CESifo WORKING PAPERS

6976
2018

April 2018

Modeling Fiscal Sustainability in Dynamic Macro-Panels with Heterogeneous Effects: Evidence from German Federal States

Lars P. Feld, Ekkehard A. Köhler, Julia Wolfinger

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6976*
Category 1: Public Finance

Modeling Fiscal Sustainability in Dynamic Macro-Panels with Heterogeneous Effects: Evidence from German Federal States

Abstract

In this paper, we extend Henning Bohn's (2008) fiscal sustainability test by allowing for slope heterogeneity and cross-sectional dependence (CD). In particular, our econometric approach is the first that allows fiscal reaction functions (FRF) to capture unobserved heterogeneous effects from business and fiscal policy cycles. We apply this econometric approach to sub-national public finance data of the German Laender between 1950 and 2015 and find that their fiscal policy only partly meets fiscal sustainability criteria. According to our results, politicians have significantly reacted to increasing debt levels by increasing budget surpluses since 1991. However, time-series evidence for longer periods does not indicate a significant and positive reaction to increasing debt levels in the West German Laender panel.

JEL-Codes: H620, H770, H720, C230.

Keywords: fiscal sustainability, public debt, panel data, cross-sectional dependence.

*Lars P. Feld**
Albert-Ludwigs-University Freiburg
Walter Eucken Institute
Goethestrasse 10
Germany – 79100 Freiburg i. Br.
feld@eucken.de

Ekkehard A. Köhler
Albert-Ludwigs-University Freiburg
Walter Eucken Institute
Goethestrasse 10
Germany – 79100 Freiburg i. Br.
koehler@eucken.de

Julia Wolfinger
Albert-Ludwigs-University Freiburg
Walter Eucken Institute
Goethestrasse 10
Germany – 79100 Freiburg i. Br.
wolfinger@eucken.de

*corresponding author

1 INTRODUCTION

Given its AAA rating and leading market opinions, the Federal Republic of Germany appears to be one of the world's most trustworthy debtors, even matching the United States. On first sight, this perception seems to be valid: Numerous studies (e.g., Schularick and Taylor 2012; Reinhart and Rogoff 2010) have found that German politicians improve (i.e., increase) the primary balance when the government debt ratio increases. Looking ahead, Germany's fiscal space is reported to range among the best in the EU (Buti and Carnot 2016; Auerbach 2016). While Germany plays a key role in primary surplus simulations of the EU15 (D'Erasmo et al. 2016), there are only a few studies that have taken a closer look on the long-term sustainability of the sub-national level of the Federal Republic of Germany, the German Laender. This is quite surprising for five reasons:

First, the 16 German Laender have accumulated a large share of total outstanding debt (about 40 percent). Second, the Laender will have to comply with the German debt brake by 2020 putting local public finances under more pressure to withstand the recently assessed "high" fiscal fatigue (Checherita-Westphal and Žďarek 2017). Third, demographic change will eventually put Germany's sub-national budgets under pressure on the expenditure and revenue sides (BMF 2017). Fourth, the German Laender are subject to a sophisticated system of fiscal equalization recently reformed by strengthening the vertical component. Fifth, the federal government already bailed out several Laender across time and wants to avoid another bailout in the future. To sum up these five reasons, we argue that public finances on the level of the Laender will become ever more important for the assessment of German general public finances. Empirical evidence is however rare.

In this paper, we address the shortcomings of the existing literature not to have properly considered the fragile relationship between the federal and the Laender levels. We analyze whether German Laender pursue sustainable fiscal policies against the backdrop of Henning Bohn's (2008) model-based sustainability (MBS) test. Our key research question is thus: Do politicians react to higher debt levels by increasing primary surplus in the following fiscal year(s)?

