Assessing non-linearities and heterogeneity in debt sustainability analysis: A panel spline approach

Benjamin Owusu   **Bettina Bökemeier**   **Alfred Greiner**

Assessing non-linearities and heterogeneity in debt sustainability analysis: A panel spline approach

Benjamin Owusu[a]   Bettina Bökemeier[b]   Alfred Greiner[c]

Abstract

This paper empirically studies public debt sustainability with the penalized panel splines approach for 25 EU economies from 2000 to 2019 by estimating the response of the primary surplus to lagged debt relative to GDP, respectively. A positive coefficient on average indicates sustainable policies, which is supported by all our results. Moreover, we show that this relationship is not homogeneous across the distribution of the debt ratios, but, varies with the magnitude of public debt to GDP. The estimations reveal a strongly increasing reaction for small and high debt ratios while it is rather flat for intermediate levels. This holds for normal times, too, whereas during years of economic crises a monotonously increasing response can be observed. Additionally, for a cluster consisting of smaller EU economies, there is indication of 'fiscal fatigue', meaning that the effort of active fiscal counter-steering peters out for high ratios of public debt to GDP. The same effect can be observed for the whole sample and for a sample including the large EU economies, once Greece is removed.

---

[a]Department of Business Administration and Economics, Bielefeld University, P.O. Box 100131, 33501 Bielefeld, Germany. E-mail: benjamin.owusu@uni-bielefeld.de

[b]Department of Business Administration and Economics, Bielefeld University, P.O. Box 100131, 33501 Bielefeld, Germany. Phone: E-mail: bboekemeier@wiwi.uni-bielefeld.de

[c]Department of Business Administration and Economics, Bielefeld University, P.O. Box 100131, 33501 Bielefeld, Germany, E-mail: agreiner@uni-bielefeld.de

1   Introduction

With the 2008/2010 European debt crisis the interest in public debt sustainability analyses has revived. Further, the current Covid-19 pandemic affecting countries around the world forces governments to react in order to protect the societies and to cushion the negative economic effects with the help of huge rescue and recovery programs that are mainly financed by issuing bonds. Thus, these policies affect both the current and the future budget as well. Therefore, studying public debt sustainability and fiscal discretionary policies is as urgent as ever.

Based on the seminal contributions by Bohn (1995, 1998) we assess public debt sustainability by analyzing the reaction of the primary budget to changes in public debt relative to the gross domestic product (GDP), respectively. If a government reacts to a rise in the public debt ratio by actively adjusting its discretionary fiscal policy in terms of higher primary surpluses, the debt policy is considered to be sustainable. Usually, the relationship is empirically tested in a single equation linear regression model. However, standard econometric specifications, assuming linearity between a response variable and explanatory variables, could lead to model misspecification if the true data generating process was non-linear and, hence, could lead to wrong inferences. Therefore, it seems to be necessary to resort to non-linear estimation techniques.

The need to increase the size of the data sets to improve statistical inference and to study the dynamic relationships between variables has made panel studies popular in applied macroeconomics research. The basic specification in panel studies is the fixed effects model, where the unobserved time-invariant individual heterogeneities are allowed to correlate with the regressors. That model assumes a linear relationship between the response variable and the co-variates and allows to measure the marginal effects of the latter on the first.

In this paper, however, we intend to get additional insight into the data generating process and we go beyond the standard linear fixed effects model that is commonly used in panel debt sustainability analyses. Thus, we question the hypothesis that the response of the discretionary fiscal policy is uniform across the distribution of the debt ratios, but, we allow it to vary with the magnitude of the debt ratios. Particularly, we can show that the reaction of the structural primary balance changes with the size of the debt ratio, implying that in situations with low debt ratios the response coefficient is expected to

be different as compared to situations with high debt ratios. In addition, we consider two different clusters of EU economies to allow for country specifics and we distinguish between times of economic crises and normal times, where we use AMECO data for 25 EU economies from 2000 to 2019. The estimations demonstrate that the reaction to higher debt ratios differs in the two clusters and that it is higher in times of crises than in normal times. Thus, our estimations provide support for the hypothesis of heterogeneity in the data.

Statistical testing of whether the inter-temporal budget constraint of a government is fulfilled began with the paper by Hamilton and Flavin (1986). Those authors tested whether the present-value borrowing constraint of the US federal government holds with annual data from 1960 to 1984. They did so by analyzing if the time series of the US federal public debt contains a bubble term that would indicate that public debt exceeded the present value of expected future primary surpluses, implying an unsustainable debt policy. Hamilton and Flavin performed several tests and found evidence for sustainability of the US federal debt policy.

The test procedure applied by Hamilton and Flavin has been criticized by Wilcox (1989) because their test does not allow for a stochastic interest rate. Therefore, he proposes a different test where the discounted time series of public debt should be analyzed and if that series converges to zero, sustainability of public debt would be given. Applying that test to the same time series Hamilton and Flavin used, Wilcox finds evidence that the US federal debt is unsustainable.

The result obtained by applying the test proposed by Wilcox heavily depends on the interest rate with which the series of public debt is discounted. Since this is a random variable and realizations of that variable in the past do not give information about future interest rates, it has been argued that tests should be resorted to that yield outcomes that are independent of the interest rate. Hakkio and Rush (1991), for example, propose to test for cointegration of revenues and expenditures. The idea behind that test that a cointegrating relation between spending and revenues implies a stationary first difference and, thus, a sustainable debt policy for a positive interest rate.

Another test that does not rely on the interest rate is the one proposed by Trehan and Walsh (1991). They test whether public deficits inclusive of interest payments grow at most linearly. If that property is fulfilled a given series of public debt is sustainable because

any time series that grows linearly converges to zero if it is exponentially discounted, provided the real interest rate is positive. Denoting by B public debt and by r the interest rate, another test proposed by Trehan and Walsh (1991) is to analyze whether a quasi-difference of public debt, $B_t - \lambda B_{t-1}$ with $0 \leq \lambda < 1 + r$, is stationary and whether public debt and primary surpluses are cointegrated. If government debt is quasi-difference stationary and public debt and primary surpluses are cointegrated, public debt is sustainable.

The dependence of sustainability tests on the interest rate has been heavily criticized by Bohn (1995, 1998), too, since future realizations of that random variable are unknown. Therefore, he proposes to test whether the primary surplus relative to GDP is a positive function of the debt to GDP ratio that rises at least linearly with higher debt to GDP ratios. The intuition behind that test is that a positive reaction of the primary surplus to higher debt relative to GDP, respectively, implies mean reversion of the debt to GDP ratio and a stationary debt to GDP ratio is sustainable in a growing economy. This test is very plausible because it has a nice economic intuition: if governments run into debt today they have to take corrective actions in the future by increasing the primary surplus. Otherwise, public debt will not be sustainable. Many applications of such a fiscal response function as suggested by Bohn have followed. A nice recent overview can be found for instance in Beqiraj et al. (2018).

The rest of this paper is organized as follows. Section 2 briefly discusses the theoretical background and section 3 presents the estimation method and the outcomes. Finally, section 4 summarizes the main results and concludes. Additional information about the empirics can be found in the appendix.

