

Parameter heterogeneity, persistence and cross-sectional dependence: new insights on fiscal policy reaction functions for the Euro area

Roberto Golinelli*, Irene Mammi** and Antonio Musolesi***

*Abstract*¹

A number of novelties have emerged in the study of the discretionary fiscal policy within the Euro area during the last decade. Among the others, the availability of up-to-date information on fiscal indicators for the years following the Great Recession, the introduction of cutting-edge econometric methods, and a renewed interest about the sustainability of fiscal policy and public debt. The aim of this paper is to address the challenges posed by the estimation of the discretionary fiscal reaction function for the Euro area. We exploit recently introduced testing and estimation strategies for heterogeneous dynamic panels with cross-sectional dependence and propose a new parsimonious approach. Using real-time data over the period 1996-2016, we investigate whether the fiscal policy reaction function is still a benchmark after the Great Recession. We find evidence of strong cross-sectional dependence in the panel, and clear support to a valid cointegration relationship among the main determinants of the function. Newly added covariates, such interest rate spreads, come out to play a relevant role in explaining discretionary actions.

This version: March 31st, 2018

Keywords: Dynamic panel models; Panel integration and cointegration; Heterogeneous parameters; Common correlated effects; Euro area countries; Fiscal policy reaction functions; real-time data.

JEL codes: E62, E61, H60, D80, C33.

* Department of Economics, University of Bologna
roberto.golinelli@unibo.it

** Corresponding author:
Department of Economics, Ca' Foscari University of Venice
San Giobbe, Cannaregio 873, 30121 Venezia
irene.mammi@unive.it

*** Department of Economics and Management, University of Ferrara, Italy
mslntn@unife.it

¹ Paper presented at the 10th International Conference on Computational and Financial Econometrics (CFE 2016), Technical School of Engineering, University of Seville, December 9-11, 2016; at the INFER Workshop on News and Fiscal Policy, March 2-3, 2017; Brussels, at the International Panel Data Conference (IPDC) at the University of Macedonia in Thessaloniki, Greece, July 7-8, 2017; at the Venice Econometrics Workshop (Monday 13th November), at the First Italian Workshop of Econometrics and Empirical Economics (IWEEE 2018): Panel Data Models and Applications, Milan (University of Milano-Bicocca), January 25-26, 2018; and at the University of Siena. We are grateful to Maria Elena Bontempi, Roberto Casarin, Wojciech Charemza, Stefano Fachin, Matteo Lippi Bruni, Victor Lopez-Perez, Francesco Ravazzolo, and Alex Tagliabracci for comments, as well as to conference participants.

NON-TECHNICAL SUMMARY

During the years of the Great Recession, macroeconomic and fiscal variables have experienced major shocks in the Euro Area. Large fluctuations in the structural balances on GDP occurred jointly with a surge in debt levels, thus raising widespread concerns about the consolidation of public finances. This debate has revived the empirical research about the determinants of discretionary fiscal policies in response to the economic cycle. New challenges to the estimation of fiscal policy reaction functions (FPRFs) arise due to the increasing interconnections across European economies, the diverging long-run dynamics across countries and the diffusion of fiscal policy shocks that require novel empirical approaches.

In this paper, we study the dynamic properties of FPRFs and cover the period from between 1996 to 2016 using real-time data that are not subject to ex-post revision, as they more properly reflect the information actually available to the policymaker when the decisions are taken. From a methodological viewpoint, we improve upon the existing literature that has often overlooked non-stationarity of fiscal policy determinants, heterogeneity in responses and cross-sectional dependence. We propose a flexible specification that accounts for such factors and that can help better capture the patterns of fiscal indicators. For instance, heterogeneity in the FPRF parameters can capture country-specific fiscal reactions. Moreover, unified monetary policy and fiscal rules in the Euro Area can result in a common latent factors structure in the fiscal shocks.

Furthermore, we extend the core linear FPRF by adding interest rate spreads to capture the effect on the policy stance of the lack of credibility of long-run debt sustainability. Our findings show that interest rate spreads play a significant role in capturing the dynamics and the short run fluctuations of the policy stance. In the light of this evidence, we recommend that in the context following the Great Recession, when modelling the discretionary behaviour of the policymaker researchers should explicitly include variables that capture possible non-linear effects related to the stock of debt.

1. Introduction

During the last decade, the global economic-financial turmoil and the upward trend in public debt have raised widespread concerns about euro area sovereign debt crises. In the same vein, critical issues emerged regarding the sustainability of fiscal policies and the consolidation of public finances under the constraints imposed by the Fiscal Compact and the European Stability and Growth Pact. This debate has revived the empirical research about the conduct of policy making on key matters, such as debt sustainability² and its trade-off with the target of economic stabilization.³ Alongside, a number of related topics have emerged, covering the effectiveness of policy rules,⁴ the budget discipline,⁵ and the fiscal fatigue.⁶ Finally, the Great Recession has also driven the attention to the cyclical responsiveness of the fiscal policies.⁷

Although this literature addresses wide-ranging policy questions, its unifying characteristic is the use of the fiscal policy reaction function (henceforth FPRF) to investigate policymakers' behaviour. The estimation of FPRF poses serious challenges in terms of both data and econometric methods. A first distinctive feature concerns the assumptions on the relevant information set when policy decisions are taken, reflected in the use of either *real time* or *ex post* revised data. With respect to the econometric specification of FPRF, the literature has usually estimated panel data models with individual heterogeneity assuming poolability for the slope parameters. Studies have often overlooked nonstationarity of fiscal policy determinants, heterogeneity in responses and cross-sectional dependence. To overcome these limitations, we advocate the use of flexible specifications that account for such factors as they can help better capture the patterns of fiscal indicators. For instance, allowing for heterogeneity in the FPRF parameters can better characterise possibly diverging fiscal performances due to country-specific fiscal reactions. Moreover, unified monetary policy and fiscal rules in the Euro Area can result in a common latent factors structure in the fiscal shocks.

Our main aim is to overcome methodological limitations by investigating integration, cointegration and dynamic properties of the FPRF determinants: we estimate fully heterogeneous

² See Fincke and Greiner (2012), Hajnovic et al. (2012), European Commission (2015), Cizkowicz et al. (2015), and Berti et al. (2016).

³ See Bankowski and Ferdinandusse (2017), and Codogno and Galli (2017).

⁴ See Afonso and Guimaraes (2015), Maltritz and Wuste (2015), and Szymanska (2016).

⁵ See Baldi and Staehr (2016), Tkacevs and Vilerts (2016), and van der Wielen (2016).

⁶ See Checherita-Westphalk, and Zdarek (2017), and Everaert and Jansen (2017).

⁷ See Benetrix and Lane (2013), Weichenrieder and Zimmer (2014), Kempkes (2014), Plodt and Reicher (2015), Caprioli et al. (2017), Combes et al. (2017), and Paloviita (2017).

panel models that allow for cross sectional dependence of fiscal shocks through the inclusion of common factors. We first test for persistence and cross-country correlation of the main fiscal variables. In the light of a strong evidence of nonstationarity, we further assess persistence by means of first- and second-generation panel unit root tests. Our findings point to stochastic trends in primary balances, debt and output gap, thus calling for a deeper investigation on whether these trends are common across countries. To this aim, we rely on a theoretically-grounded level-relationship among these variables to assess cointegration. Our FPRF estimates are obtained using various dynamic specifications that allow for different degrees of heterogeneity in the parameters and for latent factors with either homogeneous or heterogeneous loadings. We impose alternative restrictions within the general common correlated effects (CCE) framework by Pesaran (2006) and Chudik and Pesaran (2015) and we test the sensitivity of the results to provide a meaningful interpretation for these restrictions in the context of a FPRF. In doing so, we also propose an approach based on a two-step strategy where we first estimate a VECM on cross-sectional averaged variables and then we include the first-stage residuals in the heterogeneous FPRF equation to proxy for the common factors. This allows for a sharp reduction in the number of parameters, compared to the CCE approach, while also allowing for strong cross-sectional dependence in the data.

The main findings from fully heterogeneous dynamic CCE models come out to be robust after imposing more stringent restrictions on the parameters and the factor loadings. Given the short time span and the relatively small number of countries considered, our preferred specification is thus the restricted model with poolable slopes and latent factors having homogenous loadings.

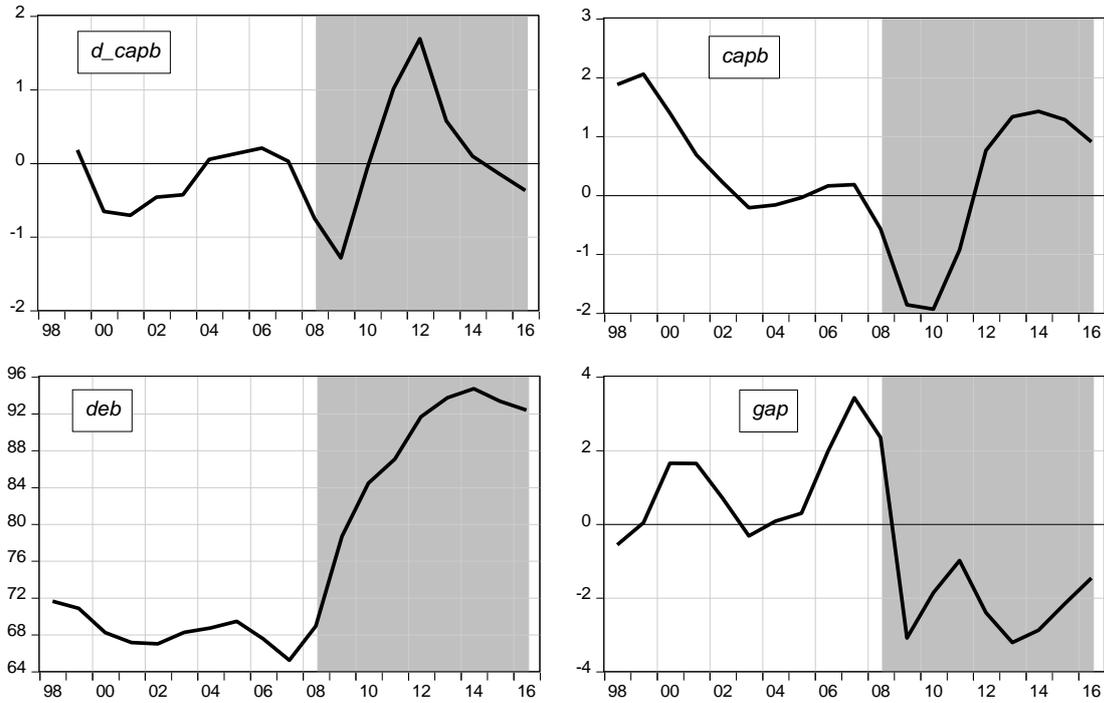
This work aims to provide a methodological contribution to the estimation of FPRFs, as well as some new insights on the underlying determinants of policymakers' discretionary behaviour in the aftermath of the Great Recession in the Euro Area. Covering the period 1996-2016, we add in the FPRF specification variables central to the policy debate, extending a core set of explanatory variables to include interest rate spreads, election years, and fiscal rule indicators. To the best of our knowledge, the role of interest rate spreads, that proxy for the credibility expressed by financial markets of the long-run sustainability of public debt, has not been investigated in the context of FPRF yet. This fills an important gap in the literature providing novel insights on the determinants of policy decision making process over an enduring phase of cyclical downturn.

2. Specification issues of the core and the extended FPRF

During the years of the Great Recession, macroeconomic and fiscal variables have experienced major shocks in the Euro Area. Since 2007 large fluctuations in both levels (*capb*) and changes

(d_capb) of the structural balances on GDP occurred jointly with a surge in debt levels. Consequently, the ratio of public debt over GDP increased by more than 20 points (from about 70% to above 90%), and since 2009 the output gap has constantly been negative. Figure 1 depicts these average trends within the Euro Area for the main fiscal variables over the last two decades.

Fig. 1 –The time pattern of selected euro area variables



Source: authors' elaboration based on data from the December 2016 issue of the OECD Economic Outlook.

Such an intense financial and economic turmoil has fuelled the policy debate and has renewed the attention on the estimation of FPRFs for the euro area since the seminal Galí and Perotti (2003).

The basic specification for the FPRF can be expressed as:

$$\Delta capb_{it} = \mu_i + \rho capb_{it-1} + \phi_{gap} gap_{it-1} + \phi_{deb} deb_{it-1} + \phi_{other} other_{it} + \varepsilon_{it}. \quad (1)$$

For each Euro Area country i , the discretionary fiscal policy in t is measured by the changes in the cyclically adjusted primary balance on potential GDP ($\Delta capb_{it}$). Being the cyclically adjusted primary balance defined as $capb_{it} = pb_{it} - \omega_i gap_{it}$, where pb_{it} is the primary balance on GDP and ω_i is the elasticity of fiscal balances to the output gap, gap_{it} , the dependent variable is net of the effect of the automatic stabilizers, and as such it proxies the discretionary fiscal policy stance.⁸

⁸ A stream of works instead measures the fiscal policy stance using the primary balance pb , thus including the effect of the automatic stabilizers. On this point see Golinelli and Momigliano (2009).

The core determinants of the target $\Delta capb_{it}$ in equation (1) are the states of the fiscal variables and of the output gap at the time the policy is set (in $t-1$, i.e. at the end of the previous fiscal year). In particular, the lagged cyclically adjusted primary balance $capb_{it-1}$ measures the structural fiscal conditions in the previous year,⁹ gap_{it} measures the business cycle, and deb_{it-1} measures the ratio of public debt over GDP.

Additional regressors are often included in empirical analyses, depending on the specific focus of each study.¹⁰ They regard common time effects τ_t , regular elections, the current account balance, financial variables, number and tightness of national fiscal policy rules, and polynomial terms of public debt to proxy for the fiscal fatigue (see Gosh et al., 2013). In model (1), we subsume these additional determinants under the term $other_{it}$.

The idiosyncratic term $\varepsilon_{it} \sim N(0, \sigma_i^2)$ captures the random fiscal shocks, i.e. those changes in legislation and other fiscal actions reflecting deviations from expectations based on past states of primary balance, economic cycle, debt, and other determinants. Although fiscal shocks can be heteroscedastic¹¹ and cross-correlated, they are usually assumed to be independently distributed across countries once their co-movements are accounted for by the common time effects τ_t or through data demeaning. The parameters $\rho, \phi_{gap}, \phi_{deb}, \phi_{other}$ are usually assumed to be invariant across countries so that potentially heterogeneous responses to changes in fiscal indicators and economic conditions have been largely disregarded so far.

The estimation of the FPRF involves both data and econometric issues. An important advantage over most of the extant literature is our use of a real-time data-set. While most studies use latest available data that are the result of several data revisions, this information is typically not available at the moment policy actions are taken (e.g. Berti et al., 2015, Checherita-Westphal and Zdarek, 2017, and Weichenrieder and Zimmer, 2014).¹² In fact, government policy is based on unrevised data that give only a preliminary knowledge of the cyclical conditions (Golinelli and Momigliano 2006,

⁹ Persistence in the FPRF dynamics is quite natural as the budgetary process often implies long parliamentary rounds to reach political consensus. Because of the inclusion of the lagged $capb$, the FPRF (1) is a simple reparameterization of the FPRF explaining $capb$ levels (rather than changes): rearranging terms of the FPRF (1), the function explaining $capb$ levels implies a different interpretation for only the lagged $capb$ parameter, $1+\rho$ instead of ρ .

¹⁰ For a thorough examination of the factors possibly causing changes in fiscal balances see Tujula and Wolswijk (2007).

¹¹ See e.g. Fatas and Mihov (2003, 2013), Fernandez-Villaverde et al. (2015), and Anzuini et al. (2017).