We estimate state specific fiscal reaction functions (FRF) across and within West German federal Laender between 1950 and 2015. In addition, we explore fiscal sustainability after reunification for all Laender between 1991 and 2015. The econometric analysis focuses on the data properties and their implications for the econometric modeling of MBS-tests. In particular, we show that heterogeneous effects remain unobserved in standard pooled OLS models with two sided fixed-

effects (FE). In addition, we show that "second generation" macro panel modeling offers a meaningful approach to grasp unobserved heterogeneity that contributes to non-independent and identically distributed (i.i.d.) errors in standard OLS models.

This paper thus contributes to the literature in two ways: First, we offer an alternative econometric approach to Henning Bohn's (2008) fiscal sustainability test by allowing for slope heterogeneity and cross-sectional dependence (CD) in a dynamic macro panel. Our modeling is the first that allows FRFs to capture unobserved heterogeneous effects from business and fiscal policy cycles. Second, we apply this econometric model to the German Laender between 1950 and 2015 and find that their fiscal policy only partly meets fiscal sustainability criteria. We conclude that politicians have significantly reacted to increasing debt levels by increasing budget surpluses only since 1991. However, time-series evidence for longer periods does not indicate a significant and positive reaction to increasing debt levels in the West German Laender panel.

The remainder of the paper is organized as follows: Section 2 briefly describes previous empirical studies and discusses their econometric approach. Section 3 presents the fiscal dataset and the econometric model proposed in this paper. In Section 4, we summarize the results. Section 5 concludes.

2 REVIEW OF LITERATURE

Empirical studies on sustainability of fiscal policy have become increasingly important, most notably for the member states of the European Economic and Monetary Union (EMU) that have been under closer scrutiny since the turbulences in European government bond markets. While there are many studies that explore fiscal policy of the Federal Republic of Germany on the central government level, evidence on the level of the German Laender is rare. The few available studies have in common that they define sustainability against the theoretical background of the intertemporal government budget constraint (IGBC). In particular, all studies (implicitly) impose the no-Ponzi game condition by assuming that outstanding debt is equal to the present value of the primary balance (see Table 1). The studies differ with respect to their econometric approaches. Motivated by Bohn (2007) and D'Erasmo et al. (2016) we identify two different approaches and thus divide the empirical literature into two categories (as depicted in Table 1 column 2).

Table 1: An Overview of Existing Studies

Study Author (Year)	Econometric Modeling of Fiscal Sustainability	Time Series Dimensions in integer numbers West (East)			Time Series Model	Empirical Finding
		t	n	t*n		Is public finance sustainable?
Kitterer and Finken (2006)	Category I: Stationarity/Unit-root and/or Cointegration tests of fiscal variables	33 (12)	1	33 (12)	Univariate	YES [HE, NW, SN] NO [others]
Kitterer (2007)		33 (12)	1	33 (12)	Univariate	See above.
Burret et al. (2017)		61 (21)	1	61 (12)	Uni- and Multivariate (Cointegration)	NO [SL, HB, RP] Ambiguous results for most Laender
Burret et al. (2016)		61	11	671	Multivariate (Pesaran MG)	NO [SL, HB, RP] YES[BY,SN,HH]
Claeys et al. (2008)	Category II: Fiscal Reaction Functions (FRF), i.e., Model-based sustainability (MBS)	35 (11)	1	35 (11)	Univariate	NO
Herzog (2010)		31 [for 2 Laender]	1	31	Univariate	NO [BE] Ambiguous [BW]
Fincke and Greiner (2011)		31	1	31	Univariate	NO [SL, HB, BE] YES[BY,BW,HH] Others rather YES
Potrafke and Reischmann (2015)		30	11	330	Multivariate (2 FE OLS)	NO[w/o transfers] YES[w transfers]

Notes: Two digit Laender codes for the German Laender can be found in Table 2.

The first category includes studies that analyze time-series characteristics of the fiscal datasets (i.e., unit root and cointegration tests). Time series characteristics allow the authors of these studies to infer fiscal sustainability. In some cases, authors develop simple indicators beyond univariate time series analysis and test whether expenditures and revenues are cointegrated with a long-term relationship vector [1;-1]. The idea of this approach is to find evidence that a one percentage point increase in expenditures is accompanied with a one percentage point increase in revenues. Econometric modeling of fiscal sustainability in this first category is criticized because it does not connect the tests to outstanding debt (D'Erasmo et al. 2016).