2   Theoretical Background

To get deeper insight into the logic of the test proposed by Bohn (1995, 1998) and the fiscal response function, we consider the accounting identity describing the accumulation of public debt in continuous time that is given by the following differential equation:

$$\dot{B}(t) = r(t)B(t) - S(t), \tag{1}$$

with $B(t) > 0$ real public debt at time t, $r(t) > 0$ the real interest rate, $S(t) \in \mathbb{R}$ the real government surplus exclusive of interest payments and the dot over a variable stands for

the derivative with respect to time d/dt. A government is said to follow a sustainable debt policy if its inter-temporal budget constraint is fulfilled, i.e. if the present value of public debt converges to zero asymptotically. The latter means that $\lim_{t\to\infty} e^{-C_1(t)} B(t) = 0$ holds, with $C_1(t) = \int_{t_0}^t r(\mu)d\mu$ the discount rate.

Assuming that the government in the economy determines the primary surplus to GDP ratio, $s(t) = S(t)/Y(t)$, such that it is a positive linear function of the debt to GDP ratio, $b(t) = B(t)/Y(t)$, and of a term that is independent of public debt, $\phi(t)$ (see Bohn, 1995, 1998, Greiner and Fincke, 2015, Afonso and Jalles, 2019), the primary surplus ratio can be written as

$$s(t) = \psi(t)\,b(t) + \phi(t), \tag{2}$$

where $\psi(t)$ is the reaction or response coefficient determining the change of the primary surplus to variation in public debt, relative to GDP respectively, and that is a continuous function of time. The parameter $\phi(t) \in \mathbb{R}$ is affected by other economic variables, such as social spending or transitory government expenditures in general. As regards $\phi(t)$ we posit that it is bounded from above and from below by a certain finite number that is constant over time. We should also like to emphasize that $\phi(t)$ cannot be completely controlled by the government. The government can influence that coefficient to a certain degree but it has not complete control over it because $\phi(t)$ is also affected by the business cycle for example that can affect temporary government outlays.

Using (2) the period budget constraint of the government (1) is obtained as

$$\dot{B}(t) = (r(t) - \psi(t))\,B(t) - \phi(t)\,Y(t). \tag{3}$$

Integrating equation (3) and multiplying both sides by $e^{-C_1(t)}$ leads to

$$e^{-C_1(t)} B(t) = e^{-C_3(t)}\, B(t_0) - e^{-C_3(t)} \int_{t_0}^t e^{-C_1(\mu)+C_2(\mu)+C_3(\mu)} Y(t_0) \phi(\mu)d\mu, \tag{4}$$

with $g(t)$ the growth rate of GDP and $C_1(\mu) = \int_{t_0}^\mu r(\nu)d\nu$, $C_2(\mu) = \int_{t_0}^\mu g(\nu)d\nu$, $C_3(\mu) = \int_{t_0}^\mu \psi(\nu)d\nu$. As regards the interest rate, we posit that the interest rate on government bonds exceeds the growth rate of GDP on average, $\int r(\mu)d\mu > \int g(\mu)d\mu$. We make this assumption because otherwise the inter-temporal budget constraint would not pose a problem for the government since it can grow out of debt in that case.

Now, assume that $\psi(t)$ is positive on average so that $\lim_{t\to\infty} \int_{t_0}^t \psi(\mu)d\mu = \lim_{t\to\infty} C_3(t) = \infty$ holds. Then, we get $\lim_{t\to\infty} e^{-C_3(t)} B(t_0) = 0$ and the first term in equation (4) converges to zero.

4

Since $|\phi(t)| < \infty$ we can set $\phi(\mu)Y(t_0) = 1$ so that it is sufficient to consider for the second term on the right hand side in (4) the following expression,

$$\Omega(t) = \frac{\int_{t_0}^t e^{-C_1(\mu)+C_2(\mu)+C_3(\mu)} d\mu}{e^{C_3(t)}}.$$

If $\int_0^\infty e^{-C_1(\mu)+C_2(\mu)+C_3(\mu)} d\mu$ remains bounded $\lim_{t\to\infty} C_3(t) = \infty$ guarantees that Ω converges to zero. In case of $\lim_{t\to\infty} \int_{t_0}^t e^{-C_1(\mu)+C_2(\mu)+C_3(\mu)} d\mu = \infty$, the limit of Ω is obtained by applying l'Hôpital as

$$\lim_{t\to\infty} \Omega(t) = \lim_{t\to\infty} \frac{e^{-C_1(t)+C_2(t)}}{\psi(t)}.$$

Since $r(t)$ and $g(t)$ are finite values, we can find a constant $\kappa > 0$ such that $\Omega \le e^{-\kappa t}/\psi(t)$. The right hand side in the former inequality does not converge to zero if $\psi(t)$ converged to zero exponentially. However, in that case $\lim_{t\to\infty} \int_{t_0}^t \psi(\mu) d\mu < \infty$ would hold. Consequently, in case that $\lim_{t\to\infty} \int_{t_0}^t \psi(\mu) d\mu = \infty$ holds, $\psi(t)$ cannot decline exponentially, and $\Omega(t)$ converges to zero.

These considerations demonstrate that a strictly positive reaction coefficient on average, such that $\lim_{t\to\infty} \int_{t_0}^t \psi(\mu) d\mu = \infty$, implies that the debt policy of a government is sustainable. It must be pointed out that the reaction coefficient $\psi(t)$ can be time-varying and it may even be negative for some time periods. However, on average that coefficient must be positive.

A shortcoming of the former analysis is that it implicitly assumes that the primary surplus can grow without upper bound. However, a positive but small reaction coefficient on average does not necessarily guarantee a bounded debt to GDP ratio. To see this, we note that the debt to GDP ratio, $b(t) = B(t)/Y(t)$, evolves according to

$$\frac{\dot{b}}{b} = \frac{\dot{B}}{B} - \frac{\dot{Y}}{Y} \tag{5}$$

that, using equation (2), can be written as

$$\dot{b}(t) = (r(t) - \psi(t) - g(t))\, b(t) - \phi(t). \tag{6}$$

Integrating (6) we obtain for the debt ratio

$$b(t) = e^{(C_1(t)-C_2(t)-C_3(t))} b(t_0) - e^{(C_1(t)-C_2(t)-C_3(t))} \int_0^t e^{-(C_1(\mu)-C_2(\mu)-C_3(\mu))} \phi(\mu) d\mu. \tag{7}$$

Equation (7) demonstrates that the public debt to GDP ratio $b(t)$ remains bounded for $\int_{t_0}^t \psi(\mu) d\mu > \int_{t_0}^t (r(\mu) - g(\mu))\, d\mu$, while it diverges to infinity in the case of $\int_0^t \psi(\mu) d\mu \le$

$\int_{t_0}^{t} (r(\mu) - g(\mu)) \, d\mu$. Thus, the reaction coefficient must exceed the difference between the interest rate on public debt and the GDP growth rate on average such that the debt to GDP ratio remains bounded.

Here, we should like to stress two aspects. First, equation (7) shows the significance of the difference $\int r(\mu) d\mu - \int g(\mu) d\mu$. If the GDP growth rate exceeds the interest rate on public debt, the debt to GDP ratio remains bounded even for $\int \psi(\mu) d\mu = 0$. In that case, the government can grow out of debt as already mentioned above. Second, a positive reaction coefficient that falls short of the difference between the interest rate and the GDP growth rate implies a rising debt to GDP ratio, if the interest rate exceeds the GDP growth rate. But, such a policy is not sustainable because it would require permanently rising primary surplus to GDP ratios which is not feasible.