¹² Conversely, Schoder (2014) studies of FPRF with medium-long spans of quarterly ex post data (partially interpolated from annual observations). High frequency data may be instead useful to answer different research questions. For example, Afonso and Toffano (2013), based on a newly built quarterly data set of ex post data, focus on the existence of fiscal and monetary regimes shifts through single-country Markov switching models.

Beetsma et al. 2009, and Cimadomo 2012). The use of short spans of data comes with several limitations. As acknowledged in Favero and Marcellino (2005), short samples make the specification and the estimation of structural models complicated and exacerbate the difficulty of modelling the joint behaviour of several variables in a period of substantial institutional and economic changes. These drawbacks call for parsimonious specifications.

As for the econometric approach, the FPRF literature is usually based on pooled-slope panel data models with individual effects estimated by either OLS or IV/GMM. As such, it has so far overlooked the most recent econometric methods which allow for heterogeneous parameters and for strong and weak cross-sectional dependence.¹³ This would instead be particularly important to capture differences across countries in terms of both initial conditions (dating back to the birth of the euro area) and of subsequent policy attitudes that can better represent diverging fiscal performances. In addition, unified monetary policy and fiscal rules in the area can induce persistent common factors across fiscal shocks.

3. The main features of the core variables

The variables included in our core *fully real time* FPRF are: the dependent fiscal policy indicator ($dCAPBrt$), the lagged explanatory cyclically adjusted primary balances on potential GDP ($CAPBrt$), the output gap ($GAPrt$), and the public debt on potential GDP ($DEBTrt$).¹⁴

To disentangle inertia from news and to assess the extent of the simultaneous correlation across country news, we perform a preliminary inspection through univariate heterogeneous autoregressive models. This delivers two main findings.¹⁵ First, the degree of cross-country dependence is strong, very similar across the four variables and not related with them being measured in levels or in first differences. Second, the dependent fiscal policy indicator (in differences) is considerably less persistent than the main explanatory levels that appear to be strongly persistent.

¹³ To the best of our knowledge, the only exception being Daniel and Shiamptanis (2013), who however focus only on unit-roots and cointegration properties of ex post data disregarding the issues of cross-section dependence and cross-member cointegration across data vintages. The full set of new panel methodologies is instead used in Afonso and Rault (2010, 2015) to run several fiscal sustainability tests in the EU, but not to estimate FPRFs.

¹⁴ Details are reported in Appendix A1.

¹⁵ Detailed outcomes are in Appendix A2.

3.1 Testing and measuring cross-sectional dependence

Neglecting sample cross-sectional dependence may seriously affect empirical outcomes. In fact, traditional panel estimators provide misleading inferences or inconsistent estimates depending on whether cross-sectional dependence is *weak* or *strong* (Chudik et al., 2011; Ertur and Musolesi, 2017).

In Table 1 we disentangle weak from strong dependence by exploiting the CD test developed by Pesaran (2004, 2015a) together with the estimated confidence bands of α , the exponent of cross-sectional dependence introduced in Bailey et al. (2016), which can take any value in the range [0,1]. Starting from a canonical decomposition of a variable as a sum of deterministic, multifactor and idiosyncratic components (i.e. $y_{it} = D_{it} + \gamma_i' \mathbf{f}_t + \varepsilon_{it}$, where D_{it} is the deterministic component, \mathbf{f}_t is a vector of latent factors with heterogeneous loadings γ_i and ε_{it} is an idiosyncratic shock), α is a measure of the strength of the factors. Suppose that the factor loadings take nonzero values for M_j out of N units, Bailey et al. (2016) define the degree of cross-sectional dependence due to the j^{th} factor by $\alpha_j = \ln(M_j)/\ln(N)$, so that only N^{α_j} of the factor loadings are individually important and the overall degree of cross-sectional dependence is given by $\alpha = \max_j(\alpha_j)$.

Tab. 1 – CD test and exponent of cross-sectional dependence of the core variables

	<i>Statistic</i>	<i>90% confidence bands for α</i>		
		$\hat{\alpha}_{0.05}$	$\hat{\alpha}$	$\hat{\alpha}_{0.95}$
dCAPBrt	12.82	0.8017	0.9892	1.1747
CAPBrt	15.98	0.8497	0.9949	1.1402
DEBrt	23.63	0.8935	1.0045	1.1154
GAPrt	17.65	0.8919	0.9974	1.1029

All variables are in logs and based on real-time data. *dCAPRrt*: change in the cyclically adjusted primary balance on potential GDP between year $t-1$ and t . *CAPBrt*: cyclically adjusted primary balance on potential GDP. *GAPrt*: output gap on potential GDP. *DEBrt*: stock of gross financial liabilities on potential GDP.

As shown in Pesaran (2015a), the implicit null of the CD test is a function of the degree to which T expands relative to N . When N and $T \rightarrow \infty$ at the same rate, as it is typical in macro panels and roughly the case for our data, the implicit null of the CD test is $0 \leq \alpha < 1/4$, while for T almost fixed as $N \rightarrow \infty$, as it is typical in micro panels, the implicit null of the CD is given by $0 \leq \alpha < 1/2$. Therefore, values of α in the range [0, 1/2] correspond to different degrees of weak cross-sectional

dependence, whereas values of α in the range $[1/2, 1]$ correspond to different degrees of strong cross-sectional dependence.

The CD statistic strongly rejects the null hypothesis for all the variables, suggesting that the exponent of cross-sectional dependence is in the range $[1/4, 1]$. The bias-corrected estimates of α , along with their 90% confidence bands, can explicitly tackle the issue of weak and strong dependence: the exponent of cross-sectional dependence is estimated at approximately one for all the variables in levels, and is close to 0.99 for the variable in first differences. In addition, the 90% confidence bands are largely above 0.5 and include unity for all variables. This evidence confirms our preliminary finding, and points to strong cross-sectional dependence in both the dependent and the explanatory variables of our core FPRF.

3.2. First- and second-generation panel unit root tests

Testing for unit roots requires properly accounting for cross-sectional dependence (Pesaran, 2007; Baltagi et al. 2007). Moreover, while weak cross-sectional dependence can be addressed with a simple correction of the tests, strong cross-sectional dependence is more problematic, causing the test statistics to be divergent (Westerlund and Breitung, 2013).

As Pesaran (2007) notes, applying first-generation panel unit root tests to demeaned data (i.e. $\tilde{y}_{it} = y_{it} - \bar{y}_i$) can mitigate the adverse impact of cross-sectional dependence on test results, although cross-sectional demeaning is effective only when the pairwise errors' covariances do not strongly differ across individuals.

Table 2 reports the outcomes of a set of first-generation tests with cross-sectionally demeaned data, namely the Levin et al. (2002, LLC) test, the Im et al. (2003, IPS) test, and the alternative P, Z, L* and Pm tests described in Choi (2001).¹⁶ Results are clear-cut: the dependent fiscal policy indicator ($dCAPrt$) is stationary, while the explanatory levels of the core FPRF seem, in almost all cases, as being generated by unit root stochastic processes.¹⁷

It is worth noting that the evidence of integrated processes generating the levels of $DEBrt$, $GAPrt$ and $CAPBrt$, can be potentially misleading, as integrated series are unbounded while these variables, being defined as ratios, are likely bounded by construction. However, results in Table 2 clearly suggest that the statistical features of $CAPBrt$, $DEBrt$ and $GAPrt$ are closer to those of $I(1)$

¹⁶ For a survey, see Westerlund and Breitung (2013). Detailed features of each test are in Appendix A3, where some robustness evidence regarding the persistence of alternative measures of fiscal policy indicators is also provided.

¹⁷ $CAPBrt$ is the only exception, being classified as $I(0)$ by LLC test at 5% but not 1%, while it is $I(1)$ on the basis of the other (heterogeneous) test outcomes. This lack of robustness can be explained by LLC restrictive alternative that no series contain a unit root. In fact, note that as few as one stationary series can lead to over-rejection (i.e. rises the rejection rate above the nominal size).

than to $I(0)$ series. Therefore, for the purposes of building and estimating FPRFs is more appropriate to treat their statistical distributions as generated by $I(1)$ variables (Hall et al 1992).

Tab. 2 – First-generation unit root tests: empirical outcomes of the core variables ^a

Variable:	LLC	IPS	P	Z	L*	P _m
dCAPBr	0.000	0.000	0.000	0.000	0.000	0.000
CAPBr	0.030	0.061	0.197	0.091	0.103	0.209
DEBr	0.225	0.798	0.348	0.902	0.885	0.383
GAPr	0.475	0.258	0.586	0.585	0.581	0.622

^(a) *p*-values. All variables are in logs and based on real-time data. *dCAPBr*: change in the cyclically adjusted primary balance on potential GDP between year $t-1$ and t . *CAPBr*: cyclically adjusted primary balance on potential GDP. *GAPr*: output gap on potential GDP. *DEBr*: stock of gross financial liabilities on potential GDP.

The first-generation tests in Table 2 handle the potential sources of cross-sectional dependence by simply demeaning the data. However, the *ad hoc* use of $\tilde{y}_{it} = y_{it} - \bar{y}_i$ instead of y_{it} imposes that the common components have the same effect on each cross-sectional unit. Conversely, the second-generation unit root tests explicitly model cross-sectional dependence through the multifactor error structure, in which the heterogeneous factor loadings allow for a greater flexibility in modelling the patterns of such dependence.

In the case of a single unobserved common factor, Pesaran (2007) suggests augmenting the standard (individual) ADF-type regression with the cross-sectional averages of the first differences and of the lagged levels of the individual series. This extension yields the cross-sectionally augmented Dickey-Fuller (CADF) statistics. The individual CADF statistics are then averaged in a modified version of the IPS test called CIPS. CIPS tests whether the parameters π_i - measuring the effect of the lagged explanatory levels on their first differences - are 0 for all individuals (i.e. all the individual time series in the panel are $I(1)$ and not cointegrated), against the alternative that $\pi_i = 0$ for some individuals and $\pi_i < 0$ for others. However, in this context, the single unobservable factor has to be assumed stationary because, otherwise, all series in the panel could be $I(1)$ irrespective of whether $\pi_i = 0$.

The results of the CADF/CIPS tests reported in the first column of Table 3 confirm previous preliminary findings. The real time measure of the fiscal policy indicator *dCAPBr* is $I(0)$ as the null is rejected even at 1%, while the explanatory levels of *CAPBr*, *DEBr* and *GAPr* are $I(1)$ as the null is never rejected even at 10%.

Note as well that under the null of CADF/CIPS test, all individual time series are $I(1)$ and not cointegrated while they could be cross-cointegrated (see Banerjee et al., 2005). In addition, the presence of a unit root in the observed y_{it} is tested by imposing that the unobserved common factor is $I(0)$, while the non-stationarity of y_{it} could be due to nonstationary common factors (Bai and Ng, 2004).

To investigate more in depth the sources of data persistence, we rely on more general tests, where both the common factors and the idiosyncratic components can be tested separately for unit-roots. For this purpose, Table 3 complements the CADF/CIPS evidence with the outcomes of Bai and Ng (2004, PANIC), and of Reese and Westerlund (2016, PANICCA) approaches.

Bai and Ng (2004) propose decomposing y_{it} in three components, i.e. the deterministic D_{it} , the common \mathbf{f}_t , and the idiosyncratic ε_{it} one. A nice feature of the decomposition $y_{it} = D_{it} + \gamma_i' \mathbf{f}_t + \varepsilon_{it}$ is that it allows determining not only whether the data are stationary, but also whether nonstationarity derives either from a common or from an idiosyncratic nonstationary component, or even from both.

In particular, Bai and Ng (2004) assume for the unobservable \mathbf{f}_t and ε_{it} the following generating processes: $(\mathbf{I} - L)\mathbf{f}_t = \mathbf{C}(L)e_t$ and $(1 - \rho_i L)\varepsilon_{it} = \mathbf{B}_i(L)v_{it}$, where $\mathbf{C}(L) = \sum_{j=0}^{\infty} C_j L^j$ and $\mathbf{B}_i(L) = \sum_{j=0}^{\infty} B_{ij} L^j$ are polynomials in the lag operator. In this context, the nonstationarity of the idiosyncratic components ε_{it} (i.e. that $\rho_i = 1$) can be tested without assuming that the factors are stationary, and vice versa.

Tab. 3 – Second-generation unit root test: empirical outcomes of the core variables

	CADF / CIPS ^a	PANIC ^b				PANICCA ^b			
		ADF	Pa	Pb	PMSB	ADF	Pa	Pb	PMSB
dCAPBrt	-4.998***	0.000	0.000	0.000	0.051	0.003	0.000	0.000	0.122
CAPBrt	-1.994	0.043	0.000	0.002	0.062	0.065	0.000	0.000	0.051
DEBrt	-1.801	0.775	0.858	0.814	0.327	0.993	0.628	0.604	0.297
GAPrt	-2.018	0.085	0.001	0.038	0.116	0.157	0.000	0.024	0.093

(^a) CIPS critical values at 1%, 5% and 10% are respectively -2.6, -2.34 and .2.21; see Pesaran (2007), Table II(b) with T=20 and N=10; *** denotes 1% rejection. (^b) p-values. ADF: standard Dickey-Fuller of the first PC (PANIC) or of the cross-sectional average of the dependent variable (PANICCA) estimating the single common factor. Pa, Pb and PMSB are the panel-modified statistics in Bai and Ng (2010) to test for unit roots in the idiosyncratic component. All variables are in logs and based on real-time data. *dCAPRrt*: change in the cyclically adjusted primary balance on potential GDP between year $t-1$ and t . *CAPBrt*: cyclically adjusted primary balance on potential GDP. *GAPrt*: output gap on potential GDP. *DEBrt*: stock of gross financial liabilities on potential GDP.

The Bai and Ng (2004) approach requires to preliminary establish the number of common factors needed to represent the cross-sectional dependence in data (see e.g. Ertur and Musolesi, 2017). A recent literature suggests that a small number of unobserved common factors is sufficient to explain most of the variation in many macroeconomic variables, while information criteria are instead prone to overestimate their number (for a comparison of the two approaches see Westerlund and Urbain, 2015). In the light of this and given the small dimension of our sample, we assume here only one common factor.

To test for the nonstationarity of the idiosyncratic components ε_{it} , Bai and Ng (2004) and Reese and Westerlund (2016) proceed by pooling individual ADF t statistics obtained on defactored residuals, with the factors proxied either by principal component estimates (Bai and Ng, 2004) or by cross-sectional averages augmentation (Reese and Westerlund, 2016). As the idiosyncratic components ε_{it} in a factor model is by construction only weakly correlated across units while factors \mathbf{f}_t involve strong correlation, the pooled tests based on defactored residuals likely satisfies the cross-sectional independence assumption.

To implement unit-root tests of the idiosyncratic components we can apply the P_a and P_b statistics¹⁸ and the PMSB test statistic (Bai and Ng, 2010). In the present case, given the assumption of a single factor, a simple ADF-type test for nonstationarity of the factors can be used. (Bai and Ng, 2004 and 2010, and Reese and Westerlund, 2016).

The outcomes in Table 3, resulting from the use of PANIC/PANICCA tests, are overall consistent with those obtained using CADF/CIPS, as they confirm that the dependent fiscal policy indicator $dCAPBrt$ is stationary since both the common and the idiosyncratic components are stationary, while the explanatory levels of $CAPBrt$, $DEBrt$ and $GAPrt$ are non-stationary.¹⁹ However, using such tests also deliver new and relevant insights as they allow understanding whether the non-stationarity of these variables comes either from idiosyncratic or from factor components, which bears relevant policy implications.