The second category includes studies that estimate FRF in fiscal datasets. Regressions of the FRF may help to infer a causal relationship between the initial debt level and the fiscal policy of policymakers. Thus, authors of these studies analyze whether politicians react to increasing debt levels by increasing budget surpluses. Papers in this second category apply the MBS test as proposed by Bohn (1995, 1998).

Importantly, Bohn (2007) doubts that econometric modeling of the first category can infer anything economically meaningful. According to his argument, stationarity and cointegration are a matter of differencing the fiscal dataset and thus cannot be taken as evidence to infer fiscal sustainability. The possibility of integration of a finite but arbitrarily high order has invalidated fiscal sustainability tests based on stationarity and cointegration. However, Bohn (1995) admits that primary balance is not unambiguous, as it is correlated with macro-fluctuations that must be controlled for.

In the case of the German Laender the studies that apply modeling of the first category are Kitterer (2007) and Kitterer and Finken (2006). They apply univariate stationarity tests of the debt-to-GDP ratio in the West (1974-2004) and East German Laender (1992-2004), respectively. They reject the hypothesis of fiscal sustainability for all Laender except Hesse, North-Rhine Westphalia and Saxony. Evidence on the East German Laender is not robust due to the small-t problem. Studies that explore a larger dataset are Burret et al. (2016, 2017) who analyze German Laender from 1950-2010 and find only weak evidence for fiscal sustainability (Burret et al. 2017). Saarland, Bremen and Rhineland-Palatinate do not meet any of the cointegration requirements. Applying panel cointegration tests and relaxing the requirements for meeting cointegration, Burret et al. (2016) find a sub-panel of several West German Laender that meet cointegration requirements. All of these studies have in common that they do not find any significant evidence for the East German Leander, which is a shortcoming of the literature. In addition, inference from time series characteristics is not without ambiguity. However, the panel study by Burret et al. (2016) overcomes many of the arguments raised against cointegration as the t*n vector is large.

Studies that apply modeling of the second category are Claeys et al. (2008), Herzog (2010), Fincke and Greiner (2011), and Potrafke and Reischmann (2015). Herzog (2010) estimates FRFs and finds evidence in favor of a significant reaction to increasing debt-to-GDP ratios in Baden-Wuerttemberg but rejects it for Berlin. Claeys et al. (2008) analyze fiscal data of the German Laender over the period 1970-2005 (West Germany) and 1991-2005 (East Germany), respectively. Their results suggest that Laender governments hardly consolidate such that regional finances are not sustainable. Fincke and Greiner (2011) conclude that the reaction of primary surpluses to changes in public debt levels is positive over the period 1975-2006 in all West German Laender but Berlin. Potrafke and Reischmann (2015) estimate a Bohn-model for West German Laender during the

period 1980-2010 in panel OLS regressions with two sided fixed effects. Their results suggest that fiscal policy is sustainable if fiscal transfers are included.

The previous studies on fiscal sustainability in the German Laender show five notable shortcomings: First, the evidence for the East German Laender is, so far, not robust due to the small-t problem. Second, most univariate studies do not analyze time series characteristics, which are meaningful to discuss given the relatively short t-dimension. Even if we disregard Bohn's arguments against time series analysis of fiscal data, the only long time series analysis exceeding 50 observations is Burret et al. (2017). However, they find heterogeneous results. Third, most of the FRF regressions are based on small-t and thus come to different (and partly contradictory) results. This calls for a larger dimensional dataset. The only study with such a multi-dimensional modeling has only 330 observations (Potrafke and Reischman 2015). In addition, Burret et al. (2017, 2016) show that time-series characteristics are heterogeneous even among the German Laender. Fourth, and based on the findings of Burret et al. (2016), CD is not controlled for in the empirical literature of FRFs. Fifth, the time period of the "economic miracle" in the 1950s is not covered by any study, so far. Finally, most studies remain tacit how to correctly model FRF.