To see the second point more clearly we note that the primary surplus relative to GDP is bounded from above because the primary surplus can never exceed GDP. Denote by $\bar{\tau} < 1$ the upper bound of the primary surplus relative to GDP and assume that the governments sets it equal to that value for ever, i.e. $s(t) = \bar{\tau}$. Then, the level of public debt evolves according to the following differential equation,

$$\dot{B}(t) = r(t) \, B(t) - \bar{\tau} \, Y(t). \tag{8}$$

Integrating that equation and multiplying both sides by $e^{-C_1(t)}$ to get present values, we obtain,

$$e^{-C_1(t)} B(t) = B(t_0) - \bar{\tau} Y(t_0) \int_{t_0}^{t} e^{-(C_1(\mu) - C_2(\mu))} d\mu. \tag{9}$$

The inter-temporal budget constraint is fulfilled for $\lim_{t \to \infty} e^{-C_1(t)} B(t) = 0$ which implies $b(t_0) = \bar{\tau} \int_{t_0}^{\infty} e^{-(C_1(\mu) - C_2(\mu))} d\mu \equiv b_{crit}$ which is the critical threshold value. If the initial debt to GDP ratio is larger than b_{crit}, sustainability of public debt is excluded. In that case the initial outstanding debt ratio is too large to be covered by the sum of discounted future primary surpluses.

To finish our theoretical considerations, we want to point out that we allowed the reaction coefficient to be time-varying. And in fact, times series analyses provide strong empirical evidence that this coefficient is not constant over time (see Greiner and Fincke, 2015, chap. 2.2-2.5). It should be noted, too, that a linear model with time-varying coefficients can be seen as an approximation of a non-linear model and the approximation is good if the parameter changes smoothly (cf. Granger, 2008). We emphasize this point

because in our empirical part we estimate a non-linear model describing the relationship between the primary surplus and public debt. More concretely, our model is a semi-parametric model and the estimation outcome consists of the average coefficient and of a smooth function such that the actual coefficient for each observation just consists of the average value plus the non-linear part given in the figure.

The use of a non-linear model is of interest because the question arises which factors are responsible for variations in the coefficients. When analyzing the response of the primary surplus to variations in public debt, there is evidence that the reaction depends on the magnitude of the public debt ratio. For example, Ghosh et al. (2013) have analyzed 23 advanced economies over the period 1970-2007 and found that the reaction decreases once a critical value has been passed, what they refer to as 'fiscal fatigue'. Checcherita-Westphal and Zdarek (2017) also explore on 'fiscal fatigue' in euro-area economies with a focus on the primary balance.

3   Empirics

3.1   Estimation Method

Regarding the methodology for the empirical estimation, we resort to the penalized splines fixed effects estimator according to Puetz and Kneib (2016). Such additive semi-parametric models have become increasingly popular in empirical works (Baltagi and Li, 2002). We argue that this approach is desirable because non-parametric modeling does not impose restrictions regarding the functional relationship between the response variables and the regressors. Rather the functional shape of the co-variate effects are derived from the datasets. Pioneering works on penalized splines can be traced back to Hastie and Tibshirani (1990) with their introduction of generalized additive models which provide a flexible way of modeling the response of a dependent variate to co-variates. Subsequent contributions were made by Wood (2000) who introduced the mixed model representation of penalized splines, as well as Ruppert et al. (2003), Eilers and Max (1996) and Kauermann (2005).

In panel data settings, the correlation between co-variates and the unobserved time-invariant factors in the error term is prevalent. Baltagi and Li (2002), Su and Ullah (2006) and Henderson et al. (2008) have all proposed alternative semi-parametric fixed effects

estimators where the non-linearity is modeled via a kernel estimator. However, Puetz and Kneib (2016) argued that modeling fixed effects and non-linearity with the Kernel estimator is computationally demanding especially with large datasets. Penalized splines on the other hand provide a flexible and convenient way of modeling the non-linearity without complications and the fixed effects in the model can be conveniently handled by way of a first difference operator.

Following Puetz and Kneib (2016), we consider the following panel additive model

$$y_{i,t} = \mu_i + \sum_{g=1}^{p} f_g(x_{gi,t}) + v_{i,t} \tag{10}$$

where $i = 1, ..., N$ represent the individual countries and $t = 1, ...T$ represent the time period, $y_{i,t}$ is the response variable, whilst μ_i accounts for the time invariant unobserved heterogeneity also known as the individual fixed effects in this case. By this, we allow the unobserved heterogeneity to correlate with the regressors instead of the error term. The variable $v_{i,t}$ represents the error term which is assumed to be normally distributed with zero mean and a constant variance. The functions $f_1(x_{gi,t}), ..., f_p(x_{gi,t})$ are the non-linear effects of the p continuous co-variates which can be approximated by a penalized B-splines according to Eilers and Marx (1996) together with a penalty term applied to penalized least squares criterion. The aim of the penalty term is to penalize too much variability in the model as a way of regularizing the estimation in order to deal with overfitting to the data.

Assuming that (10) holds for each point, we could express the lag as

$$y_{i,t-1} = \mu_i + \sum_{g=1}^{p} f_g(x_{gi,t-1}) + v_{i,t-1} \tag{11}$$

in order to eliminate the fixed effects, we subtract (11) from (10) to obtain

$$\Delta y_{i,t} = \sum_{g=1}^{p} [f_g(x_{gi,t}) - f_g(x_{gi,t-1})] + v_{i,t} - v_{i,t-1} \tag{12}$$

where Δ is the difference operator. The non linear function f_g is approximated by the weighted sum of d_g B-spline basis functions $B_{g1}, ..., B_{gd}$ such that

$$f_g(x_{gi,t}) = \sum_{g=1}^{d_g} B_{gj}(x_{gi,t})\beta_{gi,t} = z_g^T(x_{gi,t})\beta_g. \tag{13}$$

Inserting (13) into (12) and applying the difference operator once again, we have

$$\Delta y_{i,t} = \sum_{g=1}^{p} [\Delta z_g(x_{gi,t})]^T \beta_g + \Delta v_{i,t}. \tag{14}$$

For the sake of simplicity, (14) can be written in compact matrix notation as

$$\Delta y = \sum_{g=1}^{p} \Delta z_g \beta_g + \Delta v \tag{15}$$

where $\Delta y = (y_{12} - y_{11}, ..., y_{N,T} - y_{N,T-1}^T)$ is a column vector with dimension $N(T-1)$ and Δz_g is derived by taking the difference between z_g and its lags which is of dimension $N(T-1)Xd_h$.

In order to obtain a first difference penalized spline estimator, we minimize the penalized least square criterion below

$$[\Delta y - \sum_{g=1}^{p} (\Delta z_g) \beta_g]^T [\Delta y - \sum_{g=1}^{p} (\Delta z_g) \beta_g] + \sum_{g=1}^{p} \lambda_g \beta_g^T k_g \beta_g \tag{16}$$

where the matrix k is introduced and assigned to each smooth function to penalize too much variability of the adjacent coefficient in the vector β_g. This prevents overfitting the data to the model. λ_g is the smoothing parameter which is the weight placed on the penalty term in the minimization criterion in (16). In reality, the smoothing parameters are unknown ex-ante. However, they can be estimated via maximum likelihood estimation by making use of the mixed model representation of the penalized splines. To estimate the smooth function f_g, equation (10) is identified such that $\sum_{i=1}^{N} \sum_{t=1}^{T} f_g(x_{gi,t}) = z_g \beta_g = 0$, for all $g = 1, \ldots, p$. In order to account for serial correlation in the residuals which could arise as a result of first difference transformation, a Generalized Least Square (GLS) approach is used where the differenced model matrix (Δz_g) and Δy in (16) are multiplied by a block diagonal matrix.[1] This ensures that the estimator is robust against serial correlation in the residuals. Another appealing feature of the semi-parametric fixed effects estimator is the estimation of the derivative of the smooth function as well as the computation of simultaneous confidence bands for inferences. See Wiesenfarth et al. (2012) and Puetz and Kneib (2016) for details regarding the simultaneous confidence bands.