Results in Table 3 suggest that the nonstationarity of the structural balances ($CAPBrt$) and output gaps ($GAPrt$) is likely due to nonstationary factors. This outcome can be rationalized in the light of both the specific span of available data and the institutional framework. Indeed, the finding

¹⁸ These statistics are the analogs of t_a and t_b of Moon and Perron (2004) who adopt a different defactoring method and consider the factors to be just nuisance parameters

¹⁹ Some minor inconsistency across idiosyncratic components unit-root tests results of P_a , P_b and PMSB can be explained in the light of the Monte Carlo simulations in Reese and Westerlund (2016), who find evidence of P_a and P_b over-rejection and of PMSB under-rejection for $N=10$ and $T=20$.

of factors' unit-roots driving $CAPBrt$ and $GAPrt$ can be attributed to peculiar features of our short sample, rather than to the effect of permanent fiscal policy and demand-side shocks, as most of the observations relates to the years when the Great Recession took place. However, also in the light of institutional and economic considerations it is not surprising that the specification with one common nonstationary factor comes out as the most realistic source of $CAPBrt$ and $GAPrt$. Given that all the countries in panel are members of the same monetary union, and are subject to common fiscal restrictions, most of the persistence in primary balances and output gap is likely due to common factors rather than to idiosyncratic dynamics.

Differently, $DEBrt$ dynamics are driven by stochastic trends both in the single common factor and in the idiosyncratic components. The trending average debt pattern depicted in Figure 1 is consistent with nonstationary debt common factors, while the nonstationarity of debt idiosyncratic components denotes a significant portion of member countries' (e.g. the PIIGS) whose debts drift apart or at least fail to converge markedly over the sample period.²⁰ This distinctive feature of debt dynamics affects the core FPRF specification, and can be investigated by adding further determinants in our FPRF.

3.3. Testing for cointegration

In the framework of the dynamic FPRF (1) whose core regressors are non-stationary, it is crucial to test for the existence of a level-relationship between them to prevent the insurgence of spurious evidence.

Prior to formal statistical testing, we provide theoretical foundations of such level-relationship. In the model of fiscal sustainability by Bohn (1998), primary balances on GDP, pb_{it} , is explained by public debt on GDP, deb_{it} , by a vector of other relevant variables, \mathbf{z}_{it} , and by an error term measuring policy shocks ε_{it} which represents the non-systematic part of the model:

$$pb_{it} = c_i + \alpha_i^d deb_{it} + \mathbf{z}_{it}' \alpha_i + \varepsilon_{it} . \quad (2)$$

The vector \mathbf{z}_{it} of the other relevant variables can be detailed relying on Barro (1979)'s tax smoothing model (see also Claeys, 2006), where the optimal tax rates depend on the permanent government spending and on the level of debt. When subtracting the non-interest government spending from Barro's tax revenues equation, the remaining determinants \mathbf{z}_{it} of the Bohn (1998) model are temporary government spending ($gtemp_{it}$), measured by the difference between the actual non-

²⁰ In addition, it is also worth noting that both first- and second-generation statistics for $CAPBrt$ and $GAPrt$ always fail to reject the null of unit-roots with lower p-values than those for $DEBrt$.

interest spending and the permanent spending on GDP), and business cycle (gap_{it} , measured by the output gap). Equation (2) can be thus rewritten as:

$$pb_{it} = c_i + \alpha_i^d deb_{it} + \alpha_i^y gap_{it} + [\alpha_i^g gtemp_{it} + \varepsilon_{it}] \quad (3)$$

If, as in our case, the target of the FPRF is the cyclically adjusted primary balance, $capb_{it}$, defined as $pb_{it} - \omega_i gap_{it}$, model (3) becomes:

$$capb_{it} = c_i + \alpha_i^d deb_{it} + (\alpha_i^y - \omega_i) gap_{it} + u_{it} \quad (4)$$

where errors $u_{it} = \alpha_i^g gtemp_{it} + \varepsilon_{it}$ embody by definition only stationary components (fiscal shocks, temporary spending and omitted dynamics) and thus they cannot explain the strong persistence of the primary balance levels displayed in our data. Model (4) can be used to test for cointegration between the $I(1)$ variables in the FPRF performing the tests by Pedroni (1999, 2004) and Di Iorio and Fachin (2012), which are the panel counterpart of the Engle and Granger (1987) residual-based cointegration test. Pedroni's test is a first-generation test entailing cross-sectional demeaning, while the Di Iorio and Fachin's approach is a second-generation test which overcomes the main drawbacks of the static Engle-Granger approach by devising a novel bootstrap strategy. Both tests aim at detecting unit roots in the error term of the long run heterogeneous representation in (4), u_{it} . Such a disturbance is expected to be stationary by the theoretical predictions of the Bohn and Barro's models.

In the Pedroni's approach, the null of unit roots in the residuals \hat{u}_{it} is tested using two alternative statistics: the panel t - and the group t -statistic, which are the analogous of LLC and IPS tests, respectively. As shown in Table 4, both tests strongly reject the null of no-cointegration.

Tab. 4 – First- and second-generation panel cointegration tests

	Panel	Group
Pedroni ADF ^a	0.000	0.000
Westerlund ECM ^a	0.058	0.061
Gengenbach et al. ECM ^b	–	–3.827
Di Iorio-Fachin, HEG ^a	–	0.0340 ^c 0.0570 ^d

(^a) p-values; (^b) averaged t-ratios 1%, 5% and 10% critical values with $p_i=2$ explanatory variables are respectively -3.38, -3.17 and -3.06; see Gengenbach et al (2016), Table 3 (supplementary material) with N=10 in case D2 (unrestricted constant). (^c) P-value corresponding to the median of the individual statistics. (^d) P-value corresponding to the mean of the individual statistics.

In Table 4 we also present the outcomes of the panel cointegration tests developed by Westerlund (2007), and Gengenbach et al. (2016), respectively.²¹ Both tests can be seen as the panel counterpart of Banerjee et al. (1998)'s dynamic ECM regression with time series, as they test for cointegration in panel ECM heterogeneous representations.²² They are second-generation tests as they handle cross sectional dependence, although with different strategies. We consider two Westerlund's cointegration test statistics: the mean-group test (G_{τ}) that averages heterogeneous OLS estimates of the speed of adjustment on their standard errors, and the panel test (P_{τ}) that estimates the aggregate speed of adjustment and its standard error. Given that both G_{τ} and P_{τ} distributions assume that the ECM errors are independently distributed, the tests account for cross-sectional dependence *via* bootstrapped standard errors (we use 1000 replications).

In the third row of Table 4, we report the group statistic for the Gengenbach et al. (2016)'s test, i.e. the average significance of the heterogeneous estimates of the ECM speed of adjustment. Differently from Westerlund's test specifications, the Gengenbach et al. (2016)'s ECM errors are assumed to have a factor structure, which is approximated by the cross-sectional averages of both dependent and explanatory variables as suggested by Pesaran (2006).

The outcomes in Table 4 from both test statistics based on dynamic ECM representations support Pedroni's result about the existence of a balanced level-relationship among the non-stationary explanatory variables of the core FPRF (1).

Finally note that both the first-generation Pedroni tests and the dynamic second-generation tests are subject to caveats regarding their performance in short T panels.²³ The Di Iorio and Fachin (2012)'s test performs better in finite samples. As shown in an extensive Monte Carlo simulation, the test has good size and power properties even with very small T and N, and with a complex dependence structure. In the last row of Table 4, we report the p -values of the bootstrap cointegration test statistics: the results confirm the previous findings irrespectively of using the median or the mean of the individual statistics.

²¹ For a survey of the panel cointegration tests, see Breitung and Pesaran (2008), and Pesaran (2015b).

²² Note that all these single-equation approaches assume the weak exogeneity of *gap* and *deb* (see Johansen, 1995) which is tested for in Appendix A4.

²³ For an extensive discussion and a detailed list of drawbacks of all these panel cointegration tests, see Di Iorio and Fachin (2012).

4. Dynamics, heterogeneity and common factors in the core FPRF

4.1 Alternative models

In the light of the evidence above, the three core variables of our FPRF are affected by strong cross-sectional dependence. They also first-order integrated but level-cointegrated within the context of heterogeneous panels. These features suggest the adoption of a dynamic heterogeneous representation of the FPRF, where fiscal idiosyncratic shocks v_{it} have a multifactor structure:

$$\begin{aligned} \Delta capb_{it} &= \mu_i + \rho_i capb_{it-1} + \phi_{i,deb} deb_{it-1} + \phi_{i,gap} gap_{it-1} + v_{it} \\ v_{it} &= \boldsymbol{\gamma}_i' \mathbf{f}_t + \varepsilon_{it} \end{aligned} \quad (5)$$

Model (5) allows for heterogeneous slopes $\boldsymbol{\beta}_i = (\rho_i \ \phi_{i,deb} \ \phi_{i,gap})'$ and factor loadings $\boldsymbol{\gamma}_i$. Parameter heterogeneity is assumed to be randomly distributed around common means, $\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i^\beta$ and $\boldsymbol{\gamma}_i = \boldsymbol{\gamma} + \boldsymbol{\eta}_i^\delta$, with $\boldsymbol{\eta}_i^j \sim IID(\mathbf{0}, \boldsymbol{\Omega}_\eta^j)$ for $j = \beta, \delta$.

This general framework encompasses a variety of panel data models as special cases. However, when the sample is finite, model generalization comes at the price of over-parameterization and more parsimonious data-congruent models, imposing restrictions on the parameters, could be preferred within a logic of a consistency-efficiency trade-off.

DCCE^{MG}. Equation (5) parameters can be consistently estimated using the Dynamic CCE Mean Group approach by Chudik and Pesaran (2015). This approach is appealing because it remains valid under weak and strong dependence (Pesaran and Tosetti, 2011) and the presence of nonstationary factors (Kapetanios et al., 2011). The underlying idea is that, the unobservable common factors \mathbf{f}_t can be approximated by the cross-sectional averages of the dependent variables and individual specific regressors, henceforth labelled $\bar{\mathbf{z}}_t$. Such proxies are regarded as nuisance variables introduced into an auxiliary regression to allow for cross-sectional dependence and to address endogeneity due to unobservables.

In the dynamic equation (5), an appropriate number of lags, p_T , of the cross-sectional averages should be introduced into an auxiliary regression function to proxy the unobservable factors and then such a function is estimated by OLS. Because of data limitation, we assume $p_T = 0$, so that

$\bar{\mathbf{z}}_t = (\overline{\Delta capb}_t, \overline{capb}_{t-1}, \overline{deb}_{t-1}, \overline{gap}_{t-1})'$ is the vector of the cross-sectional means. According to Chudik and Pesaran (2015), the number of lags has to guarantee sufficient degrees of freedom for consistent estimation and, at the same time, the number of cross-sectional averages must be at least

as large as the unobserved common factors. The four cross-sectional averages exploited here are likely to meet these requirements, as a small number of unobserved common factors is often sufficient to explain most of the variation in macroeconomic variables (Stock and Watson, 2002).

The mean slope parameters β are estimated using the mean-group (MG) estimator $\hat{\beta}_{MG}$ from the following auxiliary regression:

$$\Delta capb_{it} = \mu_i + \rho_i capb_{it-1} + \phi_{i,deb} deb_{it-1} + \phi_{i,gap} gap_{it-1} + \left\{ \delta_{i1} \overline{\Delta capb}_t + \delta_{i2} \overline{capb}_{t-1} + \delta_{i3} \overline{deb}_{t-1} + \delta_{i4} \overline{gap}_{t-1} \right\} + e_{it} \quad (6)$$

From model (6) it is easy to see that its generality involves 8 parameters to be estimated for each of the N=11 countries over the time span T=20. This warns about risks of over-parametrization and inaccurate estimates, which have been documented in previous studies adopting the CCEMG (see, e.g., Ertur and Musolesi, 2017, and Eberhardt et al., 2013).

DCCEP. A way to address the excess of parameters is to reduce the scope of heterogeneity of model (5) by assuming poolability, i.e. common slopes $\beta_i = \beta$, and estimating the following auxiliary regression:

$$\Delta capb_{it} = \mu_i + \rho capb_{it-1} + \phi_{deb} deb_{it-1} + \phi_{gap} gap_{it-1} + \left\{ \delta_{i1} \overline{\Delta capb}_t + \delta_{i2} \overline{capb}_{t-1} + \delta_{i3} \overline{deb}_{t-1} + \delta_{i4} \overline{gap}_{t-1} \right\} + e_{it} \quad (7)$$

The common slope assumption in (7) entails 30 restrictions and, consequently, the number of parameters to be estimated drops down from 88 to 58. In the CCEP, the geometry of least squares requires that the following equalities hold:

$$\frac{1}{N} \sum_{i=1}^N \hat{\delta}_{i1} = 1; \quad \frac{1}{N} \sum_{i=1}^N \hat{\delta}_{i2} = -\hat{\rho}; \quad \frac{1}{N} \sum_{i=1}^N \hat{\delta}_{i3} = -\hat{\phi}_{deb}; \quad \frac{1}{N} \sum_{i=1}^N \hat{\delta}_{i4} = -\hat{\phi}_{gap}$$

The homogeneity of the slope parameters implicit in the use of a pooled specification as in (7) has, however, been questioned from many standpoints. Among others, Brock and Durlauf (2001) warn that assuming homogeneous parameters may be inappropriate when the unit of observations are countries.

CD^{MG}. Models (6) and (7) are extremely flexible in terms of cross-sectional dependence as both allow for heterogeneous factor loadings. Conversely, Cross-sectionally Demeaning (CD) provides a parsimonious representation of cross-sectional dependence. Indeed, within the framework of Pesaran

et al. (1999)'s dynamic heterogeneous panels, the demeaning procedure leads to the restricted model:

$\Delta capb_{it} = \mu_i + \rho_i capb_{it-1} + \phi_{i,deb} deb_{it-1} + \phi_{i,gap} gap_{it-1} + e_{it}$ (where tilde superscript represents demeaned data e.g. $\Delta capb_{it} = \Delta capb_{it} - \overline{\Delta capb_t}$), or equivalently:

$$\Delta capb_{it} = \mu_i + \rho_i capb_{it-1} + \phi_{i,deb} deb_{it-1} + \phi_{i,gap} gap_{it-1} + \left\{ \overline{\Delta capb_t} - \rho_i \overline{capb_{t-1}} - \phi_{i,deb} \overline{deb_{t-1}} - \phi_{i,gap} \overline{gap_{t-1}} \right\} + e_{it} \quad (8)$$

The term in curly brackets indicates how cross-sectional dependence is handled by demeaning without the need of estimating additional nuisance parameters. Model (8) is nested in (6), as it imposes the following *untested* restrictions to the parameters of the cross-sectional averages:

$$\delta_{i1} = 1; \delta_{i2} = -\rho_i; \delta_{i3} = -\phi_{i,deb}; \delta_{i4} = -\phi_{i,gap}$$

These restrictions entail a major reduction, from 88 to 44, in the number of parameters in model (8). Finally note that the term are the averages across Euro area countries and where the parameters are restricted to be the same as those in curly brackets can be interpreted as the residuals of a hypothetical FPRF whose explanatory variables of each country's FPRF.