3 THEORETICAL FRAMEWORK AND EMPIRICAL MODELING

Is there evidence that politicians react to higher debt levels by increasing primary surpluses in the following fiscal year(s)? Like many others in the field, we rely on a priori theoretical reasoning of the FRF that builds on the IGBC (D'Erasmo et al. 2016). In accordance with Bohn (2008), we estimate the following regression across all Laender to infer fiscal sustainability as follows.

$$Primary\ Surplus_{it} = \alpha Public\ debt_{it-1} * b_i + \beta Control_{it} + u_{it}$$

The dependent variable measures the primary surplus-to-GDP ratio of each German Land excluding interest payments as suggested by Bohn (1998). We include vertical and horizontal transfers for the following reason. We want to find out whether the German fiscal equalization system lives up to its own objective to smooth out differences in economic performance and corresponding fiscal capacities of the Laender. If we cannot find evidence for a positive and significant reaction of fiscal policy conditional on an increase of the lagged debt-to-GDP ratio, we conclude that even vertical and horizontal transfers cannot bolster regional public finances. In accordance, the main explanatory variable is the lagged debt-to-GDP ratio of each Land. To

account for tax smoothing (Barro 1981, 1986), we control for the business cycle (YVAR) and fluctuations of government spending (GVAR). The YVAR variable has negative values if current output (y_t) is above its Hodrick-Prescott-filtered trend (y_t), while the GVAR variable is negative if current expenditures are below its trend. Note that the variables are filtered with the very same method. As an alternative to Barro's approach and in line with Potrafke and Reischmann (2015), we also vary the output and expenditure gap controls as suggested by Bohn (2008). According to theory, we expect a negative correlation of YVAR, GVAR and expenditure with primary surplus.

In addition to the above-mentioned modeling, some studies include a lagged dependent variable in the FRF to control for the persistence of fiscal policy. Claeys (2006) argues that fiscal instruments (in this case the primary surplus) can adjust only over several periods to the targeted levels. The adjustment is, however, time lagged due to parliamentary processes, sunk (decision) costs or other institutional settings. Theofilakou and Stournaras (2012) contend that the lagged primary surplus should always be controlled for in FRF estimations to overcome the omitted variable bias. Thus, we end up in a dynamic panel data model. We will show the difference of excluding and including a lagged dependent variable in each of the applied estimators.

The dataset is a macro panel dataset with 16 cross-sections of fiscal data on revenues, expenditures, debt, and primary surplus comprising the 10 West German Laender from 1950 to 2015 and 6 East German Laender from 1991 to 2015. With more than 650 observations in the West German case and 125 observations in the East German case, the panel dimensions are moderate but more than twice as large as the panel analysis by Potrafke and Reischmann (2015).

Choosing the "correct" estimator may be subject to model uncertainty. Motivated by Durlauf et al. (2016), we analyze time series characteristics to specify an econometric modeling that is closer to the "true" modeling, instead of setting a model a priori, as most studies in the FRF literature do (D'Erasmo et al. 2016).

In a first step, we analyze time series characteristics of the variables of each cross-section and for the panel as a whole. This evidence is important to inform the modeling of the FRF regression. The benefit of this strategy is twofold: First, we gain insights into fiscal policy in each Land (time series analysis) and in the dynamics of all Laender together, before we specify the econometric modeling of the FRF. Second, we disclose every step of the modeling in a transparent manner and document evidence on each step of the analysis.