---

[1]See Puetz and Kneib (2016) for further details.

3.2 Empirical results

In this paper, we model a fiscal reaction function in analogy to Bohn (1998) where the primary surplus is assumed to be a linear function of debt, relative to GDP respectively, and of other macroeconomic variables which serve as control variables. In contrast to Bohn, we do not posit that the response of the primary surplus to public debt is linear, but, we estimate a semi-parametric model of the form

$$S_{i,t} = \mu_i + f(X_{i,t}) + \gamma Z_{i,t}^T + v_{i,t} \tag{17}$$

where the vector $S_{i,t}$ is the response variable, $X_{i,t}$ represents the variables which enter the model non-linearly, $Z_{i,t}$ are other variables which are modeled linearly, γ measures the impact of the linear regressors on the response variable and $v_{i,t}$ is the uncorrelated error term assumed to be centered around zero with a constant variance. We shall apply (17) to test the fiscal reaction function in this paper.

As a first step, all variables are modeled non-parametrically, meaning that we do not impose any restrictions on the relationship between the response variables and the co-variates. For instance, if the true relationship between any of the co-variates and the response variable is indeed linear, the spline model will impose a linear relationship via the effective degree of freedom of the co-variate in question.[2] This is particularly advantageous since we thus avoid misspecification of the model that would result if a priori restrictions were imposed that are inconsistent with the true data generating process.

The source of the data set used for the empirical study is the European Union AMECO website. Regarding the response variable, we use the cyclically adjusted primary balance of the governments. This represents the structural aspect of the primary balance which depicts the behavior of the policymakers since it is devoid of shocks or one-off fluctuations. It is expressed relative to GDP and denoted by PBratio with positive (negative) values indicating surpluses (deficits). The explanatory variables include the lagged debt ratio of the previous period $t-1$ which constitutes our main co-variate of interest. Following the Barro (1979) tax smoothing principle[3], we use the business cycle variable (YVAR) and

---

[2]For Inflation this turned out to be the case, so it has been included in a linear manner. The effective degree of freedom is a meaningful and scale-free measure of the complexity of the penalized spline fit (Harezlak et al. 2018).

[3]According to the tax smoothing principle, the public deficit should be used such that the tax rate is constant over time to minimize the excess burden of taxation.

the public expenditure gap (GVAR) as control variables. YVAR which is also known as the output gap is computed as the deviation of output (GDP) from its long-term trend. Similarly, we compute the GVAR as the deviation of real government spending from its long term trend. We use the HP filter to estimate the trend components of output and real spending. In order to capture the influence of international trade, we include the net exports ratio that is given by the difference between exports and imports as a ratio to GDP. In addition, we include Inflation in order to explore its influence on the primary balance as a proxy for monetary policy effects.[4] A total of 25 of the EU countries constitutes the sample with the exception of Lithuania and Croatia due to unavailability or missing data. The sample period is from 2000 to 2019, hence a total of 500 observations in annual frequency were generated.

Before we present our estimation results we give a summary statistics of the variables in table 1. It can be realized that the average primary balance is positive, although rather small, and that the two variables of interest, the primary balance ratio and the debt to GDP ratio, are left-skewed and right-skewed, respectively.

Table 1: Summary Statistics

Statistic	PBratio	Debt/GDP	GVAR	YVAR	Inflation	Net Exportratio
Mean	0.000107	0.585548	1.80E-11	-3.40E-11	0.905804	0.022806
Median	0.002366	0.540408	-0.136394	-0.255106	0.938120	0.015596
Maximum	0.097171	1.862386	52.88208	96.16156	1.166769	0.360148
Minimum	-0.276920	0.037655	-36.70577	-106.6082	0.441318	-0.206930
Std. Dev.	0.030345	0.343700	6.337278	16.20572	0.128713	0.087258
Skewness	-1.700262	0.903345	1.808153	0.221095	-1.067245	1.257040

The next table gives the outcomes of our estimations for the full sample, both for the linear and for the non-linear model.[5]

Table 2 presents the standard linear fixed effects results for model 1 through 3, with the latter giving the full model including all our control variables. It shows that the response coefficient is positive and statistically significant for all three parametric specifications, indicating sustainable debt policies. Additionally, public spending (GVAR) turns out to have a small negative influence on the structural primary balance, while the business cycle effect represented by YVAR is rather weak and significant only at the 10% level in model

---

[4]We used the GDP price deflator as a proxy for inflation with 2015 the base year.

[5]The standard errors were rounded to the third and forth decimal place, respectively.

Table 2: Estimation results - Full sample

	Panel Linear Fixed Effects			*Panel Penalized Spline*				
				Parametric				
Variables	1	2	3	4	5	6	7	8
Laggeddebtratio	0.073***	0.077***	0.078***	0.049***	0.055***	0.042***	0.042***	0.024***
	(0.017)	(0.017)	(0.017)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
GVAR	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.0015***
	(0.0001)	(0.0001)	(0.0001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
YVAR	-0.0001	-0.0001*	-0.0001*	0.00001	0.00004	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.000)	(0.000)	(0.000)	(0.0001)	(0.000)	(0.000)
Inflation		0.044	0.042		-0.022**	-0.050***	-0.049***	-0.043***
		(0.035)	(0.035)		(0.009)	(0.009)	(0.009)	(0.009)
Net exportratio			-0.024			0.154***	0.181***	0.172***
			(0.039)			(0.029)	(0.030)	(0.029)
				Non-parametric				
edf(laggeddebtratio)				4.856***	5.38***	5.102***	5.806***	6.987***
edf (GVAR)							6.995***	6.986***
edf (YVAR)							2.987**	3.399***
edf (Net exportratio)							6.360***	6.423***
Adj R^2	0.292	0.293	0.292	0.382	0.379	0.409	0.609	0.574
Observ	475	475	475	475	475	475	475	456

Standard error are indicated in parenthesis. *, ** and *** indicates statistical significance at 10%, 5% and 1%, respectively.

2 and 3. The latter result may be due to the fact that we use the cyclically adjusted primary balance. Inflation and trade did not turn out to be statistically significant.

Next, we argue that allowing for a more flexible approach, where the effects are modeled in a non-parametric manner, turns out to be more informative. The second part of table 2 with specifications 4 to 7 presents the results for the penalized spline model accounting for non-linear effects of the debt ratio, of GVAR, of YVAR and of net exports.[6] Net exports and Inflation are now statistically significantly positive and negative, respectively. The results for the debt ratio show a significantly positive average response that, however, is a little smaller compared to the outcome of the linear estimations. But the estimated degrees of freedom of the smooth term for the debt ratio with values around 5 indicate an alternating behavior across the distribution of the debt ratios and is statistically significant. The same holds for the other variables, except Inflation. The shape of the smooth functions is depicted in figure 1.

Figure 1 shows that there are indeed variations in the effects of the variables across the distribution of the explanatory variables. It should be recalled that negative (positive) values in the figure 1 imply that the actual parameter is below (above) the average and

---

[6]As mentioned above, Inflation exerts a linear effect, thus, no smooth function is needed. Hence, the model is rather a semi-parametric one.