CD^P. The most parsimonious specification imposes common slopes, $\beta_i = \beta$, and common time effects, $\gamma_i = \gamma$, so that the multifactor structure reduces to $\gamma' \mathbf{f}_i = \gamma_t$ and the Cross-sectional Demeaning procedure of this Pooled specification can be expressed as:

$$\Delta capb_{it} = \mu_i + \rho capb_{it-1} + \phi_{deb} deb_{it-1} + \phi_{gap} gap_{it-1} + \left\{ \overline{\Delta capb_t} - \rho \overline{capb_{t-1}} - \phi_{deb} \overline{deb_{t-1}} - \phi_{gap} \overline{gap_{t-1}} \right\} + e_{it}. \quad (9)$$

Model (9) is also nested in (6), as it is obtained imposing the following restrictions on the parameters of equation (6):

$$\beta_i = \beta; \delta_{i1} = 1; \delta_{i2} = -\rho; \delta_{i3} = -\phi_{deb}; \delta_{i4} = -\phi_{gap}.$$

With $T = 20$, the dynamic FE estimates of model (9) parameters are potentially subject to the Nickell's (1981) bias due to the incidental parameter problem. To address this issue, we exploit the System GMM approach by Blundell and Bond (1998). In this context, the existence of a cointegration relation is of crucial importance to preserve the asymptotic properties of IV estimators (see, e.g. Hsiao, 1997). Moreover, it has been shown that the System GMM provides a viable approach in the case of very persistent series (Blundell and Bond, 1998).

CAS. Finally, we propose a novel approach, labelled Common Average Shocks (CAS), based on a two-step procedure. In the first-step, a proxy for the common factors is obtained using the residuals of a FPRF, estimated using cross-sectional averaged variables within a VECM framework. These residuals can be interpreted as the fiscal policy shocks affecting the whole Euro area. In the second-step, we plug these residuals into the heterogeneous-slope FPRF to account for common factors and allow for strong cross-sectional dependence.

As far as the first step is concerned, the preliminary unit root analysis supports the view that each explanatory level of the FPRF is characterized by a nonstationary common factor. Therefore, to jointly model such a common nonstationarity and to investigate more in depth relevant data features, such as the existence of a cointegrated relationship for cross-sectional averages, we adopt a VECM framework and explain the changes $\overline{\Delta capb}_t$ with the lagged levels \overline{capb}_{t-1} , \overline{deb}_{t-1} , and \overline{gap}_{t-1} . In particular, in the *first-step*, we adopt a multivariate VAR(1) representation of the three nonstationary levels in VECM form:

$$\begin{pmatrix} \overline{\Delta capb}_t \\ \overline{\Delta deb}_t \\ \overline{\Delta gap}_t \end{pmatrix} = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{pmatrix} \begin{pmatrix} \overline{capb}_{t-1} \\ \overline{deb}_{t-1} \\ \overline{gap}_{t-1} \end{pmatrix} + \begin{pmatrix} s_{1t} \\ s_{2t} \\ s_{3t} \end{pmatrix} \quad (10)$$

where the errors of the first equation $s_{1t} = \overline{\Delta capb}_t - \theta_{11}\overline{capb}_{t-1} - \theta_{12}\overline{deb}_{t-1} - \theta_{13}\overline{gap}_{t-1}$ are a model-based representation of the reduced-form fiscal shocks affecting the whole European core FPRF and can be estimated with the residuals \hat{s}_{1t} of the first equation in model (10).

As for as the interpretation of the fiscal policy shocks s_{1t} , the DCCE^{MG} approach in model (6) can be rewritten as:

$$\begin{aligned} \Delta capb_{it} = & \mu_i + \rho_i capb_{it-1} + \phi_{i,deb} deb_{it-1} + \phi_{i,gap} gap_{it-1} + \\ & \delta_{i1} \left(\overline{\Delta capb}_t + \frac{\delta_{i2}}{\delta_{i1}} \overline{capb}_{t-1} + \frac{\delta_{i3}}{\delta_{i1}} \overline{deb}_{t-1} + \frac{\delta_{i4}}{\delta_{i1}} \overline{gap}_{t-1} \right) + \varepsilon_{it} \end{aligned} \quad (6')$$

where the sum within brackets has a similar interpretation to the fiscal shocks s_{1t} because the VAR θ s parameters can be seen as reduced-form estimates of the parameters in brackets of model (6'), i.e.:

$$\theta_{11} = -\frac{\delta_{i2}}{\delta_{i1}}, \theta_{12} = -\frac{\delta_{i3}}{\delta_{i1}}, \text{ and } \theta_{13} = -\frac{\delta_{i4}}{\delta_{i1}}$$

At the *second-step*, the residuals $\hat{s}_{1t} = \overline{\Delta capb}_t - \hat{\theta}_{11} \overline{capb}_{t-1} - \hat{\theta}_{12} \overline{deb}_{t-1} - \hat{\theta}_{13} \overline{gap}_{t-1}$ are plugged into the main equation of interest, that is the core heterogeneous FPRF,

$$\Delta capb_{it} = \mu_i + \rho capb_{it-1} + \phi_{deb} deb_{it-1} + \phi_{gap} gap_{it-1} + \delta_i \hat{s}_{1t} + \varepsilon_{it} \quad (10)$$

This procedure preserves the heterogeneous-slopes assumption of the DCCE^{MG}, but imposes a number of restrictions on the parameters that lies between the DCCE^{MG} (6) and the CD^{MG} (8). Indeed, model (10) imposes restrictions on the parameters of equation (6) by using the first-step VECM residuals but permits flexibility by allowing the estimation of the loading parameters δ_i , while cross-sectional demeaning through model (8) *a priori* imposes restrictions to the parameters without any additional estimation and inference.

The CAS approach presents some appealing features. First, it allows reducing the number of parameters to be estimated, compared to the DCCE^{MG} approach. This is particularly fruitful in cases where the cross-sectional dimension is limited. Second, VAR modelling in the first-step relies on the assumption that the cross-sectional averages are informative about the common behaviour of the whole area. Finally, while, due to the small sample size, we assume a VAR(1)²⁴, in general, the choice of the number of lags can be based on standard VAR-lag selection criteria (Kilian and Lütkepohl, 2018), thus overcoming the *ad hoc* lag selection of the DCCE^{MG} approach.

4.2 Estimation results

As for the sign of the parameters in $\beta = (\mu \ \rho \ \phi_{deb} \ \phi_{gap})'$, first note that the dynamic FPRF converges to a steady state solution only if $\rho < 0$, with the long-run effects of debt and gap given by

$\alpha^d = \frac{\phi_{deb}}{-\rho}$ and $\alpha^y - \omega = \frac{\phi_{gap}}{-\rho}$, respectively. Second, as shown by Bohn (1998), $\alpha^d > 0$ is a sufficient

condition for a positive initial stock of debt to converge to zero in present value terms, this implying that the debt stock on GDP is sustainable in the long run, and this condition holds only if $\phi_{deb} > 0$.

The intuition is that, if $\phi_{deb} > 0$, the policymaker reacts to increases in public debt by systematically

²⁴ More explicitly, the choice above of a VAR(1) model is due to the assumption that, because of our small sample, $p_T = 0$. Reversing the point, we could say that, in the context of the average-VAR modelling, the choice of p_T can be seen as a problem of VAR lag-selection.

raising the primary balances, this being a sufficient condition to satisfy the intertemporal budget constraint.²⁵ Third, looking at effect of the economic cycle on the discretionary fiscal policy, if policymakers adopt a passive approach relying on automatic stabilizers only, thus implying an a-cyclical discretionary policy, then we would have $\alpha^y = \frac{\phi_{gap}}{-\rho} = \omega$ (i.e. $\phi_{gap} = -\rho \omega$). Conversely, $\alpha^y - \omega > 0$ ($\phi_{gap} > -\rho \omega$) would instead point to the implementation of additional discretionary policy intervention to stabilize the cycle (i.e. the non-automatic policy is counter-cyclical); finally, with $\alpha^y - \omega < 0$ ($\phi_{gap} < -\rho \omega$) a pro-cyclical discretionary policy would end up amplifying economic fluctuations.²⁶

We present in Table 5 the results for the alternative modelling options listed above. Along the columns, the assumptions about the degree of heterogeneity become more restrictive from the left to the right, and the number of constraints increases accordingly.

Column (1) reports the MG estimates of model (6). The absence of cross-sectional dependence in model's residuals reveals that even the most parsimonious specification within the class of DCCE models is still able to fully account for cross-country correlations. However, over-parameterization is a serious concern as confirmed by evidence of over-fitting, suggested by implausibly high values for the R^2 and low values for the RMSE, along with largely imprecise parameter estimates not significant in most cases. Slope and loading parameters are, in pairs, similar and of opposite sign, suggesting the viability of recovering efficiency through data congruent restrictions.

Model (7) addresses the excess of parameters assuming pooled-slopes but still allowing for heterogeneous loadings. Pooled estimates of slopes for lagged *capb* and *deb* reported in column (2) are very close to those in column (1), while policy pro-cyclicality emerges. However, the drop in explanatory power and the significant cross-sectional dependence in residuals suggest that the slope poolability assumption is too restrictive and raises serious statistical concerns.

Alternatively, estimating the parameters of the heterogeneous model (8) on demeaned data is an effective way to recover efficiency.

²⁵ However, the $\phi_{deb} > 0$ condition has been subsequently labelled as weak because, given the existence of fiscal limits, unbounded debt ratios on GDP may also be consistent with the transversality condition. Daniel and Shiamptanis (2013) show that, in this context, a responsible fiscal reaction function must exhibit cointegration, as we tested above. In addition, following the intuition in Mendoza and Ostry (2008), in Section 5 we will extend the determinants of model (5) to assess whether the perception that the financial markets have about the credibility of policymakers' actions and announces actually plays an explanatory role in our FPRF.

²⁶ Regarding interpretation and measurement of the fiscal stance cyclicality, see also the discussions in Claeys (2006), Golinelli and Momigliano (2009), and Plodt and Reicher (2015).

Considering first the estimates for model (6) shown in column (1), the corresponding 44 joint restrictions are not rejected (p-value=0.6201). Such evidence supports the choice of the more parsimonious model (8) whose results are reported in column (3): here the MG estimates of both lagged dependent and debt parameters are significant and with the expected sign, while – albeit not significant – the output gap parameter points to counter-cyclicality²⁷.

Tab. 5 – Alternative estimates of the core FPRF parameters ^a

	(1)	(2)	(3)	(4)	(5)	(6)
Model: ^b	DCCE ^{MG}	DCCE ^P	CD ^{MG}	CAS ^{MG}	CD ^P	CD ^P
Estimator:	OLS	OLS	OLS	OLS	OLS	GMM
CSD as: ^c	AVG	AVG	DM	CAFSF	TD	TD
<i>Slopes:</i>						
CAPBrt	-0.9000 ***	-0.9029 ***	-0.5891 ***	-0.3361 ***	-0.4326 ***	-0.5165 ***
DEBrt	0.0065	-0.0036	0.0452 **	0.0203 ***	0.0424 ***	0.0423 **
GAPrt	0.0468	-0.2161 ***	0.1353	0.1140 **	-0.0157	-0.0160
<i>Factor loadings:</i>						
dCAPBrtT	1.1142 ***	1.0000 ***	1.0000 (⊖)	0.9612 ***	1.0000 (⊖)	1.0000 (⊖)
CAPBrtT	1.0997 ***	0.9029 **	0.5891 (⊖)	0.1358 (⊖)	0.4326 (⊖)	0.5165 (⊖)
DEBrtT	-0.0155	0.0036	-0.0452 (⊖)	-0.0198 (⊖)	-0.0424 (⊖)	-0.0423 (⊖)
GAPrtT	-0.0993	0.2161	-0.1353 (⊖)	0.0446 (⊖)	0.0157 (⊖)	0.0160 (⊖)
Cons	0.3816	-0.1350	0.6567	-0.8830	-1.7483 *	-1.3441
N	11	11	11	11	11	11
T	20	20	20	20	20	20
NT	220	220	220	220	220	220
# parameters	88	58	44	58	33	33
R2	0.872	0.775	0.586	0.665	0.488	0.425
RMSE	0.977	1.170	1.525	1.429	1.644	1.775
TD (p-vals)					0.000	0.000
CD ^d	-1.542	-2.352 **	-0.114	0.957	-0.484	-1.188
Frees ^d	0.485 ***	0.494 ***	0.084	0.095	0.133 *	0.316 ***

(^a) In Italic mean-group estimates, in bold pooled estimates.

(^b) DCCE^{MG} (heterogeneous dynamic common correlated effects, and heterogeneous slopes, see model (6)); DCCE^P (heterogeneous dynamic common correlated effects, and homogeneous slopes see model (7)); CD^{MG} (common time effects, heterogeneous OLS slope estimates see model (8) for DM; CAS model (10) with CAFS;F CD^P (common time effects, OLS with pooled slopes and individual fixed effect estimated with OLS or GMM, see model (9)).

(^c) Cross sectional dependence (CSD) is accounted with AVG (model variables' averages across countries); CAS (common average fiscal shocks); DM (variables' demeaning); TD (time dummies).

(^d) Cross sectional dependence tests of Pesaran (2004, 2015a, CD) and Frees (1995).

All variables are in logs and based on real-time data. *dCAPBrt*: change in the cyclically adjusted primary balance on potential GDP between year *t-1* and *t*. *CAPBrt*: lagged cyclically adjusted primary balance on potential GDP. *GAPrt*: lagged output gap on potential GDP. *DEBrt*: lagged stock of gross financial liabilities on potential GDP.

The labels corresponding to the factor loadings refer to cross-sectional averages, whose parameters are estimated or imposed depending on the approach.

²⁷ As for the output gap, this result is right opposite to the one obtained in column (2) under the assumption of homogeneity in the slope parameters.

This finding suggests that the common time effects of model (8) are reliable proxies of the sources of cross sectional dependence because they capture the impact of common fiscal shocks that might affect all the Euro-Area countries in the same way. These common fiscal shocks may arise from their common area membership, which entails fiscal agreements and treaties.

The estimates of model (10) reported in column (4) are in line with those of model (8) in column (3), in terms of both the effects of fiscal determinants and diagnostic tests of cross-sectional dependence. This suggests that both strategies to proxy the unobservable common factors are consistent with the data at hand. However, some differences emerge in the point estimates of slope parameters. In particular, the estimated dynamics for *CAPB* in model (10) is slightly slower than in model (8). As far as factor loadings estimates are concerned, model (10) representation gives an average estimate of $-\hat{\delta}_i \hat{\theta}_{11}$ (i.e. of the \overline{gap}_{t-1} parameter) which is opposite in sign to that in column (3), i.e. to $-\hat{\phi}_{i,gap}$. The less restrictive specification of model (10), where $\delta_i \theta_{11} \neq \phi_{i,gap}$, leads to a $\hat{\phi}_{i,gap}$ estimate in column (4) which suggests significant counter-cyclical discretionary fiscal policy.

The comparison between columns (1)-(2) and (3)-(4) suggests two general considerations concerning the role of debt and output gap.

First, the heterogeneity in factor loadings is associated with non significant estimates of the parameters of debt (columns (1)-(2)), while the debt coefficient is significant and positive when imposing restrictions on common time (or fiscal) effects (columns (3)-(4)). The lack of significance in the former specification might be due either to the larger number of estimated parameters to allow for a higher degree of heterogeneity in the factor loading, or to unaccounted non-linearity in debt effects. Even under the assumption of slope poolability, the debt effect is not significant, while the gap comes up with a significantly pro-cyclical impact. Given that after the Great Recession financial markets lost confidence in the solvency of some Euro Area countries, these Governments likely implemented more restrictive policies when facing high levels of debt despite the phase of economic slowdown (as captured by negative output gaps).

Second, the estimated output gap effects come out as erratic both in sign and significance level. This outcome reinforces the concern that failing to account for non linear or threshold effects of debt might affect the reliability of the estimated output gap parameter.²⁸

²⁸ More on this issue is explicitly investigated in the next section.