Table 2: Laender-Specific Time Series Evidence

Land	Code	Period	Stationarity[1]				Estimation of fiscal reaction of surplus on debt[2]
		Years	Debt	Primary surplus	Revenue	Expenditure	rho coefficient
Baden-Wuerttemberg	BW	1950-2015	~	~	~	~	0,044***
Bavaria	BY	1950-2015	~	√	~	~	0,092***
Berlin	BE	1950-2015	X	~	X	X	0,022
Brandenburg	BB	1991-2015	~	~	~	~	0,236***
Bremen	HB	1950-2015	X	X	X	~	0,013
Hamburg	HH	1950-2015	~	√	~	~	0,030**
Hesse	HE	1950-2015	~	~	~	X	0,016
Mecklenburg-Western Pomerania	MV	1991-2015	~	~	~	~	0,195***
Lower Saxony	NI	1950-2015	X	√	~	X	0,022**
North Rhine-Westphalia	NW	1950-2015	~	~	~	~	0,019***
Rhineland-Palatinate	RP	1950-2015	~	~	~	~	0,005
Saarland	SL	1960-2015	X	X	X	X	0,015
Saxony	SN	1991-2015	~	~	~	~	0,086
Saxony-Anhalt	ST	1991-2015	~	X	~	~	0,155***
Schleswig-Holstein	SH	1950-2015	~	~	~	~	0,018*
Thuringia	TH	1991-2015	~	~	X	~	0,150***

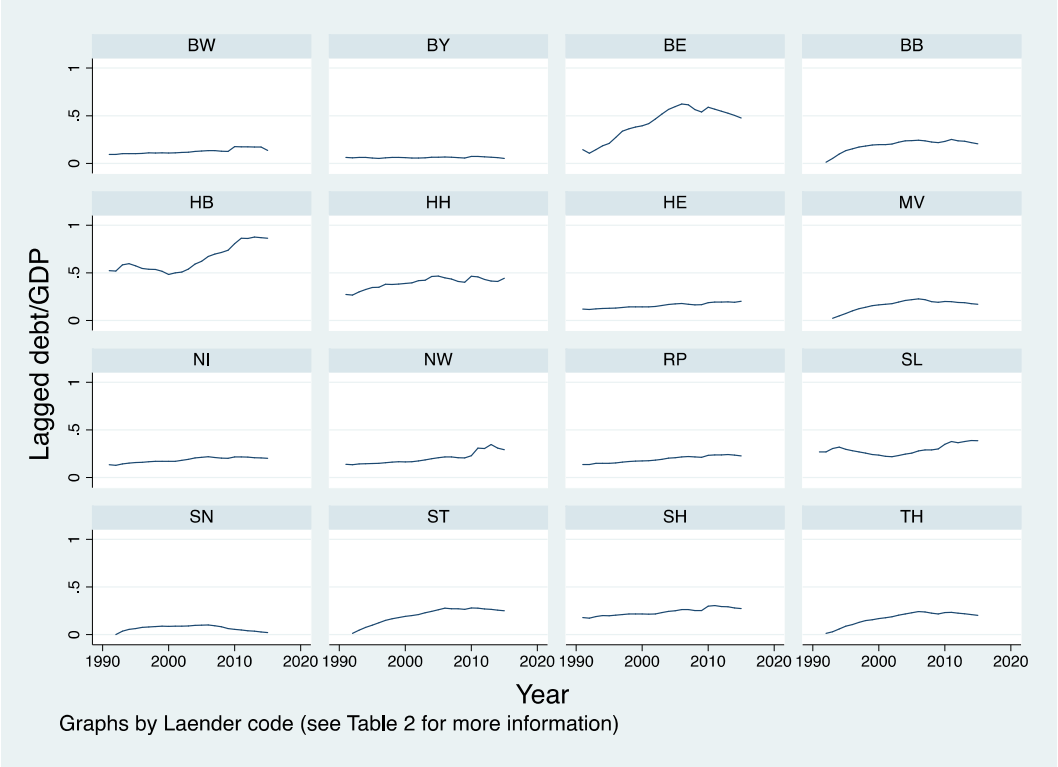
[1] Stationarity: Unit-root tests Augmented Dickey-Fuller, Phillips-Perron und Kwiatkowski, Phillips, Schmidt und Shin. √ if Null is rejected in Levels in any test, i.e., if time series is stationary in levels. X if any of the tests cannot reject the null. ~ if we obtain mixed evidence.
[2] Bohn MBS Test with cross-sectional regressions: pb = alpha + rho*d(-1) + gamma*GVAR + delta*YVAR + epsilon. Robust standard errors.