Figure 1: Full sample - Smooth Functions
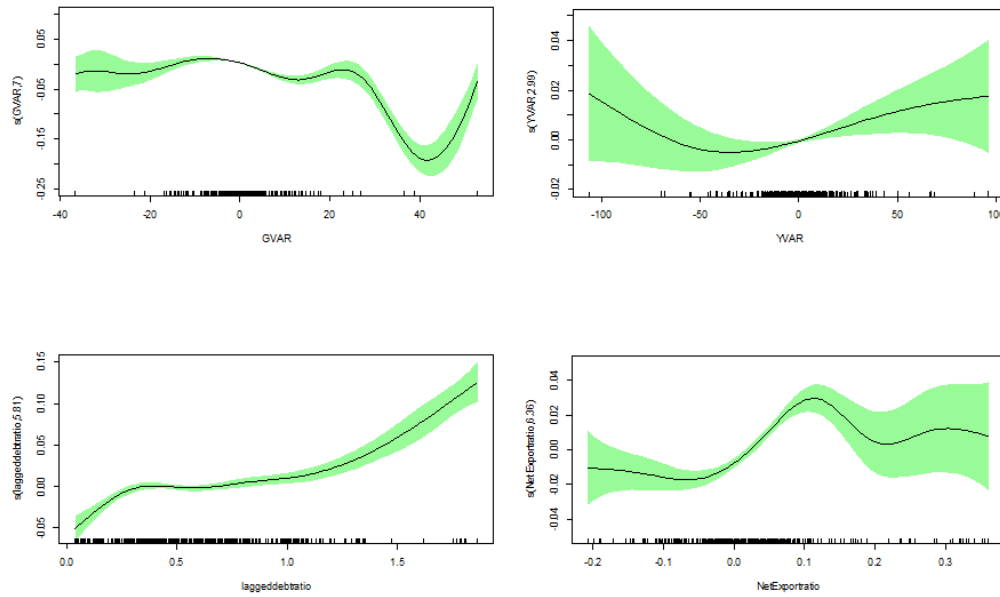
the actual coefficient just equals the estimated average plus the respective value given in the figure. For our variable of interest, the lagged debt ratio, it can be realized that there is an increase in the reaction coefficient until the debt ratio reaches about 30 percent, followed by some flat part and again a distinctive increase once the debt to GDP ratios rise above 100 percent. These information from the non-linear modeling yield additional insight compared to the estimated parameters obtained in the standard linear fixed effects model. It clearly reveals heterogeneity in the relationship.

Model 8 in table 2 presents the estimation outcome for the full sample without Greece. It can be realized that the average reaction coefficient is clearly smaller now. This demonstrates that Greece had achieved very high primary surpluses, with the help of other EU economies of the Euro-zone and of the IMF, during its consolidation after its public debt crisis. The elimination of Greece from the sample also changes the shape of the smooth function, describing the reaction to higher debt ratios. Figure 2 demonstrates that the response to higher debt ratios rises with higher debt, but, not monotonously. Rather, there are oscillations indicating phases of 'fiscal fatigue', meaning that the effort of counter-steering peters out and declines once the debt to GDP ratio exceeds certain thresholds.

All in all, our empirical analysis yields a statistically significant positive reaction coefficient on average and, thus, supports the view of sustainable debt policies in the countries under consideration. However, based on the spline estimations it reveals that there is

Figure 2: Full sample - Without Greece
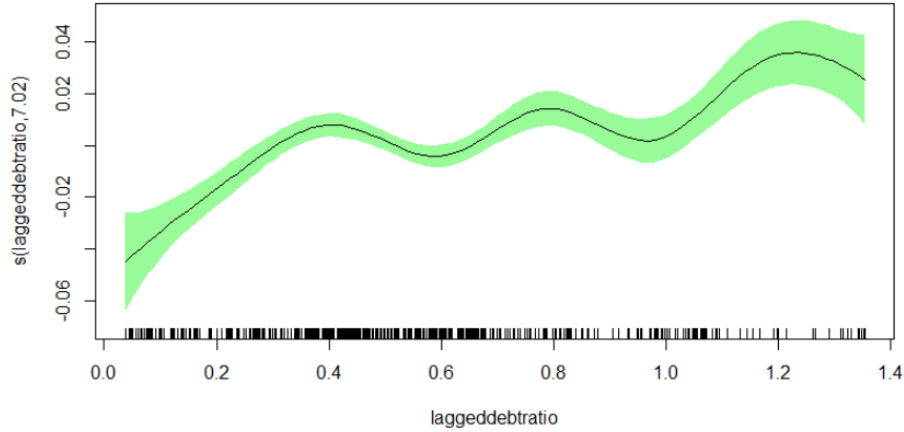
more information contained in the data. The smooth functions show that the response varies considerably across the distribution of the debt ratios. This effect is particularly pronounced for low and high values of public debt relative to GDP.

Figure 6 in the appendix depicts the residual diagnostics plot where we compare the residuals of the panel linear fixed effects model (specification 3 in table 2) against the penalized spline model (specification 7 in table 2). Since both models assume normally distributed residuals, it would be plausible to ascertain if this is indeed the case. The upper panel in figure 6 shows a quantile-quantile plot of the residuals for the two models. It can be observed that the panel spline residuals satisfy the normality assumption better than the linear fixed effects model since most of the residuals lie on the straight diagonal line between the theoretical quantile of a normal distribution and the sample quantile. Hence, the spline model performs better than the linear model. In the lower panel of figure 6, we plotted the fitted values against the residuals to explore the presence of heteroskedasticity. However, the residual plots for both models are similar and do not exhibit any pattern.

As a robustness test, we estimated another semi-parametric model by way of a panel mixed effects model according to Ruppert et al. (2003), where the fixed effects represent the linear estimates and the random effects depict the non-linearities. Results of the panel mixed effects model are shown in table 5 in the appendix. Comparing the panel mixed effects model to the panel spline fixed effects model (full model in table 2) demonstrates that the estimates are almost the same and hence a confirmation for the robustness of

our results.

3.3 Results of subsamples

Due to the potential heterogeneity in the datasets, we explore the possibility of dividing the countries into clusters where each cluster is made up of countries with similar characteristics. We employ the K-mean approach (see Hartigan and Wong (1979)) which entails finding clusters and cluster centers in a set of unlabeled data. Using this machine learning clustering technique, we choose two cluster centers and allow the K-mean algorithm to iteratively move the centers in a way as to minimize the sum within the variance of the cluster. Firstly, the algorithm chooses random centers of the data points and identifies data points that are closer to the center points. Secondly, it computes the means of each cluster which becomes the new center points. The two steps are iterated until there is convergence. Details regarding K-mean clustering is thoroughly discussed in Hastie et al. (2017, page 509).

We estimate the clusters based on our variables of interest, notably the primary balance and the lagged debt ratio. The results of the two cluster characteristics can be found in table 4 in the appendix. Cluster 1 has an average debt of 105% and a primary surplus of 0.7% of GDP, respectively. Cluster 2, on the other hand, has a relatively smaller debt ratio of 42% and a primary surplus to GDP ratio of -0.25%, i.e. a primary deficit.[7] Cluster 1 is made up of mainly bigger economies in the EU namely: Belgium, Germany, Greece, France, Italy, Austria and Portugal. Countries that make up cluster 2 are as follows: Bulgaria, Czechia, Denmark, Estonia, Ireland, Spain, Cyprus, Latvia, Luxembourg, Hungary, Malta, Netherlands, Poland, Romania, Slovenia, Slovakia, Finland and Sweden.