The last two columns of Table 5 report the pooled-slope FE and GMM estimates, where time dummies accounts for common time effects. The FE estimates in column (5) are obtained from model (8) under the assumption of slope poolability. With $T = 20$, the FE estimates are potentially subject to the Nickell's (1981) bias due to an incidental parameter problem. To address this issue, we exploit in column (6) the System GMM approach by Blundell and Bond (1998).²⁹

The assumption of poolable slopes and the restrictions on common time effects lead to estimates that are in line with those obtained without imposing such constraints. Our results so far, based on approaches allowing for parameter and loadings heterogeneity and cross-sectional dependence, support the use of classic dynamic panel data estimators relying on slope poolability and common time effects. This being suitable in particular for small samples for which less restrictive specifications face identification issue due to over-parameterization. The use of standard panel methods is corroborated by the evidence of cointegration among the persistent explanatory levels of the core FPRG determinants.

5. Revisiting the pooled-slope dynamic panel approach: extending the core FPRF

The previous section displayed mixed evidence about the response of the discretionary policy stance to the economic cycle and the uncertain linear effect of debt. This suggests that the omission of additional fiscal policy determinants might contribute to explain the frailty of the estimates of a parsimonious specification of the FPRF. Therefore, in this section we investigate the role of other potential determinants of the discretionary fiscal policy action. Having acknowledged the advantages of a parsimonious approach, we will opt for pooled slope dynamic panel specification with common time effects.

In particular, we address the non linear effects of debt and the possible (omitted variable) bias in the estimates of the reaction of fiscal policy to the cycle by extending the set of the explanatory variables of the core FPRF. The increase in the number of parameters in the extended FPRF is made possible thanks to the parsimony of the pooled-slope dynamic panel approach with common time effects which is a valid reduction of the over-parameterized heterogeneous slope model with dynamic common correlated effects.³⁰

²⁹ We assess the robustness of the findings resulting under the assumption of slope poolability in Appendix A6.

³⁰ The evidence in Section 4 suggests that the slopes' poolability and of common time effects are neither statistically nor economically rejected in our sample.

The regressors that we include in the vector $other_{it}$ of equation (1) are: the fiscal rules' indicator (fri), the election dummies ($elect$) in t and in $t-1$, the $spreads$ between the sovereign bond yields in each member country against the German Bund dated in May and September, i.e. occurring respectively few months before ($m5$) and while ($m9$) the fiscal law is set in $t-1$ ³¹; in symbols:

$$\begin{aligned}
 other_{it} &= \left(fri_{it-1} \quad elect_{it} \quad elect_{it-1} \quad spread_{it-1}^{m5} \quad spread_{it-1}^{m9} \quad \lambda_t \right)' \\
 \phi_{other} &= \left(\vartheta \quad \theta_0 \quad \theta_1 \quad \lambda_{m5} \quad \lambda_{m9} \quad 1 \right)
 \end{aligned} \tag{11}$$

While the inclusion of fiscal rules' indicator fri and election dummies $elect$ has a long tradition in the literature, to the best of our knowledge this is the first time that $spreads$ information is exploited in a FPRF.³²

Tkacevs and Vilerts (2016) list the sovereign bond yields as a regressor in their FPRF. Yet, they aim at measuring the effects of changing borrowing costs on the policy setting and interpret bond yields merely as a measure of the change in the fiscal space. Conversely, we first define the sovereign bond yields in terms of spreads against the German Bund, then we measure with the latter the extra fiscal discipline induced by investors' fears about debt repayment in a phase of confidence crisis. Instead of modelling the non-linear effect of debt in the FPRF as in Gosh et al. (2013) and Fournier and Fall (2015), we extended the core linear FPRF adding interest rate $spreads$ to capture the lack of investor confidence. In line with Bohn (2011), when confidence crisis is triggered by doubts about fundamental solvency, further cuts in current and projected deficits represent a credible stance of debt aversion. Of course, in this context, yield differentials are assumed to reflect the market assessment about the creditworthiness of borrowers. Several empirical findings support this view: Codogno et al. (2003) provide evidence of the dominance of credit-risk factors (such as debt to GDP ratios) over the traditional liquidity indicators, and Cimadomo et al. (2016) show that a projected improvement of the fiscal outlook significantly reduces the expected sovereign spreads. De Grauwe and Ji (2016) start from a baseline model in which the interest rate spreads in the Eurozone are explained, among the others, by non-linear debt-GDP ratios effects, and provide evidence that spreads' increases can impose fiscal discipline through austerity programs.

³¹ For robustness, the timing of the two explanatory $spreads$ (here in May and September) will be further assessed in Appendix 7 by allowing the two specific months to vary along the year $t-1$.

³² For the use of fri , see e.g. the recent Afonso and Guimaraes (2015), Maltritz and Wuste (2015), Szymanska (2016), Checherita-Westphal and Zdarek (2017), and Combes et al. (2017). For the use of $elect$ and, in general, the existence of political business cycles, see the seminal Alesina et al (1997) and the papers listed in Table A.1 of Checherita-Westphal and Zdarek (2017).

More explicitly, in our FPRF, spreads are leading indicators of the fiscal policy stance changes due to the market belief of excessive debt to GDP ratios.

In the extended FPRF (1) with the additional *other* regressors defined as in (11), the fiscal policy shocks ε_{it} are assumed to be uncorrelated across countries and over time once that common correlated effects have been accounted for through *T-1* time dummies. Robust standard errors address potential heteroskedasticity.³³

Estimation results of extended FPRF are included in Table 6.

Tab. 6 – Alternative panel estimates of the fully real-time reaction function ^a

	(1) MG	(2) FE	(3) LSDVc	(4) GMM-dif	(5) GMM-sys
ρ	-0.648 *** 0.099	-0.445 *** 0.071	-0.358 *** 0.068	-0.451 *** 0.063	-0.439 *** 0.079
ϕ_{gap}	0.176 0.177	0.079 0.080	0.079 0.065	0.080 0.073	0.031 0.054
ϕ_{deb}	0.050 * 0.031	0.039 *** 0.012	0.034 *** 0.011	0.040 *** 0.010	0.026 ** 0.010
ϑ	0.453 0.535	0.085 0.147	0.056 0.203	0.115 0.136	0.340 ** 0.162
θ_0	-0.201 0.143	-0.110 0.140	-0.130 0.339	-0.119 0.126	-0.053 0.132
θ_1	-0.345 ** 0.135	-0.307 * 0.159	-0.323 0.317	-0.315 ** 0.142	-0.238 0.152
λ_{m5}	-0.536 0.374	-0.293 * 0.153	-0.315 0.219	-0.291 ** 0.132	-0.317 ** 0.154
λ_{m9}	0.899 * 0.561	0.550 ** 0.188	0.557 ** 0.237	0.547 *** 0.164	0.529 *** 0.173
N×T	220	220	220	209	220
R ²	0.525	0.441	0.441	0.435	0.476
AR1, p-val	--	--	--	0.132	0.114
AR2, p-val	--	--	--	0.295	0.295
AR3, p-val	--	--	--	0.377	0.392
Sargan	--	--	--	0.338	0.303
τ_t , p-val	demeaned	0.000	0.000	0.000	0.000

$$^a \Delta capb_{it} = \mu_i + \rho capb_{it-1} + \phi_{gap} gap_{it-1} + \phi_{deb} deb_{it-1} + \tau_t + \vartheta fri_{it-1} + \theta_0 elect_{it} + \theta_1 elect_{it-1} + \lambda' spreads_{it-1} + \varepsilon_{it}$$

^b Hausman (1978) test statistic which compares the MG and FE does not reject the null hypothesis with a p-value = 0.6371. All the variables are in logs.

³³ The inclusion of these additional variables should not bias the long run level relationships detected above for the core (very persistent) variables, as they are not much persistent by definition. Instead, these additional variables are expected to improve the explanatory power of dynamics and short term fluctuations of the policy stance.

Column (1) reports the MG estimates of the demeaned dynamic ARDL model parameters (Pesaran et al. 1999). MG estimates close to the fixed-effects estimates, (FE in column 2) signal that heterogeneity is not empirically relevant. Slope poolability is supported by not significant Hausman (1978) test.

However, with $T = 20$, the FE estimates in column (2) of Table 6 may suffer from the Nickell (1981) bias. This notwithstanding, Judson and Owen (1999) Monte Carlo results suggest that Nickell's bias is not very relevant for $T=20$ provided that model's dynamics is not highly persistent, e.g. below 0.7-0.8. Further, the authors show that the Kiviet (1995) bias-corrected LSDV (LSDVc) is the best choice overall. On such bases, we present LSDVc estimates in column (3) of Table 6.

As these biases come from the presence of the individual effects μ_i in the extended FPRF, GMM estimators take FPRF variables in differences to get rid of them and instrument the explanatory variables in differences with their levels, see the GMM-dif of Arellano and Bond (1991) in column (4). Alternatively, one could instrument the levels with both levels and differences in a system combining the equation in differences with that in levels, see the GMM-sys of Blundell and Bond (1998) in column (5) Table 6.³⁴

Given the significance of the lagged state of primary balances and of debt (initial conditions), we can conclude that the stability conditions are met (Bohn 1998). Regarding the output gap, positive and not significant ϕ_{gap} estimates denote weak counter-cyclicalities. These clear-cut indications of counter-cyclicalities conflict with the erratic findings in Table 5. Such evidence can be rationalized considering that the inclusion of interest rate spreads in the extended FPRF may have attenuated the problems due to the omission of non linear debt effects³⁵. It is worth noticing that we estimate the effect of the output gap on the discretionary policies, net of the role of automatic stabilizers; to estimate the overall output gap effect on the unadjusted primary balances, as in Bohn (1998), we would need to account for an additional response by around $\omega_i \approx 0.5$. Given our results, the overall gap effect is estimated to be below one, this figure being consistent with a moderate tax-smoothing model.

³⁴ Given that GMM estimators are often bound to be affected by an excess of identification restrictions, we will test these over-identification restrictions with the more resistant Sargan (1958) test, rather than with that of Hansen (1982). On instrument proliferation see also Bontempi and Mammi (2015).

³⁵ In addition, note that the lack of significance of the output gap effects is quite typical with OECD data, as also noted in Golinelli and Momigliano (2009).

The additional explanatory variables (*spreads*, *elect* and *fri*) are jointly significant, denoting their relevance to better explain dynamics and the short run fluctuations of the policy stance. In the autumn period, when the fiscal law is set, the ten-years interest rate spread of each country against the German Bund exerts a statistically significant role: an increase of 100 basis points in the spread of September with respect to May induces a fiscal restriction in the following year of around 0.5% of GDP and a lasting 0.3% restrictive effect that operates until the spread goes to zero. The occurrence of regular elections in the previous year induces an increase in the Government deficit, while the simultaneous effect is never significant. Finally, the restrictive effect of a growing level of the fiscal rules index (*fri*) is barely significant.³⁶

7. Concluding remarks

In this paper we have addressed the estimation of fiscal policy reaction functions for the Euro area in the context over and after the Great Recession. The threats to the sustainability of public debt, the increasing interconnections across European economies, the diverging long-run dynamics across countries and the diffusion of fiscal policy shocks have raised a number of issues and challenges that call for novel empirical approaches. Despite the shared institutional framework, the Euro Area has witnessed heterogeneous reactions to changes in the main fiscal determinants and to raising trends in public debt and drops in the output gap. Moreover, the economic environment following the onset of the financial turmoil has brought new factors such as interest rate spreads to the forefront of the agenda of policymakers. Our aim has been to tackle all these new facts and to account for the changes in the underlying determinants of fiscal policy action relying on up-to-date econometric methods and data.

From the methodological point of view, recently developed estimation and inferential methods allow to explicitly account for cross sectional dependence and for parameter heterogeneity and to properly test for stochastic trends and long-run relationships across fiscal and macroeconomic variables. We take advantage of these innovations to provide novel insights on the key policy determinants and on the discretionary response of policymakers, using real-time data over the period 1996-2016 for the Euro Area. We find evidence of strong cross-sectional dependence in the panel, nonstationarity in the main fiscal variables and clear support to a valid cointegration relationship among the main determinants of the reaction function. We estimate fiscal policy reaction functions

³⁶ The robustness of all the findings in this section is assessed in Appendix A7 by using alternative measures for debt and the output gap, and a different timing of the spread (Table A7.1), by letting the fiscal rules indicator interact with the state of other policy determinants (Table A7.2), and by using the alternative measures of fiscal stance indicator listed in Appendix A1 (Table A7.3).

by exploiting alternative dynamic specifications which allow for varying degrees of heterogeneity in the parameters and in the factor loadings.

The short time span and the small number of countries in our sample require to limit the scope for heterogeneity in order to avoid overparametrization. The main findings are robust to the imposition of an increasing number of restrictions on the parameters and the factor loadings. Overall, our results either point to a moderate procyclicality of discretionary fiscal policies or to acyclicity. As for debt, we find evidence of a shift in its aggregate level within the Euro Area which also comes along with diverging long-run debt trajectories by country. Based on our estimates, the transversality condition on debt sustainability is not always met, unless the model is extended with additional factors.

Moving from the consideration that public debt might have nonlinear effects on primary balances, we extended the core linear FPRF by adding interest rate spreads to capture the lack of credibility of long-run debt sustainability (as expressed by the investors on the financial markets). Assuming poolable slopes and homogeneous loadings for the latent factors, we find that interest rate spreads actually play a significant role in capturing the dynamics and the short run fluctuations of the policy stance, this suggesting that the extended set of covariates helps to attenuate the problems due to the omission of nonlinear debt effects. In this new context, the transversality condition on debt sustainability is always met.

References

- Afonso, A. and A. S. Guimaraes (2015), "The relevance of fiscal rules for fiscal and sovereign yield developments", *Applied Economic Letters*, Vol. 22, No. 11, pp. 920-924.
- Afonso, A. and C. Rault (2010), "What do we really know about fiscal sustainability in the EU? A panel data diagnostic", *Review of World Economics*, Vol. 145, pp. 731-755.
- Afonso, A. and C. Rault (2015), "Multi-step analysis of public finances sustainability", *Economic Modelling*, Vol. 48, pp. 199-209.
- Afonso, A. and P. Toffano (2013), "Fiscal regimes in the EU", *ECB Working Paper Series*, No. 1529.
- Alesina, A., N. Roubini and G. Cohen, *Political Cycles and the Macroeconomy*, MIT Press, Cambridge.
- Anzuini, A., L. Rossi and P. Tommasino (2017), "Fiscal policy uncertainty and the business cycle: time series evidence from Italy", *Bank of Italy Temi di Discussione*, No. 1151.
- Arellano, M. and S. Bond (1991), "Some tests of specification for panel data: Monte-Carlo evidence and an application to employment equations", *Review of Economic Studies*, Vol. 38, pp. 277-297.
- Bai, J. and S. Ng (2004), "A panic attack on unit roots and cointegration", *Econometrica*, Vol. 72, No. 4, pp. 1127-1177.
- Bai, J. and S. Ng (2010), "Panel unit root tests with cross-section dependence: a further investigation", *Econometric Theory*, Vol. 26, pp. 1088-1114.
- Bailey, N., G. Kapetanios and M. H. Pesaran (2016), "Exponent of cross-sectional dependence: estimation and inference", *Journal of Applied Econometrics*, Vol. 31, pp. 929-960.