We start with a time series analysis of the variables in each cross-section to explore data properties. We report the findings in Table 2. We find few debt time series in the Laender that can be described by a variation of their past realizations and a normally distributed error. Therefore, stationarity of debt can be rejected across most Laender. In addition, revenues and expenditures are not stationary in levels either. We have mixed evidence for primary surpluses. In Bavaria, Hamburg and Lower Saxony we find stationary deficits. We reject stationarity of all fiscal variables in the Saarland. However, we refrain from inferring fiscal (un-)sustainability from the abundant verdict against stationarity of fiscal variables due to the objections raised by Bohn (2007) against assessing fiscal sustainability with unit-root tests.

We depict the estimated coefficients of the FRF in the last column of Table 2. For the sake of robustness, we have conducted estimations using Barro's GVAR and YVAR controls as well as

Bohn's output and expenditure gaps. We cannot find a significant and positive reaction to increasing initial debt levels in Berlin, Bremen, Hesse, Rhineland-Palatinate, Saarland and Saxony. Saxony is an outlier with continuously reducing debt-to-GDP levels since its peak at 10 percent. Figure 1 depicts the time series of debt-to-GDP for all German Laender since 1991. Inferring fiscal unsustainability due to an insignificant reaction coefficient in Saxony as depicted in Table 2 would be misleading given the concave debt series as shown in Figure 1 ("SN").

Figure 1: Panel Time Series of Main Explanatory Variables of All German Laender 1991-2015