Applying the penalized splines approach from above to these subsamples gives interesting results presented in table 3. The first part (column two to four) considers the differences in the two clusters, the second part (column five and six) presents the results distinguished by crisis and non-crisis years. We define the period 2008-2014 as crisis years because of the strong increase of the public debt to GDP ratios in that period, see figure 7 in the appendix, and all other years as non-crisis years. Further, in those years most of the EU countries were under the excessive deficit procedure due to the violation of the

---

[7]A graphical view of the clusters plot is shown by figure 5 in the appendix

Table 3: Panel Penalized Spline Fixed Effects - Clusters, Crises and Non-Crises period

| | | | Response variable: PBratio | | |
| | | | *Parametric* | | |
Variables	*Cluster 1*	*Cluster 1A*	*Cluster 2*	*Crises*	*Non-Crises*
Laggeddebt	0.084***	0.025*	0.020**	0.076***	0.045***
	(0.015)	(0.013)	(0.009)	(0.017)	(0.007)
GVAR	-0.0005***	-0.0003**	-0.0029***	-0.002***	-0.0006***
	(0.000)	(0.000)	(0.000)	(0.0002)	(0.0002)
YVAR	-0.00002	-0.00003	0.0003***	-0.00003	0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Inflation	-0.125***	-0.059***	-0.044***	-0.022	-0.049***
	(0.022)	(0.019)	(0.010)	(0.050)	(0.009)
Net exportratio	0.30***	0.365***	0.173***	-0.039	0.170***
	(0.093)	(0.074)	(0.031)	(0.069)	(0.029)
			Non-Parametric		
edf(laggeddebt)	2.898***	4.676***	4.509***	3.507***	4.831***
edf (GVAR)	4.976***	1.861*	8.166***	6.467***	3.936***
edf (YVAR)	2.042	1.938	2.451**	2.150	1.977*
edf (Net exportratio)	2.342**	2.530***	5.905**	3.861**	5.818***
Adj R^2	0.728	0.551	0.652	0.707	0.52
Observ	133	114	342	150	300

Standard error are indicated in parenthesis. *, ** and *** indicates statistical significance at 10%, 5% and 1% respectively.

Maastricht Treaty criteria.[8]

For all specifications the results in table 3 mainly support the general findings for the full sample from above. Active fiscal policy reacts in a positive manner to changes in debt, indicating sustainable policies. Higher spending exerts a negative effect on the primary balance and YVAR is statistically insignificant in most cases or positive but rather small for Cluster 2. Net exports and Inflation are again statistically significant and positive and negative, respectively, except for the crisis years.

Looking closer at the estimation outcomes reveals some interesting insight. The upper part of table 3 again presents the parametric part. For the cluster containing the bigger countries (Cluster 1) the debt coefficient turns out to be much larger compared to the smaller countries (Cluster 2), unless Greece is removed. Cluster 1A in table 3 gives the estimations when Greece is eliminated from the sample. This yields a considerably lower estimate for the average reaction coefficient, just as in the case of the full sample, and the average reaction coefficient is only slightly larger than for Cluster 2 then. This shows once more that Greece had achieved large primary surpluses through its austerity programs during the crisis. Similarly, the negative inflation effect (our proxy for monetary

---

[8]See European Commission (2021) for information as regards the excessive deficit procedure

policy) is stronger when Greece is included. The coefficient for net exports is positive and significant, just as for the full sample.

Additionally, we split the data set in order to study the behavior of the reaction function in crises periods and non-crises periods. Regarding the crises period, we choose the period between 2008 and 2014 that coincides with the global financial crises which affected all EU countries.[9] Figure 7 in the appendix depicts the time series plot of the debt ratio to GDP ratios for selected European countries. Strong upward sloping debt ratios can be observed until 2014 (mostly beginning 2008), which begin to decline in 2014 in most of the countries. Therefore, the crises period chosen is supported by the data.

Regarding the reaction function, the response of the primary surplus to public debt is stronger in crisis years than in non-crisis years which can be explained by austerity measures during crises, while eventually less emphasis is put on debt sustainability in normal times. Moreover, the GVAR variable has a negative effect on the primary balance that is much stronger in crisis years compared to non-crisis years. Business cycles, YVAR, are statistically insignificant[10] independent of whether crisis years or normal times are considered which may be due to the fact that we use the cyclically adjusted primary balance. Inflation and net exports have the same effect as in the case of the full sample in non-crisis years, but, are statistically insignificant in crisis years. This may be due to unconventional policy measures during the crisis years, with interest rates near the lower bound among other things.

The smooth functions with estimated degrees of freedom higher than 1 demonstrate that the relationship between the primary surplus ratio and the explanatory variables is characterized by non-linearities in most of the cases. Focusing on the public debt and primary balance to GDP ratio, figure 3 depicts the shape of the smooth function for our specifications Cluster 1 and Cluster 2 as well as for crisis and non-crisis periods.

Obviously, the response behavior is different in the two clusters: while the response in the case of Cluster 1 is almost monotonously increasing with higher public debt to GDP ratios, Cluster 2 with the smaller economies shows rather a flat curve that reminds somehow of a shallow inverse U-shaped form. For the small economies the response increases for low values of the debt ratio up to about 30%, then flattens until about 90 %

---

[9]Additional estimations show that the results are quite robust with regard to the cut in 2014, they do not differ much if the time span for the crisis is slightly extended or shortened.

[10]Again, the standard errors are rounded to the third and forth decimal place, respectively, in table 3.

Figure 3: Smooth Functions
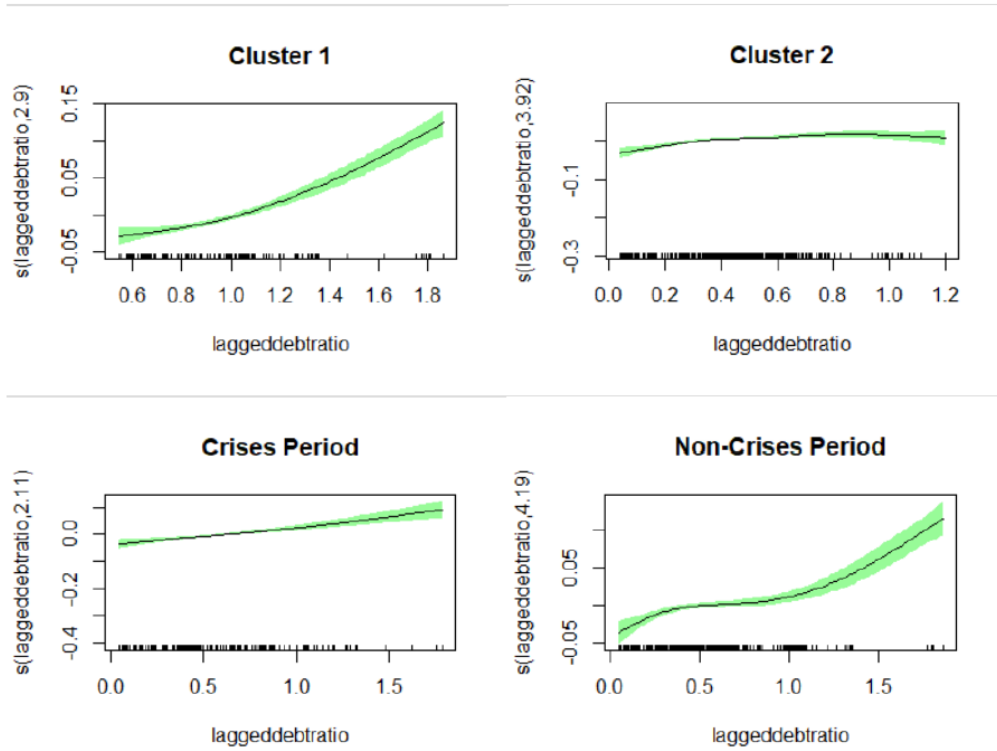
and for higher debt ratios the response decreases. This declining reaction of the primary balance for high values of the debt ratio is in line with the 'fiscal fatigue' hypothesis as suggested in Gosh et al. (2013), for instance. Comparing Crisis and Non-Crisis periods in figure 3 the latter supports the findings from above for the full sample, where the rise of the response becomes stronger for higher values of the debt ratio, while during crisis years there seems to be a constantly increasing reaction.