- Baldi, G. and K. Staehr (2016), "The European debt crisis and fiscal reactions in Europe, 2000-2014", *International Economics and Economic Policy*, Vol. 13, No. 2, pp. 297-317.
- Baltagi, B. H., G. Bresson and A. Pirotte (2007), "Panel unit root tests and spatial dependence", *Journal of Applied Econometrics*, Vol. 22, No. 2, pp. 339-360.
- Banerjee, A., J. J. Dolado and R. Mestre (1998), "Error-correction mechanism tests for cointegration in a single-equation framework", *Journal of Time Series Analysis*, Vol. 19, pp. 267-83.
- Banerjee, A., M. Marcellino and C. Osbat (2005), "Testing for PPP: should we use panel methods?", *Empirical Economics*, Vol. 30, pp. 77-91.
- Bankowski, K. and M. Ferdinandusse (2017), "Euro area fiscal stance", *ECB Occasional Papers*, No. 182.
- Barro, R. J. (1979), "On the determination of the public debt", *The Journal of Political Economy*, Vol. 87, No. 5, pp. 940-971.
- Beetsma, R., M. Giuliadori and P. Wierts (2009), "Planning to cheat: EU fiscal policy in real time", *Economic Policy*, October, pp. 753-804.
- Benetrix, A. S. and P. R. Lane (2013), "Fiscal cyclicality and EMU", *Journal of International Money and Finance*, Vol. 34, pp. 164-176.
- Berti, K., E. Colesnic, C. Despouts, S. Pamies, and E. Sail (2016), "Fiscal reaction functions for European Union countries", *European Economy Discussion Papers*, No. 28.
- Blundell, R. W. and S. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, Vol. 87, pp. 115-143.
- Bohn, H. (1998), "The behavior of U. S. public debt and deficits", *The Quarterly Journal of Economics*, Vol. 113, No. 3, pp. 949-963.
- Bohn, H. (2011), "The economic consequences of rising U.S. government debt: privileges at risk", *FinanzArchiv*, Vol. 67, No. 3, pp. 282-302.
- Bontempi, M. E. and I. Mammi (2015), "Implementing a strategy to reduce the instrument count in panel GMM", *The Stata Journal*, Vol. 15, No. 4, pp. 1075-1097.
- Breitung, J. and M. H. Pesaran (2008) "Unit roots and cointegration in panels", in L. Mátyás and P. Sevestre (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Springer, pp. 279-322.
- Breusch, T. S. and A. R. Pagan (1980), "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics", *The Review of Economic Studies*, Vol. 47, No. 1, pp. 239-253.
- Brock, W. A. and S. N. Durlauf (2001), "What have we learned from a decade of empirical research on growth? Growth empirics and reality", *World Bank Economic Review*, Vol. 15, pp. 229-272.
- Busetti, F. (2006), "Preliminary data and econometric forecasting: an application with the Bank of Italy Quarterly Model", *Journal of Forecasting*, Vol. 25, No. 1, pp. 1-23.
- Caprioli, F., M. Romanelli and P. Tommasino (2017), "Discretionary fiscal policy in the euro area: past, present and future", *Bank of Italy Occasional Papers*, No. 398.
- Checherita-Westphal, C. and V. Zdarek (2017), "Fiscal reaction function and fiscal fatigue: evidence for the euro area", *ECB Working Paper Series*, No. 2036.
- Choi (2001), "Unit root tests for panel data", *Journal of International Money and Finance*, Vol. 20, No. 2, pp. 249-272.
- Chudik, A. and M. H. Pesaran (2015), "Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors", *Journal of Econometrics*, Vol. 188, No. 2, pp. 393-420.
- Chudik, A., M. H. Pesaran and E. Tosetti (2011), "Weak and strong cross-section dependence and estimation of large panels", *Econometrics Journal*, Vol. 14, No. 1, pp. 45-90.

- Cimadomo, J. (2012), "Fiscal policy in real time", *The Scandinavian Journal of Economics*, Vol. 114, No. 2, pp. 440-465.
- Cimadomo, J., P. Claeys and M. Poplawski-Ribeiro (2016), "How do experts forecast sovereign spreads?", *European Economic Review*, Vol. 87, pp. 216-235.
- Cizkowicz, P., A. Rzonca and R. Trzeciakowski (2015), "Membership in the euro area and fiscal sustainability. Analysis through panel fiscal reaction functions", *NBP Working Paper*, No. 203.
- Claeys, P. (2006), "Policy mix and debt sustainability: evidence from fiscal policy rules", *Empirica*, Vol. 33, pp. 89-112.
- Codogno, L., C. Favero and A. Missale (2003), "Yield spreads on EMU government bonds", *Economic Policy*, Vol. 18, No. 37, pp. 503-532.
- Codogno, L. and G. Galli (2017), "Can fiscal discipline be counterproductive?", *Economia Italiana*, Vol. 1-3, pp. 9-44.
- Combes, J. L., A. Minea and M. Sow (2017), "Is fiscal policy always counter- (pro-) cyclical? The role of public debt and fiscal rules", *Economic Modelling*, Vol. 65, No. 138-146.
- Corradi, V., A. Fernandez and N. R. Swanson (2009), "Information in the revision process of real-time datasets", *Journal of Business & Economic Statistics*, Vol. 27, No. 4, pp. 455-467.
- Daniel, B. C. and C. Shiamptanis (2013), "Pushing the limit? Fiscal policy in the European Monetary Union", *Journal of Economic Dynamics and Control*, Vol. 37, pp. 2307-2321.
- De Grauwe, P. and Y. Ji (2016), "How much fiscal discipline in a monetary union?", *Journal of Macroeconomics*, vol. 39, pp. 348-360.
- Di Iorio, F and S. Fachin (2012), "Savings and investments in the OECD: a panel cointegration study with a new bootstrap test", *Empirical Economics*, Vol. 46, pp. 1271-1300.
- Eberhardt, M, C. Helmers and H. Strauss (2013), "Do spillovers matter when estimating private returns to R&D", *Review of Economics and Statistics*, Vol. 95, pp. 436-448.
- Engle, R. F. and C. W. J. Granger (1987), "Co-Integration and error correction: representation, estimation, and testing", *Econometrica*, Vol. 55, No. 2, pp. 251-276.
- European Commission (2015), "Fiscal reaction functions and public debt sustainability thresholds for member states", *Fiscal Sustainability Report*, Annex A4, pp. 167-173.
- Everaert, G. And S. Jansen (2017), "On the estimation of panel fiscal reaction functions: heterogeneity or fiscal fatigue?", *Universiteit of Gent Working Paper*, No. 930.
- Ertur, C. and A. Musolesi (2017), "Weak and strong cross-sectional Dependence: a panel data analysis of international technology diffusion", *Journal of Applied Econometrics*, Vol. 32, No. 3, pp. 477-503.
- Fatás, A. and I. Mihov (2003), "The case for restricting fiscal policy discretion", *The Quarterly Journal of Economics*, Vol. 118, No. 4, pp. 1419-1447.
- Fatás, A., and I. Mihov (2013), "Policy volatility, institutions and economic growth", *The Review of Economics and Statistics*, Vol. 95, No. 2, pp. 362-376.
- Favero, C. A. and M. Marcellino (2005), "Modelling and forecasting fiscal variables for the Euro area", *Oxford Bulletin of Economics and Statistics*, Vol. 67, pp. 755-783.
- Fernandez-Villaverde, J., P. Guerron-Quintana, K. Kuester and J. Rubio-Ramirez (2015), "Fiscal volatility shocks and economic activity", *The American Economic Review*, Vol. 105, No. 11, pp. 3352-3384.
- Fincke, B. and A. Greiner (2012), "How to assess debt sustainability: Some theory and empirical evidence for selected euro area countries", *Applied Economics*, Vol. 44, pp. 3717-3724.
- Fournier, J.-M. and F. Fall (2015), "Limits to government debt sustainability", *OECD Economics Department Working Papers*, No. 1229.

- Frees, E. W. (1995), "Assessing cross-sectional correlation in panel data", *Journal of Econometrics*, Vol. 69: pp. 393-414.
- Gali, J. and R. Perotti (2003), "Fiscal policy and monetary integration in Europe", *Economic Policy*, October, pp. 533-572.
- Gengenbach, C., J.-P. Urbain and J. Westerlund (2016), "Error correction testing in panels with common stochastic trends", *Journal of Applied Econometrics*, Vol. 31, No. 6, pp. 982-1004.
- Golinelli, R. and S. Momigliano (2006), "Real-time determinants of fiscal policies in the euro area", *Journal of Policy Modeling*, Vol. 28, pp 943-964.
- Golinelli, R. and S. Momigliano (2009), "The cyclical reaction of fiscal policies in the Euro Area: the role of modelling choices and data vintages", *Fiscal Studies*, Vol. 30, No. 1, pp 39-72.
- Gosh, 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*, Vol. 123, pp. 4-30.
- Hajnovic, F., J. Zeman and J. Zilinsky (2012), "Fiscal space in the euro zone", *NBS Working Paper*, August.
- Hall, A., H. Anderson and C. W. J. Granger (1992), "A cointegration analysis of Treasury Bill yields", *The Review of Economics and Statistics*, Vol. 74, No. 1, pp. 116-26.
- Hansen, L. P.(1982), "Large sample properties of generalized method of moments estimators", *Econometrica*, Vol.50, No. 4, pp. 1029-1054.
- Hartwig Lojsch, D., M. Rodriguez-Vives and M. Slavik (2011), "The size and composition of government debt in the euro area", *ECB Occasional Paper Series*, No. 132.
- Hausman, J. A. (1978), "Specification tests in econometrics", *Econometrica*, Vol. 46, No. 6, pp. 1251-1271.
- Hsiao, C. (1997), "Statistical properties of the two-stage least squares estimator under cointegration", *The Review of Economic Studies*, Vol. 64, No. 3, pp. 385-398.
- Im, K. S., M. H. Pesaran and Y. Shin (2003), "Testing for unit roots in heterogeneous panels", *Journal of Econometrics*, Vol. 115, No. 1, pp. 53-74.
- Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.
- Judson, R. A. and A. L. Owen (1999), "Estimating dynamic panel data models: a guide for macroeconomists", *Economic Letters*, Vol. 65, pp. 9-15.
- Kapetanios, G., M. H. Pesaran and T. Yamagata (2011), "Panels with non-stationary multifactor error structures", *Journal of Econometrics*, Vol. 160, No. 2, pp. 326-348.
- Kempkes, G. (2014), "Cyclical adjustment in fiscal rules: some evidence on real-time bias for EU-15 countries", *FinanzArchiv*, Vol. 70, pp. 278-315.
- Kilian, L. and H. Lütkepohl (2017), *Structural Vector Autoregressive Analysis*, Cambridge University Press.
- Kiviet, J. F. (1995), "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models", *Journal of Econometrics*, Vol. 68, No. 1, pp 53-78.
- Levin, A., C.-F. Lin and C.-S. J. Chu (2002), "Unit root tests in panel data: asymptotic and finite-sample properties", *Journal of Econometrics*, Vol. 108, No. 1, pp. 1-24.
- Maddala, G. S. And S. Wu (1999), "A comparative study of unit root tests with panel data and a new simple test", *Oxford Bulletin of Economics and Statistics*, Vol. 61, No. S1, pp. 631-652.
- Maltritz, D. and S. Wuste (2015), "Determinants of budget deficits in Europe: the role and relations of fiscal rules, fiscal councils, creative accounting and the Euro", *Economic Modelling*, Vol. 48, pp. 222-236.
- Mendoza, E. G. And J. D. Ostry (2008), "International evidence on fiscal solvency: is fiscal policy responsible?", *Journal of Monetary Economics*, Vol. 55, pp. 1081-1093.
- Nickell, S. (1981), "Biases in dynamic models with fixed effects", *Econometrica*, Vol. 49, No. 6, pp. 1417-1426.

- Osterwald-Lenum, M. (1992), "A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics", *Oxford Bulletin of Economics and Statistics*, Vol. 54, No. 3, pp. 461–472.
- Paloviita, M. (2017), "Real time uncertainty in fiscal planning and debt accumulation in the euro area", *International Economics and Economic Policy*, Vol. 43, pp. 43-59.
- Pedroni, P. (1999), "Critical values for cointegration tests in heterogeneous panels with multiple regressors", *Oxford Bulletin of Economics and Statistics*, Vol. 61, No. S1, pp. 653-670.
- Pedroni, P. (2004), "Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis", *Econometric Theory*, Vol. 20, pp. 597-625.
- Pesaran, M. H. (2004), "General diagnostic tests for cross section dependence in panels", *Cambridge Working Papers in Economics*, No. 0435.
- Pesaran, M. H. (2006), "Estimation and inference in large heterogeneous panels with a multifactor error structure", *Econometrica*, Vol. 74, pp. 967-1012.
- Pesaran, M. H. (2007), "A simple panel unit root test in the presence of cross-section dependence", *Journal of Applied Econometrics*, Vol. 22, No. 2, pp. 265-312.
- Pesaran, M. H. (2015a), "Testing weak cross-sectional dependence in large panels", *Econometric Reviews*, Vol. 34, No. 6-10, pp. 1089-1117.
- Pesaran, M. H. (2015b), *Time Series and Panel Data Econometrics*, Oxford University Press.
- Pesaran, M. H., Y. Shin and R. P. Smith (1999), "Pooled mean group estimation of dynamic heterogeneous panels", *Journal of the American Statistical Association*, Vol. 94, pp. 621-634.
- Pesaran, M.H. and E. Tosetti (2011), "Large panels with common factors and spatial correlation", *Journal of Econometrics*, Vol. 161, No. 2, pp. 182-202.
- Plodt, M. And C. A. Reicher (2015), "Estimating fiscal policy reaction functions: the role of model specification", *Journal of Macroeconomics*, Vol.46, pp. 113-128.
- Reese, S. and J. Westerlund (2016), "Panikka: panic on cross-section averages", *Journal of Applied Econometrics*, Vol. 31, No. 6, pp. 961-981.
- Sargan, J. D. (1958), "The estimation of economic relationships using instrumental variables", *Econometrica*, Vol. 26, No. 3, pp. 393-415.
- Schoder, C. (2014), "The fundamentals of sovereign debt sustainability: evidence from 15 OECD countries", *Empirica*, Vol. 41, pp. 247-271.
- Szymanska, A. (2016), "Effects of fiscal rules on budgetary outcomes: the case of the European Union member states", *Acta Oeconomica*, Vol. 66, No. 4, pp. 597-616.
- Tkacevs, O. and K. Vilerts (2016), "The impact of sovereign bond yields on fiscal discipline", *Latvijas Banka Working Paper*, No. 5.
- Tujula, M. and G. Wolswijk (2007), "Budget balances in OECD countries: what make them change?", *Empirica*, Vol. 34, pp. 1-14.
- Van der Wielen, W. (2016), "Stochastic characterization of EMU budget discipline with policy uncertainty", *mimeo*. <http://www.vanderwielen.eu/Version%20Post%204.1.pdf>
- Weichenrieder, A. J. and J. Zimmer (2016), "Euro membership and fiscal reaction functions", *International Tax and Public Finance*, Vol. 21, pp. 598-613.
- Westerlund, J. (2007), "Testing for error correction in panel data", *Oxford Bulletin of Economics and Statistics*, Vol. 69, No. 6, pp.709-748.
- Westerlund, J., and J. Breitung (2013), "Lessons from a Decade of IPS and LLC", *Econometric Reviews*, Vol. 32, No. 5-6, pp. 547-591.

Westerlund, J. and J.-P. Urbain (2015), “Cross-sectional averages versus principal components”, *Journal of Econometrics*, Vol. 185, No. 2, pp. 372-377.

Zellner, A. (1962), “An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias”, *Journal of the American Statistical Association*, Vol. 57, No. 298, pp. 348-368.

FURTHER RESULTS INTENDED FOR ON LINE APPENDICES

Appendix A1 – The real-time data-set and alternative fiscal policy indicators

Each winter issue of the OECD Economic Outlook (EO) reports the time series of the variables used in this study at annual frequency. We define as $Y_{\text{calendar}|\text{issue}}$ the observation of Y for year "calendar" published in the year "issue", where Y can be alternatively primary balances, output gap, and debt. Table A1.1 reports the main features of the data-set for any variables of interest, Y .