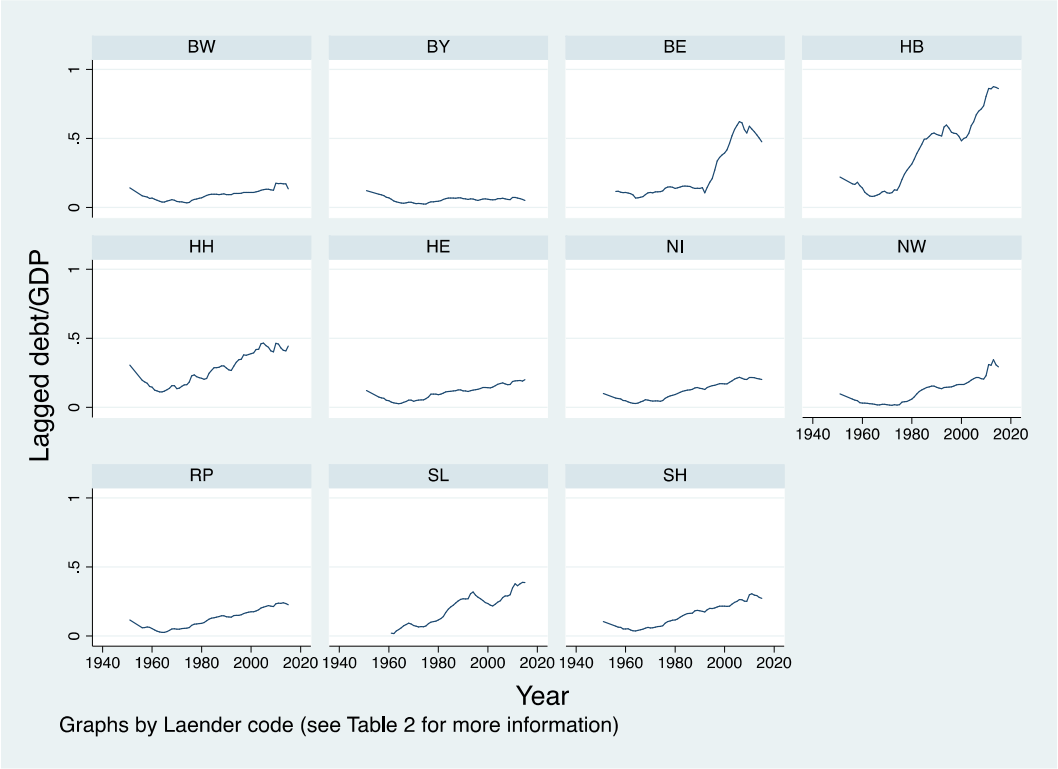


We also doubt that the insignificant coefficient for Hesse is evidence for unsustainable public finances. If we control for a trend in the debt series, we cannot reject unit-roots. We show the Laender specific time series of the main explanatory variable in Figure 2 since 1950 (and since 1960 for Saarland that only became part of the Federal Republic of Germany in 1955 and was not reporting until 1960).

Highlighting the heterogeneity of the different Leander with regard to their time series characteristics, we refrain from further analyzing fiscal sustainability in each cross-section and rather concentrate on the panel data analysis. We are interested in exploring as to how to control for slope heterogeneity of the FRF regressions. Thanks to the advances in macro panel

econometrics, we have developed a test procedure to analyze the dataset of all Laender since 1991 and the long-time series panel of the West German Leander since 1950. We proceed in three consecutive steps:

Figure 2: Panel Time Series of Main Explanatory Variable of West German Laender 1950-2015



First, we test for CD as suggested by Pesaran (2004) and applied to the German Laender by Burret et al. (2016). As depicted in Table 3, we can reject the null hypothesis of cross-sectional independence for any panel time series at the 1 % significance level. This is evidence in favor of controlling for CD in the panel regressions. Second, we explore panel time-series stationarity and choose Pesaran's CADF test as we have evidence for CD in the panel as a whole for all series. We report the results in Table 3 with and without allowing for a trend in the series.

The test is estimated with Laender specific lag lengths that are determined by 16 VAR models for each series (with and without a time trend). The null hypothesis assumes that all series are non-stationary. Without a time trend, the null hypothesis of non-stationarity can be rejected for all variables at least on the 10% significance level except for the expenditure series, that are I(1). Allowing for a time trend, we cannot reject the null hypothesis of non-stationarity. This is evidence

that we can model the variables in levels in the panel estimation without causing interference by I(1) panel variables. Note that YVAR and GVAR are both stationary and thus meaningful controls.

Table 3: Panel Time Series Characteristics

Variable	CD[1]	Non-Stationarity[2] (without a trend)	Non-Stationarity[2] (including a trend)	Westerlund ECM cointegration test[3]
Revenue to GDP	28.66 (0.000)	-1.523 (0.064)	0.338 (0.632)	Gt: -3.316 (0.000) Ga: -18.898 (0.000) Pt: -13.013 (0.000) Pa: -16.771 (0.000)
Expenditure to GDP	43.88 (0.000)	0.958 (0.831)	-0.401 (0.344)	
Primary Surplus to GDP	28.55 (0.000)	-3.307 (0.000)	0.722 (0.765)	Gt: -3.232 (0.000) Ga: -23.408 (0.000) Pt: -8.112 (0.650) Pa: -14.759 (0.000)
Lagged Debt to GDP	49.27 (0.000)	-1.534 (0.063)	0.879 (0.810)	
YVAR	29.09 (0.000)	-3.064 (0.001)	-7.931 (0.000)	
GVAR	67.15 (0.000)	-10.246 (0.000)	-2.087 (0.018)	

Note: p-values are reported in parentheses.
[1] CD is tested with the Pesaran (2004) CD test. The null hypothesis assumes cross-section independence.
[2] Pesaran's unit root test that controls for CD with the null of non-stationarity in the panel time series.
[3] Westerlund error-correction-based panel cointegration test. Under the null hypothesis the series are not cointegrated.

In a third step, we explore cointegration between the dependent variable (primary surplus) and its main explanatory variable of the FRF (lagged debt-to-GDP). In addition, we also explore whether expenditures and revenues are cointegrated. We apply Westerlund's (2007) error-correction panel cointegration procedure. The four regressions test for the absence of cointegration by determining whether there exists error correction for individual panel members or for the panel as a whole. We reject the null hypothesis of no cointegration at the 1% significance level for all FRFs and the expenditure-revenue relationship.