We should like to point out that the estimation for Cluster 1 without Greece, Cluster 1A in table 3, yields a smooth function that again resembles an (left skewed) inverse U, implying a declining response for high debt to GDP ratios, as can be seen from figure 4. Hence, we see a tendency for 'fiscal fatigue' in those economies, just as for the ones of Cluster 2, when Greece is left out. This underlines once more that Greece had achieved very high primary surpluses in its process of fiscal consolidation after the beginning of the debt crisis.

Summing up our additional results, there are some interesting findings beyond the standard fixed effects and the full sample estimation. First, all specifications mainly support the full sample outcomes and, especially, reveal sustainable public debt policies.

Figure 4: Cluster 1A - Model without Greece
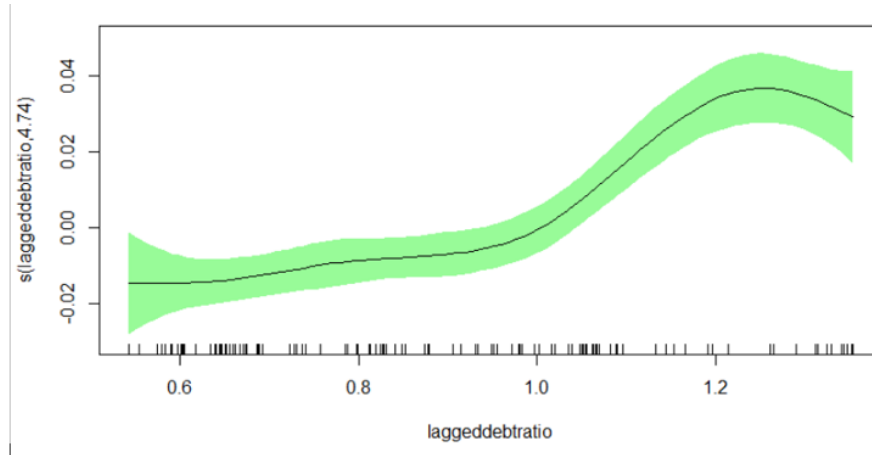
By distinguishing two (data-driven) clusters, it turns out that the fiscal response differs. For the cluster with mostly small countries, in particular, 'fiscal fatigue' characteristics appear, meaning that the response to higher debt peters out and, finally, decreases as the debt to GDP ratios become larger. For the cluster containing the three largest economies, the response to higher debt is a monotonically increasing function of the debt to GDP ratio, thus, resembling the shape obtained for the full sample. However, once Greece is eliminated from that sample, that cluster reveals 'fiscal fatigue', implying that the response declines for high debt to GDP ratios in that cluster, too. Moreover, in non-crisis years there is some phase of easing for intermediate debt ratios, but during crisis years, the response is monotonously increasing, meaning a stronger reaction to higher debt if governments face large debt ratios.

4   Conclusion

This paper has studied non-linearities in a debt sustainability analysis by resorting to penalized panel splines technique. Based on data for 25 EU economies over the years from 2000 to 2019, we estimated the fiscal response function by analyzing the relationship between the discretionary fiscal policy in terms of the cyclically adjusted primary balance to GDP ratio and the lagged debt to GDP ratio. A positive coefficient on average indicates a sustainable public debt policy, which is supported by our results for the full sample. However, we have seen that the relationship is not homogeneous across the distribution

of the debt ratios, but, rather varies across the distribution. It reveals strongly increasing reactions for small and high debt ratios, while it is rather flat for intermediate values. When Greece is eliminated from the sample, the response to higher debt ratios rises, but, not monotonously. Rather, there are phases of 'fiscal fatigue', meaning that the effort of counter-steering peters out and declines once the debt to GDP ratios exceed certain thresholds.

In a next step, we have split the full sample into different categories to explore the heterogeneity in the data set. We have classified the data set into two categories to ascertain the reaction function in times of debt crises and in normal times. The reaction function in normal times exhibits a similar pattern as for the full sample, while during years of crises the reaction to public debt has been found to be stronger. Additionally, we have used a data-driven algorithm to cluster the data set into two distinct groups. Results for the first cluster, consisting of mainly larger EU economies, shows a strong response of the primary balance to increases in the debt ratio yielding strong evidence for fiscal sustainability as long as Greece is included in that sample. Once Greece is taken out, however, the average reaction coefficient takes on a much a lower value and the smooth function is characterized by an inverted U-shaped relationship: it rises with increasing debt to GDP ratios, reaches a maximum and, then, declines again. That shape was found for the second cluster, too, that consists of mostly smaller EU economies. To test the robustness of our results, we have estimated a fiscal reaction function using spline mixed effects models for our panel data set and the main results remain unchanged, pointing out the robustness of our estimations.

References

Afonso, A. and J.T. Jalles (2019), "Fiscal Reaction Functions Across the World: In Quest of Statistical (In)significance", *FinanzArchiv/Public Finance Analysis*, Vol. 75: 207-228

AMECO (2021), "European Commission's Directorate General for Economic and Financial Affairs macro-economic database", Available at `https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/macro-economic-database-ameco/ameco-database_en`, [Accessed on 08.03.2021]

Baltagi, B.H. and Li, D. (2002), "Series Estimation of Partially Linear Panel Data Models

with Fixed Effects", *Annals of Economics and Finance, Society for AEF*, vol. 3(1), pp 103-116

Barro, R. (1979), "On the Determination of Public Debt", *Journal of Political Economy,* vol. 87, no. 5, pp. 940 - 971

Beqiraj, E., Fedeli, S. and Forte, F. (2018), "Public debt sustainability: An empirical study on OECD countries", *Journal of Macroeconomics*, vol. 58: 238-248

Bohn, H. (1995), "The sustainability of budget deficits in a stochastic economy", *Journal of Money, Credit and Banking*, Vol. 27: 257-271.

Bohn, H. (1998), "The behavior of U.S. public debt and deficits", *Quarterly Journal of Economics*, Vol. 113: 949-963.

Checherita-Westphal, C., Zdarek, V. (2017), "Fiscal reaction function and fiscal fatigue: evidence for the euro area" , *ECB Working Paper,* No 2036/March 2017.

Eilers, P. H. C. and Marx, B. D. (1996), "Flexible Smoothing with B-Splines and Penalties", *Statistical Science*, vol. 11(2), pp. 89 - 102

European Commission (2021), "Excessive deficit procedures - overview", Available at https://ec.europa.eu/info/business-economy-euro/economic-and-fiscal-policy-coordination/eu-economic-governance-monitoring-prevention-correction/stability-and-growth-pact/corrective-arm-excessive-deficit-procedure_en, [Accessed on 26.04.2021]

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

Granger, C.W.J. (2008), "Non-linear models: Where do we go next - Time varying parameter models?" *Studies in Nonlinear Dynamics & Econometrics,* Vol. 12, Issue 3, Article 1: 1-9.