Tab. A1.1 – A sketch of the real time data-set

calendar date:	EO issue:	t-1	t	t+1
...		$Y_{\dots t-1}$	$Y_{\dots t}$	$Y_{\dots t+1}$
t-1		$Y_{t-1 t-1}$	$Y_{t-1 t}$	$Y_{t-1 t+1}$
T		$Y_{t t-1}$	$Y_{t t}$	$Y_{t t+1}$
t+1		<i>n.a.</i>	$Y_{t+1 t}$	$Y_{t+1 t+1}$
t+2		<i>n.a.</i>	<i>n.a.</i>	$Y_{t+2 t+1}$

Along the columns, a portion of the time series published on different winter issues (of year $t-1$, t , $t+1$, ...) is reported. Therefore, the column "issue" summarizes the available information at the time the issue is released. The variability of Y along each column is due to economic events, policy and institutional changes, etc. The variable Y in each column spans from the beginning of data availability to the one-year ahead forecast with respect the issue year. In particular, those observations for which the date of calendar is smaller than the date of issue, e.g. $Y_{t-1|t}$, denote historical data: the observation for year $t-1$ available in winter of year t (that corresponds to the latest historical information). Those observations for which the date of calendar is the same as the date of issue, e.g. $Y_{t|t}$, denotes partly estimated data: the observation for year t available in winter of year t . Those observations for which the date of calendar is larger than the date of issue, e.g. $Y_{t+1|t}$, denote predictions: the forecast of Y for year $t+1$ made in t (i.e. one-year ahead with respect the latest available historical observation of the issue t).

Along the rows, for each calendar year $t-1$, t , $t+1$, $t+2$..., a portion of different data releases of the same observation is reported. The variability of Y along each row is due to data revisions, a process that improves the quality of the observations as soon as time passes by (because they have been subject to a larger number of revisions).

Our purpose is to explain the evolution over time of a general fiscal policy indicator, $fpi_{t|t}$, by using end-of-year t observations. To do so, we assume that the relevant explanatory variables are dated at the time the fiscal law is set, i.e. in autumn/winter of the previous issue year $t-1$ when the latest available observations are dated $t-1$. Therefore, we denote with $x_{t-1|t-1}$ the state of the knowledge at the time the fiscal law is set, i.e. the vector of the explanatory variables of $fpi_{t|t}$.

In particular, the $x_{t-1|t-1}$ vector is made of three variables measured in real-time: CAPBrt, DEBrt and GAPrt.

Our estimates of the target fiscal policy indicator, i.e. the dependent variable in FPRF (1), can be computed by taking the first differences of $capb$ in t from the previous period $t-1$. Depending on the $capb$ observations used, we can construct four alternative indicators.

The policy intentions or plans for t refer to the fiscal law voted at time $t-1$ and measure the expected (predicted) one year ahead change of $capb$ from its actual level in the issue $t-1$ (see Beetsma et al, 2009). Formally, they are defined as: **PCAPBrt** = $capb_{t|t-1} - capb_{t-1|t-1}$.

Relatedly, the differences between plans for t of the issue $t-1$ and their actual realization measured one year later in the issue t (i.e. the deviations between first released observations and one-year-before reported forecasts) are labelled as policy implementation: **ICAPBrt** = $capb_{t|t} - capb_{t|t-1}$. ICAPBrt is due to all decisions occurred after the fiscal law is approved at the end of year $t-1$ until the end of year t .

A second measure of fpi is the change of actually observed $capb$ between the calendar dates $t-1$ and t in the same issue t : **DCAPBrt** = $capb_{t|t} - capb_{t-1|t}$.

DCAPBrt differs from previous PCAPBrt in two respects: the amount of policy implementation (ICAPBrt) and the data revisions from the first release in $t-1$ to the second release in t , i.e. **RCAPBrt** = $capb_{t-1|t-1} - capb_{t-1|t}$. Therefore, DCAPBrt can be represented by the sum of three components: PCAPBrt, ICAPBrt and RCAPBrt.

Given that the latter component of DCAPBrt concerns revisions and not policy measurements, a third alternative measure of fpi only embodies policy components: plans (PCABrt) and implementations (ICAPBrt): **dCAPBrt** = $capb_{t|t} - capb_{t-1|t-1}$.

With the latter fpi measure, the FPRF (1) is a "fully real time" dynamic model, as both its dependent variable (in differences) and the lagged explanatory levels of $capb$ are both obtained from CAPBrt, see Cimadomo (2012).

Finally, as in Golinelli and Momigliano (2009), we can measure fpi with the change in the most recently published ("latest available", l.a.) time series for $capb$: **dCAPB** = $capb_{t|l.a.} - capb_{t-1|l.a.}$. dCAPB has the advantage of using the best statistical information given that its observations have been revised many times. However, the observations of $capb$ latest available series have been revised a different number of times and, because of this, are heterogeneous over time, see e.g. Busetti (2006), and Corradi et al (2009). More explicitly, the change from one year to the other of dCAPB might be affected not only by policies but also by different degrees of data revisions.

Despite the different definitions, the four alternative measures are highly correlated (with coefficients usually above 0.8) denoting quite similar empirical features, see Table A1.2.

Tab. A1.2 - Degree of correlation across different fiscal policy indicators

	dCAPBrt	dCAPB	DCAPBrt	PCAPBrt
dCAPBrt	1			
dCAPB	0.7852	1		
DCAPBrt	0.8505	0.8371	1	
PCAPBrt	0.6455	0.6098	0.7051	1

The PCABrt is a partial exception to the rule: coefficients with actual (and not planned) dCAPB and dCAPBrt are about 0.6 because plans ignore policy implementations. The correlation between PCABrt and DCAPBrt is slightly higher (0.73) because data revisions counterbalance the implementation effect.

Appendix A2 – Preliminary inspection to data: detailed results

The analysis of the persistence for the real-time levels of the cyclically adjusted primary balances on potential GDP ($CAPBrt$), the output gap ($GAPrt$), and debt on GDP ($DEBTrt$), and for the alternative fiscal policy indicators (defined in Appendix A1) is conducted in the context of an univariate AR(2) framework to disentangle inertia (in squared brackets) from news v_{it} :

$$y_{it} = [\beta_{0i} + \beta_{1i}y_{it-1} + \beta_{2i}y_{it-2}] + v_{it} \quad (A2.1)$$

where y_{it} corresponds to the variables of interest. In the classical context of SUR equations approach (Zellner, 1962), inertia parameters are heterogeneous (β s are i -specific), and news can be heteroskedastic, $E(v_{it}) = 0; Var(v_{it}) = \sigma_{ii}$, independent over time $Cov(v_{it}, v_{is}) = 0; \forall t \neq s$, and simultaneously cross-country correlated, $Cov(v_{it}, v_{jt}) = \sigma_{ij}; \forall i \neq j$. Compactly, $V(v) = \Sigma_t \otimes I_T$,

where: $\Sigma_t = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{N1} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \dots & \sigma_{NN} \end{pmatrix}$. Inertia (persistence) by country i is measured by the sum $(\beta_{1i} + \beta_{2i})$.

Cross-sectional dependence is preliminarily investigated by looking at the average cross-correlation coefficient by country that are obtained from the AR(2) specification in (A2.1) and are

given by $cc_i = \sqrt{\frac{\sum_{\forall j \neq i} \rho_{ij}^2}{N-1}}$, where $\rho_{ij}^2 = \frac{\sigma_{ij}^2}{\sigma_{ii} \times \sigma_{jj}}$ is the squared cross-country (i and j) correlation

coefficient between news.

Table A2.1 reports the main results.

Tab. A2.1 – Persistence and cross-country dependence: preliminary evidence

y_{it} :	Persistence			Cross-country dependence		
	$(\beta_{1i} + \beta_{2i})$			$cc_i = \sqrt{\frac{\sum_{\forall j \neq i} \rho_{ij}^2}{N-1}}$		
	<i>min</i>	<i>max</i>	<i>Avg</i>	<i>min</i>	<i>max</i>	<i>Avg</i>
dCAPBr	-0.759	0.573	-0.288	0.176	0.575	0.452
dCAPB	-0.594	0.335	-0.173	0.183	0.383	0.327
DCAPBr	-0.518	0.602	0.000	0.299	0.548	0.446
PCAPBr	-0.444	0.692	0.297	0.262	0.412	0.351
CAPBr	0.255	0.969	0.543	0.222	0.556	0.459
DEBr	0.623	1.017	0.852	0.212	0.487	0.396
GAPr	0.038	0.901	0.507	0.249	0.628	0.570

All variables are in logs and based on real-time data. *dCAPRr*: change in the cyclically adjusted primary balance on potential GDP between year $t-1$ and t . *CAPBr*: cyclically adjusted primary balance on potential GDP. *GAPr*: output gap on potential GDP. *DEBr*: stock of gross financial liabilities on potential GDP.

The alternative measures of fiscal policy indicators in the upper part of the table are considerably less persistent than the explanatory variables of the FPRF in the lower part. Moreover, the degree of cross-country dependence is quite stable across different variables. Given that $N=11$, the LM test of Breusch and Pagan (1980) can be applied. It is based on $\hat{\rho}_{ij}^2$ and tests the null that all the pairwise error covariances are equal to 0. The null hypothesis is strongly rejected in all cases.

Finally, Table A2.2 reports measures of cross-sectional dependence discussed in the main text for alternative indicators of fiscal policy: the robustness of the outcomes clearly emerges.

Tab. A2.2 – CD test and exponent of cross-sectional dependence of alternative indicators

	<i>statistic</i>	<i>90% confidence bands for α</i>		
		$\hat{\alpha}_{0.05}$	$\hat{\alpha}$	$\hat{\alpha}_{0.95}$
	CD			
dCAPBr	12.82	0.8017	0.9892	1.1747
dCAPB	8.32	0.7761	0.9106	1.0451
DCAPBr	11.75	0.8297	0.9853	1.1410
PCAPBr	11.97	0.7975	0.9331	1.0687

Appendix A3 – Unit roots evidence: details and robustness

Table A3.1 summarizes the main statistical features of the first-generation tests whose outcomes are reported in the main text.

Tab. A3.1 – First-generation unit root tests: summary of the main statistical features

Test	Asymptotics	H_0	H_1
Levin et al (2002), LLC	$N/T \rightarrow 0$	$\pi_i=0$	$\pi_i < 0$
Im et al (2003), IPS	$(T, N) \rightarrow_{\text{seq}} \infty$	$\pi_i=0 \forall i$	$\pi_i < 0 \forall i \in N_s$ $N_s/N \rightarrow \delta$ as $N_s, N \rightarrow \infty$
Maddala and Wu (1999), Choi (2001) ^a	$T \rightarrow \infty$, N finite or $\text{seq } \infty$	$\pi_i=0 \forall i$	$\pi_i < 0 \forall i \in N_s$ $N_s/N \rightarrow \delta$ as $N_s, N \rightarrow \infty$

(^a) Four test are computed (see Choi, 2001, pp. 253-256): P, Z, L* and P_m. All of them combine p-values from the ADF test by country

In addition, Tables A3.2a and A3.2b report the outcomes of alternative measures of fiscal policy indicators. Again, the robustness of the outcomes of different variables and approaches is evident.

Tab. A3.2a – First-generation tests: robustness across fiscal policy indicators ^a

\tilde{y}_{it} :	LLC	IPS	P	Z	L*	P _m
dCAPBrt	0.000	0.000	0.000	0.000	0.000	0.000
dCAPB	0.000	0.000	0.000	0.000	0.000	0.000
DCAPBrt	0.000	0.000	0.000	0.000	0.000	0.000
PCAPBrt	0.000	0.000	0.000	0.000	0.000	0.000

(^a) p-values.

Tab. A3.2b – Second-generation tests: robustness across fiscal policy indicators

	CADF / CIPS ^a	PANIC ^b				PANICCA ^b			
		ADF	Pa	Pb	PMSB	ADF	Pa	Pb	PMSB
dCAPBrt	-4.998***	0.000	0.000	0.000	0.051	0.003	0.000	0.000	0.122
dCAPB	-5.232***	0.000	0.000	0.004	0.083	0.005	0.002	0.045	0.179
DCAPBrt	-4.523***	0.000	0.000	0.000	0.058	0.016	0.000	0.000	0.045
PCAPBrt	-3.785***	0.000	0.000	0.000	0.045	0.038	0.000	0.029	0.151

(^a) CIPS critical values at 1%, 5% and 10% are respectively -2.6, -2.34 and .221; see Pesaran (2007), Table II(b) with T=20 and N=10; *** denotes 1% rejection. (^b) p-values. ADF: standard Dickey-Fuller of the first PC (PANIC) or of the average (PANICCA) estimating the single common component; Pa, Pb and PMSB are the panel-modified statistics in Bai and Ng (2010) to test for unit roots in the idiosyncratic component.

Appendix A4 – Weak exogeneity tests

Given that in principle there could be a feedback from the level relationship disequilibria to the changes in the output gap and in debt, we must validate the weak exogeneity assumption of the latter in the context of the VECM estimate:

$$\begin{pmatrix} \Delta capb_{it} \\ \Delta deb_{it} \\ \Delta gap_{it} \end{pmatrix} = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{pmatrix} \begin{pmatrix} 1 & -\beta_2 & -\beta_3 \end{pmatrix} \begin{pmatrix} capb_{it-1} \\ deb_{it-1} \\ gap_{it-1} \end{pmatrix} + \begin{pmatrix} v_{1it} \\ v_{2it} \\ v_{3it} \end{pmatrix} \quad (\text{A4.1})$$

The vector white noise assumption for the errors v can be met by augmenting the model with lagged differences. In the VECM specification (A4.1), the weak exogeneity of deb and gap in explaining $capb$ long run dynamics can be assessed by testing for $\pi_2 = \pi_3 = 0$, while single-equation cointegration evidence needs that $\pi_1 < 0$.

VECM results in Table A4.1 support the cointegration results through single-equations in the main text because of the evidence of deb and gap weak exogeneity, as the null hypothesis $\pi_2 = \pi_3 = 0$ is almost always not rejected, and the null hypothesis $\pi_1 = 0$ is always rejected independently from the parameter estimates used to define the disequilibria.

In addition, we tested for unit roots (with LLC and IPS) all the nine ECM disequilibria corresponding to the alternative long run (cointegrated) parameters. In our analysis the null of no-cointegration is always rejected. In this regard, note that the only ECM disequilibrium to be inspected for stationarity in Table 4 is that for $\beta_2 = 0.061$ and $\beta_3 = 0.073$ through Pedroni test.

Tab. A4.1 – Weak exogeneity in VECM ^a

	$gap, \hat{\beta}_3 =$	0.07		0.10		0.46	
$deb, \hat{\beta}_2 =$	$\pi_2 = \pi_3 = 0$	$\pi_1 = 0$	$\pi_2 = \pi_3 = 0$	$\pi_1 = 0$	$\pi_2 = \pi_3 = 0$	$\pi_1 = 0$	
0.03	0.161	0.000	0.142	0.000	0.022	0.001	
0.06	0.411	0.003	0.349	0.003	0.098	0.002	
0.11	0.646	0.031	0.631	0.028	0.382	0.016	

^(a) p -values. In bold, the cases corresponding to the long run estimates from the tests in Table 4.

Appendix A5 – Modelling the core FPRF for the average euro area country

Tests for cointegration rank and weak exogeneity in Table A5.1 suggest that between the three non-stationary averages \overline{capb}_t , \overline{deb}_t and \overline{gap}_t there is one (weakly exogenous) long-run relationship. This fact, accompanied by sound long-run and speed of adjustment estimates in Table A5.1, allow us to interpret the residuals of the first equation of the VAR as a way to estimate the common fiscal shocks affecting the average country.

