕ_2	$\mathbf{-0.393}$	0.147	$\mathbf{-0.330}$	0.136	0.066	0.167		
ϕ_3	-0.080	0.157	$\mathbf{-0.238}$	0.141	-0.104	0.163		
ϕ_4	-0.077	0.159			-0.051	0.046		
				Ireland				
	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	0.057	0.404	0.042	0.183	2.223	3.189	-0.395	1.390
time	$-8e{-4}$	0.015			-0.108	0.111		
ϕ_1	0.110	0.154			**0.351**	0.152	**0.507**	0.123
ϕ_2	-0.244	0.154	-0.233	0.146	0.123	0.152		
ϕ_3	-0.036	0.152			0.024	0.153		
ϕ_4	-0.009	0.150			-0.108	0.111		

Table A.6: Primary Balance and Debt: AR models with constant and trend, 1961-2007

	Portugal							
	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	0.128	0.394	0.027	0.156	0.712	1.238	**1.137**	0.561
time	−0.004	0.014			0.018	0.044		
ϕ_1	**−0.297**	0.145	**−0.302**	0.134	0.202	0.145	0.242	0.143
ϕ_2	**−0.343**	0.149	**−0.420**	0.133	0.158	0.149		
ϕ_3	0.067	0.152	−0.220	0.140	0.056	0.149		
ϕ_4	0.182	0.145			−0.018	0.044		
	Spain							
	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	−0.410	0.282	−0.403	0.270	0.162	1.979	−0.100	1.230
time	0.017	0.010	0.017	0.009	−0.003	0.070		
ϕ_1	−0.166	0.148	−0.165	0.145	**0.411**	0.143	**0.399**	0.140
ϕ_2	0.004	0.152			**0.326**	0.157	0.243	0.143
ϕ_3	0.062	0.153			−0.065	0.156		
ϕ_4	−0.014	0.150			−0.132	0.144		
	New Zeland							
	Δb				Δd			
	AR(4)		Best model		AR(4)		Best model	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
const	−0.040	0.255	0.067	0.114	0.623	1.422	**−0.988**	0.602
time	0.004	0.009			−0.065	0.052		
ϕ_1	0.133	0.130			0.155	0.130		
ϕ_2	−0.070	0.134			0.021	0.130		
ϕ_3	−0.037	0.131			−0.021	0.127		
ϕ_4	**−0.423**	0.128	**−0.440**	0.123	**−0.445**	0.123	**−0.444**	0.123

Table A.7: A dynamic FRF for UK, 1961-2007: model selection and misspecification tests

Model specification	Diagnostics
Lag order	*No autocorrelation order 2*
$F(\alpha_3 = \beta_3 = 0) = 0.18\,[83.7]$	$LM = 0.14\,[86.6]$
$F(\alpha_2 = \beta_2 = 0) = 0.84\,[44.1]$	*No heteroskedasticity (White)*
$t(\alpha_1 = 0) = 0.57\,[57.5]$	$TR^2 = 4.69\,[86.1]$

p-values$\times$100 in brackets.

Table A.8: Long-run non-linear FRF's, CPR tests 1961-2007

	Belgium	*Canada*	*Finland*	*Greece*	*Iceland*	*Ireland*
P_u	11.14	14.9	11.69	**34.72**	23.99	9.81
Shin	0.07	0.26	0.35	0.08	0.05	0.21
LM	**13.59** [0.9]	**104.83** [0.0]	0.94 [91.8]	1.41 [84.3]	1.62 [80.5]	0.002 [100.0]
	Italy	*New Zealand*	*Norway*	*Portugal*	*Spain*	
P_u	**37.06**	18.58	22.45	26.58	11.48	
Shin	0.04	0.07	0.14	0.12	0.08	
LM	0.37 [98.5]	1.99 [73.8]	5.82 [21.3]	0.01 [100.0]	0.43 [98.0]	

Model: $pb_t = \theta + \rho_1 d_{t-1} + \rho_2 d_{t-1}^2 + \rho_3 d_{t-1}^3 + e_t$.
P_u, H_0 : "no cointegration", critical values (10%, 5%, 1%): 27.78, 34.10, 48.03;
Shin, H_0 : "cointegration", critical values (10%, 5%, 1%): 0.20, 0.37, 0.49;
LM, test variable d_{t-1}^4, H_0 : "cointegration", p-value $\times$100 in brackets;
Tests significant at 5% in bold case.

Table A.9: Medium-Long term Government bond yields: country summary statistics, 1961-2007

min	median	max	min	median	max	min	median	max
Non-FRF countries								
Australia			*Austria*			*Denmark*		
4.58	6.95	15.38	3.39	7.40	10.61	3.40	8.99	20.39
Finland			*France*			*Germany*		
3.35	5.72	13.30	3.41 7	7.54	15.85	3.35	6.83	10.37
Iceland			*Japan*			*Netherlands*		
3.73	5.12	7.75	1.01	6.09	9.26	3.37	6.59	11.55
New Zealand			*Norway*			*Sweden*		
5.07	6.69	17.71	3.28	6.38	13.56	3.38	7.39	13.49
Switzerland			*USA*					
2.07	4.29	7.15	3.88	6.58	13.91			
FRF countries								
Belgium			*Canada*			*Greece*		
3.43	7.29	13.44	4.30	7.58	15.22	3.59	6.30	23.27
Ireland			*Italy*			*Portugal*		
3.33	8.94	18.09	3.56	9.01	20.22	3.44	6.01	21.50
Spain			*UK*					
3.39	11.27	16.91	4.41	8.90	14.77			

Values $\times$*100*.
For some countries the data begin after 1961: Austria (1970), Finland (1988), Greece (1993), Iceland (1992), Japan (1967), Spain (1979).
Portugal: 1974 and 1975 not available.
Source: authors' calculations on IMF data.