Greiner, A. and Fincke, B. (2015), "Public Debt, Sustainability and Economic Growth. Theory and Empirics,"*Springer Verlag, Cham.*.

Hakkio, C.S. and Rush, M. (1991), "Is the Budget Deficit 'Too Large?" *Economic Inquiry,* Vol. 29: 429-445.

Hamilton, J.D. and Flavin, M. (1986), "On the limitations of government borrowing: A framework for empirical testing", *The American Economic Review*, Vol. 76: 808-19.

Hartigan, J. A. and Wong, M. A. (1979), "Algorithm AS 136: A K-means clustering algorithm", *Applied Statistics*, vol. 28, pp 100 - 108

Harezlak, J., Ruppert, D. and Wand M.P. (2018), "Semiparametric Regression with R", *Springer-Verlag New York*.

Hastie, T. and Tibshirani, R. (1990), "Generalized Additive Models", Chapman and Hall, London

Hastie, T, Tibshirani,R. and Friedman, J. (2017), "The Elements of Statistical Learning, Data Mining, Inference, and Prediction", Second edition, *Springer series in statistics New York*

Henderson, D. J., Carroll, R. J. and Li, Q. (2008), "Nonparametric estimation and testing of fixed effects panel data models", *Journal of Econometrics*, vol. 144(1), 257 - 275

Kauermann, G. (2005), "Penalized spline smoothing in multivariable survival models with varying coefficients", *Computational Statistics & Data Analysis,* vol. 49 (1), pp. 169 - 186

Puetz, P. and Kneib, T. (2016), "A Penalized Spline Estimator for Fixed Effects Panel Data Models", *SOEP papers on Multidisciplinary Panel Data Research* 827, DIW Berlin

Ruppert, D., Wand, M.P. and Carroll, R.J. (2003), "Semiparametric Regression",*Cambridge University Press*

Su, L. and Ullah, A. (2006),"Profile Likelihood Estimation of Partially Linear Panel Data Models with Fixed Effects", *Economics Letters*, vol. 92, 75-81

Trehan, B. and Walsh, C.E. (1991), "Testing intertemporal budget constraints: theory and applications to U.S. Federal budget and current account deficits", *Journal of Money, Credit and Banking*, Vol. 23: 206-223.

Wilcox, D.W. (1989), "The sustainability of government deficits: Implications of the present-value borrowing constraint", *Journal of Money, Credit and Banking*, Vol. 21: 291-306.

Wiesenfarth, M., Krivobokova, T., Klasen, S. and Sperlich, S. (2012), "Direct Simultaneous Inference in Additive Models and its Application to Model Undernutrition", *Journal of the American Statistical Association* vol. 107(500), 1286- 1296.

Wood, S. N. (2000), "Modelling and smoothing parameter estimation with multiple quadratic penalties", *Journal of the Royal Statistical Society*, Series B 62, pp 413 - 428

5 Appendix

Table 4: Clustering Characteristics, measured in % of GDP

Cluster	laggeddebtratio	PBratio
1	105	0.7
2	42.2	-0.25

Figure 5: K-mean cluster plot
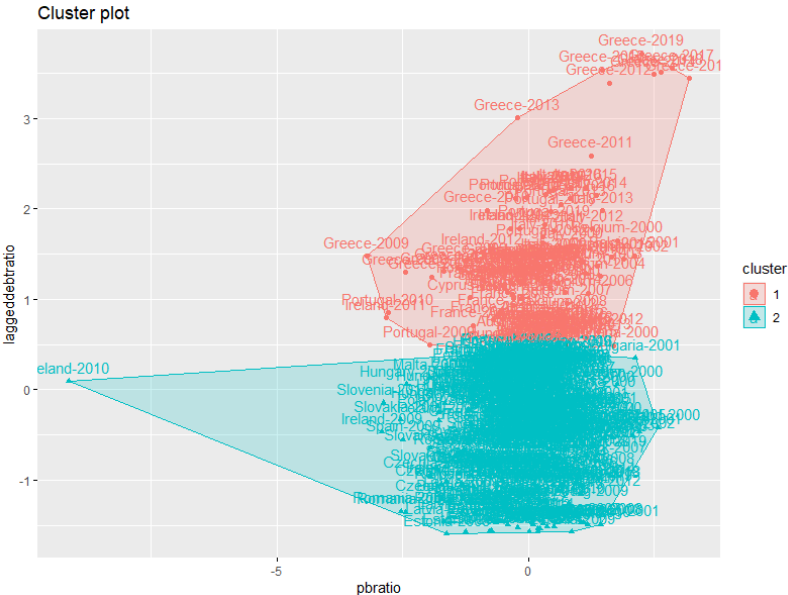
Table 5: Robustness Test - Spline Mixed Effect Model

| | | Response variable: PBratio | | |
| | | Parametric | | |
Variables	1	2	3	4
Laggeddebt	0.056***	0.066***	0.054***	0.058***
	(0.001)	(0.007)	(0.007)	(0.007)
GVAR	-0.002***	-0.002**	-0.002***	0.048***
	(0.000)	(0.000)	(0.000)	(0.000)
YVAR	0.0001	-0.000002	-0.0001	-0.003
	(0.000)	(0.000)	(0.000)	(0.000)
Inflation		-0.027**	0.049***	-0.056***
		(0.008)	(0.001)	(0.007)
Net exportratio			0.145***	-0.001***
			(0.000)	(0.000)
		Non-Parametric		
edf(laggeddebt)	6.462***	7.083***	6.828***	7.206***
edf (GVAR)				8.510***
edf (YVAR)				2.909*
edf (Net exportratio)				6.881***
Adj R^2	0.174	0.156	0.30	0.471
Observ	500	500	500	500

Standard error are indicated in parenthesis. *, ** and *** indicates statistical significance at 10%, 5% and 1% respectively.

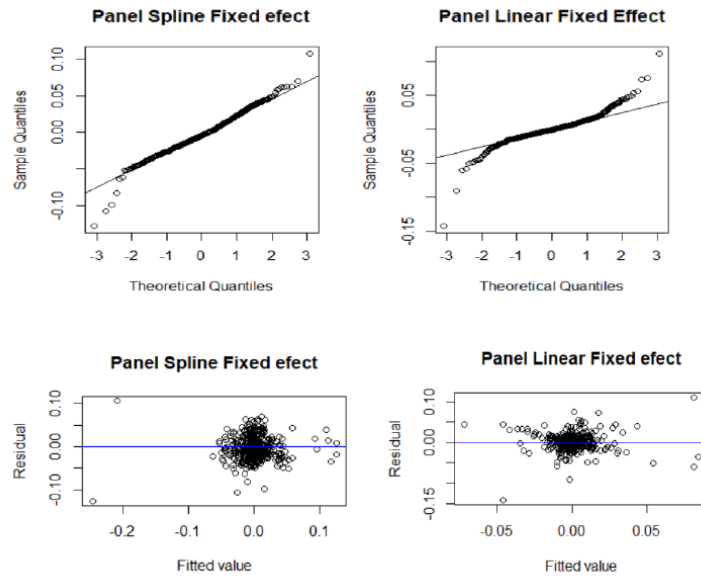Figure 6: Residual diagnostic plots

Figure 7: Laggeddebtratio with shaded grey area (2008 and 2014) indicating crises period