Instead of using the restricted model residuals to obtain τ_t in equation (A5.2), we can also make a step backward and estimate τ_t with the residuals of a unrestricted VAR(1) model for the three variables of interest, i.e. without imposing any reduced-rank and weak-exogeneity restrictions to VAR model parameters. Given that we are interested in obtaining at least reduced form shocks, we use this second VAR-based measure in the main text (without the need of imposing any restriction), and leave the first one to check for robustness.

Tab. A5.1 - Johansen (1995) rank cointegration and weak exogeneity tests

Sample: 1998 2016

No deterministic trend, restricted intercept and year dummies for 2008 and 2010

Series: cross-section averages of: *CAPBrt DEBrt GAPrt*

1 lag of first differences in VAR

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Max-Eigen Statistic	With 5% c.v. of Ostervald-Lenum (1992)
None	0.6829	28.67	21.82	Reject
At most 1	0.2757	6.85	6.13	Not reject
At most 2	0.0373	0.72	0.72	Not reject

Tests of weak exogeneity restrictions:

Hypothesized No. of CE(s)	Restricted Log-likelihood	LR Statistic	Degrees of Freedom	Probability
1	-77.931	6.821	2	0.0330

Restricted cointegrating vector (standard error in parentheses)

<i>CAPBrt</i>	<i>DEBrt</i>	<i>GAPrt</i>
1.000	-0.0406	-0.836
(-)	(0.006)	(0.328)

Loading parameters (standard error in parentheses)

Δ <i>CAPBrt</i>	-0.620 (0.137)
Δ <i>DEBrt</i>	0.000 (-)
Δ <i>GAPrt</i>	0.000 (-)

Appendix A6 – Pooled-slope core FPRF estimates across estimators

	FE	LSDVc	GMM-dif	GMM-lev	GMM-sys
L.CAPBrt	-0.433*** (0.0726)	-0.344*** (0.0706)	-0.511*** (0.1071)	-0.424*** (0.0749)	-0.517*** (0.0977)
L.GAPrt	-0.016 (0.0972)	-0.019 (0.0491)	-0.051 (0.0734)	-0.024 (0.0933)	-0.016 (0.0713)
L.DEBrt	0.042*** (0.0112)	0.036*** (0.0108)	0.053*** (0.0197)	0.026** (0.0130)	0.042** (0.0181)
N	11	11	11	11	11
T	20	20	19	20	20
NT	220	220	209	220	220

Robust standard errors in brackets. * p<0.10, ** p<0.05, *** p<0.01

Appendix A7 – Extended FPRF estimates across approaches, measures and indicators

The robustness of the findings reported in Table 6 of the main text is first assessed here by using alternative measures for the debt and the output gap, and a different timing of the spread.

The second column in Table A7.1 reports the results from the use of debt-GDP data as published by the European Commission instead of those from the data published by OECD as “General government gross financial liabilities” (reported in column (1) that corresponds to column (3) GMM-dif estimates in Table 6).³⁷ The third column of Table A7.1 reports the results from the use of one-step ahead forecast of the gap at the time the policy is set instead of those from the actual gap known at the same time (reported in column (1) that corresponds to column (3) of Table 6). Finally, the fourth column of Table A7.1 reports the results from the use of the spread of the month of February instead of May. Note that the first four columns in Table A7.1 use GMM-dif and the other four use GMM-sys. The strong significance of initial conditions of balance and debt is still valid independently of the proxy-measures used for them, and this conclusion is clearly confirmed by both GMM approaches. The GMM-dif effect of gap is slightly more counter-cyclical compared to GMM-sys, and tends to reach its maximum intensity when debt is measured with EU data. The expansive effect of lagged elections is robust, and the same occurs for the evidence of not significant elections’ simultaneous effects. Finally, the September level of the spread is always positive and strongly significant, higher if specified together with that of May (which is significantly negative), while if it

³⁷ For an explanation of the different definitions of government debt which are in use see Hartwig Lojsch et al (2011).

is combined with e.g. (not significant) February, its parameter is lower. Overall, these results suggest a long lasting effect of spread level of about 0.3% of GDP per 100 basis points, against an impact effect higher if counterbalanced by the negative effect played by the spread of May. An increase of one standard error in fiscal rules positively impacts on fiscal policy, and the effect is stronger and more significant with GMM-sys. The reason for the restrictive view of the rules is related with the policies historically induced by the Stability Pact around 2010, when a number of countries were outside the Treaty thresholds.

Albeit relevant, it could be that the fiscal rules index (*fri*) has not only a direct impact on fiscal policy but also an indirect one, through its interactions with the state of other policy determinants. To address this, Table A7.2 explicitly inspects the role of such interactions between *fri* and the primary balance, debt and the output gap. In this extended context, the significant role of primary balance, debt, and spreads, and the a-cyclical nature of the output gap are confirmed. Regarding interactions, it seems that an increase in the fiscal policy rules number and enforcement induces a significant drop in the effect played by the initial budgetary conditions, while the other interactions are never significant. However, despite the significance of this effect of rules, the columns of Table A7.2 labelled as (4) report the GMM-dif and GMM-sys estimates of the parameters of a policy reaction function where any role played by the fiscal rules is ignored, and suggest that this exclusion does not substantially affect the findings regarding the estimates of the other determinants. This stability of the reaction function core parameters estimates and significance suggests that the relevant institutional changes occurred over the sample did not break the relationships among the main indicators, and supports the evidence of cointegration emerging in the first part of this paper, i.e. of stable level relationships.

A final assessment regards the robustness of the results above to the use of the alternative measures of fiscal stance indicator listed in Appendix A1. In Table 6 (and in Tables A7.1-A7.2 too) the fiscal stance is always measured by the change in the real time *capb* data (labelled as dCAPBrt), while the availability of other three measures rather than dCAPBrt enables the alternative estimates reported along the columns of Table A7.3 that reports alternative GMM-dif and GMM-sys estimates of the fiscal policy reaction function parameters where *fpi* is also measured by planned (forecast) changes in the primary borrowing (PCAPBrt), by real time changes of *capb* data (DCAPBrt), and by changes of *capb* latest available data (dCAPB), i.e. after a number of data revisions occurred.

The robustness of the effects of the fiscal initial conditions (*capb* and *deb*) and the significance of the additional variables (mainly *spreads*) is evident, although *spreads*' timing is *fpi* specific. Albeit

with different timings (as sometimes the fiscal loosening occurs in the year of elections, and in other cases the year after the elections), regular elections affect policies only in the implementation phases, as they are completely disregarded when making plans, i.e. using PCAPBrt.

The mild counter-cyclicalities of the policies, that we noted measuring *fpi* with the changes in real time *capb*, disappears with the other indicators and a-cyclicalities emerge. The latter outcome confirms the findings in Golinelli and Momigliano (2009) with pre-crisis data, and might suggest possible asymmetries of the stance cyclicalities that we will inspect in future research.

Table A7.1 – Robustness to alternative: output gap and debt measures, and spread timing (both GMM-dif and GMM-sys estimates)

	GMM-dif				GMM-sys			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
L.CAPBrt	-0.451*** (0.0626)	-0.500*** (0.0756)	-0.478*** (0.0692)	-0.455*** (0.0652)	-0.439*** (0.0791)	-0.465*** (0.0837)	-0.447*** (0.0793)	-0.445*** (0.0762)
L.GAPrt (effective)	0.080 (0.0732)	0.135* (0.0800)		0.092 (0.0724)	0.031 (0.0535)	0.056 (0.0551)		0.046 (0.0558)
L.FGAPrt (predicted)			0.079 (0.1103)				0.045 (0.0997)	
L.DEBrt (OECD debt)	0.040*** (0.0102)		0.039*** (0.0072)	0.040*** (0.0100)	0.026** (0.0104)		0.026** (0.0102)	0.027** (0.0107)
L.debue (EU debt)		0.047*** (0.0104)				0.028** (0.0109)		
L.FRI	0.115 (0.1360)	0.162 (0.1758)	0.138 (0.1452)	0.116 (0.1359)	0.340** (0.1619)	0.379** (0.1918)	0.346** (0.1691)	0.296* (0.1524)
elections	-0.119 (0.1256)	-0.031 (0.1429)	-0.141 (0.1557)	-0.120 (0.1274)	-0.053 (0.1315)	0.014 (0.1338)	-0.075 (0.1603)	-0.054 (0.1376)
L. elections	-0.315** (0.1416)	-0.369** (0.1584)	-0.403*** (0.1545)	-0.326** (0.1371)	-0.238 (0.1516)	-0.294** (0.1467)	-0.315** (0.1594)	-0.254* (0.1518)
L.spread_m2 (Feb)				-0.091 (0.1011)				-0.120 (0.1178)
L.spread_m5 (May)	-0.291** (0.1322)	-0.204 (0.1582)	-0.287** (0.1328)		-0.317*** (0.1536)	-0.260 (0.1771)	-0.305** (0.1503)	
L.spread_m9 (Sept)	0.547*** (0.1637)	0.420** (0.1701)	0.557*** (0.2066)	0.339*** (0.1304)	0.529*** (0.1725)	0.438** (0.1879)	0.532** (0.2152)	0.330** (0.1388)
NT	209	187	187	209	220	198	198	220
ar1p	0.132	0.131	0.134	0.136	0.114	0.118	0.120	0.123
ar2p	0.295	0.294	0.295	0.298	0.295	0.294	0.293	0.301
ar3p	0.377	0.429	0.354	0.342	0.392	0.404	0.360	0.361
Sargan test p-value	0.3383	0.3273	0.3544	0.3526	0.3032	0.4550	0.5396	0.3023
Time Dummies test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2	0.435	0.422	0.448	0.434	0.476	0.477	0.475	0.472

* p<0.10, ** p<0.05, *** p<0.01

Table A7.2 – The analysis of alternative specifications of the effect played by the Fiscal Rules Indicator (FRI)

	GMM-dif				GMM-sys			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
L.CAPBrt	-0.441*** (0.0605)	-0.435*** (0.0609)	-0.434*** (0.0577)	-0.445*** (0.0599)	-0.411*** (0.0534)	-0.413*** (0.0661)	-0.413*** (0.0659)	-0.423*** (0.0726)
L.GAPrt	0.095 (0.0914)	0.101 (0.0790)	0.092 (0.0688)	0.084 (0.0710)	0.048 (0.0679)	0.050 (0.0517)	0.050 (0.0471)	0.024 (0.0491)
L.DEBrt	0.038*** (0.0094)	0.039*** (0.0117)	0.039*** (0.0111)	0.042*** (0.0104)	0.025*** (0.0084)	0.025*** (0.0096)	0.025** (0.0102)	0.024** (0.0110)
L.FRI	0.597* (0.3426)	-0.050 (0.1854)	0.052 (0.1349)		0.962* (0.5396)	0.247*** (0.0954)	0.252 (0.1632)	
L.CAPB*L.FRI	0.102*** (0.0266)	0.052** (0.0220)	0.045* (0.0263)		0.123*** (0.0289)	0.071** (0.0312)	0.070** (0.0324)	
L.GAP*L.FRI	-0.097 (0.0888)	-0.028 (0.0441)			-0.079 (0.0873)	-0.002 (0.0324)		
L.DEB*L.FRI	-0.011 (0.0075)				-0.012 (0.0092)			
elections	-0.176 (0.1361)	-0.131 (0.1145)	-0.134 (0.1219)	-0.130 (0.1237)	-0.147 (0.1492)	-0.087 (0.1222)	-0.087 (0.1247)	-0.066 (0.1290)
L. elections	-0.466*** (0.1620)	-0.340** (0.1495)	-0.345** (0.1419)	-0.323** (0.1397)	-0.425** (0.1669)	-0.289* (0.1630)	-0.290* (0.1622)	-0.257* (0.1452)
L.spread_m5	-0.287** (0.1246)	-0.313** (0.1391)	-0.289** (0.1232)	-0.288** (0.1392)	-0.279* (0.1433)	-0.320** (0.1577)	-0.319** (0.1392)	-0.291* (0.1707)
L.spread_m9	0.536*** (0.1777)	0.573*** (0.1861)	0.537*** (0.1560)	0.545*** (0.1683)	0.457*** (0.1584)	0.521*** (0.1768)	0.519*** (0.1560)	0.484*** (0.1812)
NT	187	209	209	209	198	220	220	220
ar1p	0.132	0.133	0.138	0.134	0.125	0.121	0.122	0.122
ar2p	0.299	0.308	0.297	0.296	0.307	0.313	0.296	0.301
ar3p	0.373	0.413	0.376	0.387	0.403	0.449	0.384	0.447
Sargan test p-value	0.2778	0.2784	0.3038	0.3728	0.4788	0.2653	0.2784	0.2950
Time Dummies test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2	0.460	0.438	0.445	0.420	0.498	0.487	0.487	0.467

* p<0.10, ** p<0.05, *** p<0.01

Table A7.3 – Estimation results from the use of alternative measures of fiscal policy indicators

	GMM-dif				GMM-sys			
	(1) dCAPBrt	(2) PCAPBrt	(3) DCAPBrt	(4) dCAPB	(1) dCAPBrt	(2) PCAPBrt	(3) DCAPBrt	(4) dCAPB
L.CAPBrt	-0.451*** (0.0626)	-0.533*** (0.0862)	-0.457*** (0.0451)	-0.422*** (0.0489)	-0.439*** (0.0791)	-0.515*** (0.0992)	-0.448*** (0.0588)	-0.420*** (0.0637)
L.GAPrt	0.080 (0.0732)	-0.059** (0.0277)	-0.085 (0.1134)	-0.078 (0.1385)	0.031 (0.0535)	-0.059** (0.0236)	-0.104 (0.0889)	-0.132 (0.1202)
L.DEBrt	0.040*** (0.0102)	0.020*** (0.0065)	0.036*** (0.0098)	0.040*** (0.0113)	0.026** (0.0104)	0.021*** (0.0053)	0.030*** (0.0087)	0.025*** (0.0096)
L.FRI	0.115 (0.1360)	0.237** (0.1111)	0.108 (0.1092)	0.111 (0.1503)	0.340** (0.1619)	0.136* (0.0828)	0.132 (0.1353)	0.261* (0.1386)
regular	-0.119 (0.1256)	-0.031 (0.0782)	-0.316** (0.1514)	-0.631*** (0.2012)	-0.053 (0.1315)	0.012 (0.0859)	-0.212 (0.1617)	-0.544*** (0.2005)
L.regular	-0.315** (0.1416)	0.136 (0.0920)	-0.100 (0.1781)	0.049 (0.2065)	-0.238 (0.1516)	0.172* (0.1012)	-0.007 (0.1874)	0.099 (0.2191)
L.spread_m5	-0.291** (0.1322)	-0.410** (0.1654)	-0.178 (0.1216)	-0.704*** (0.1084)	-0.317** (0.1536)	-0.411** (0.1718)	-0.204 (0.1352)	-0.723*** (0.1055)
L.spread_m9	0.547*** (0.1637)	0.577*** (0.2082)	0.389* (0.2020)	0.813*** (0.1568)	0.529*** (0.1725)	0.581*** (0.2140)	0.452** (0.2095)	0.801*** (0.1561)
NT	209	209	209	209	220	220	220	220
ar1p	0.132	0.058	0.121	0.050	0.114	0.055	0.110	0.044
ar2p	0.295	0.648	0.142	0.223	0.295	0.610	0.159	0.357
ar3p	0.377	0.383	0.148	0.465	0.392	0.360	0.142	0.386
Sargan test p-value	0.3383	0.0439	0.4383	0.5051	0.3032	0.0044	0.5439	0.6158
Time Dummies test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R2	0.435	0.722	0.486	0.343	0.476	0.725	0.492	0.367

* p<0.10, ** p<0.05, *** p<0.01