Following our procedure from step 1 to step 3, we have evidence to control for short and long run dynamics in the modeling of FRFs in a panel dataset of CD and multi-sloped cross-sections. We will therefore apply the standard OLS with two-sided FE with Kraay-Driscoll errors that are robust to CD as well as Pesaran's Common Correlated Effects Mean Group estimator (CCEMG).

Pesaran's (2006) CCEMG estimator allows for CD, time-variant unobservables with heterogeneous impact across panel members and problems of identification. The CCEMG estimator augments the group-specific regression equation apart from the standard regressors and the intercept by adding cross-section averages of the dependent and independent variables as additional regressors. Thus, in practical terms, cross-section averages for all observable variables in the model are computed (using the data for the entire panel) and then added as explanatory variables in each of the N regression equations. Thereby it can account for unobserved common factors. Compared to a FE estimator, which is controlling for time-invariant unobservables only, the CCEMG estimator can control for time-variant unobservables, too.

4 RESULTS

Table 4 shows the results for the long time-series panel estimation of the West German Laender from 1950-2010. The upper part of Table 4 displays the results of the FE modeling with Driscoll-Kraay standard errors reported in parentheses. We have selected Driscoll-Kraay standard errors because they are robust to CD as well as a heteroskedastic and autocorrelated errors. We control for year FE, the time dummies are excluded from the output table (Table 4). If we do not include a lagged dependent variable and estimate a two-way FE OLS with standard errors robust to CD, we cannot find a significant fiscal reaction coefficient (see specifications (1)-(3), Table 4). This is evidence against a conditional and significant response to increasing initial debt levels. Note that we find a significant coefficient of the expenditure gap in the data and a positive and significant constant.

If we control for persistence in fiscal policy as suggested by Claeys (2006) for sub-national data and include the lagged dependent variable (see specifications (4)-(6) Table 4), we find a positive and significant reaction parameter on the 5% level of significance. Thus, estimations shown in column (1) to (3) suffer from an omitted variable bias: Including the lagged dependent variable as depicted in columns (4) to (6) shows a significant impact of this variable as well as a changed significance of the main explanatory variable (lagged debt-to-GDP).

Table 4: FRF Evidence for German Laender, 1950-2015

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Lagged debt to GDP	0.021	0.021	0.021	0.017**	0.017**	0.017**
	(0.015)	(0.015)	(0.015)	(0.008)	(0.008)	(0.008)
YVAR		0.224			0.044	
		(0.387)			(0.321)	
GVAR		-0.210*			-0.158*	
		(0.122)			(0.086)	
Output gap			-0.115			0.360***
			(0.113)			(0.109)
Expenditure gap			-0.204			-0.158*
			(0.124)			(0.086)
Lagged Primary surplus to GDP				0.685***	0.679***	0.681***
				(0.052)	(0.049)	(0.049)
Number of observations	655	655	655	645	645	645
Within R-squared	0.325	0.336	0.335	0.652	0.658	0.659
Variables	(7)	(8)	(9)	(10)	(11)	(12)
Lagged debt to GDP	0.067***	0.067***	0.061***	0.062***	0.059***	0.056***
	(0.013)	(0.019)	(0.016)	(0.013)	(0.017)	(0.015)
YVAR		4.147***			1.747	
		(1.433)			(1.176)	
GVAR		-0.424***			-0.445***	
		(0.099)			(0.060)	
Output gap			0.121			0.387
			(0.638)			(0.381)
Expenditure gap			-0.437***			-0.461***
			(0.105)			(0.063)
Lagged Primary surplus to GDP				0.432***	0.360***	0.360***
				(0.065)	(0.078)	(0.079)
Cross-section average of primary surplus to GDP	1.011***	1.009***	1.013***	1.006***	0.991***	0.987***
	(0.193)	(0.207)	(0.207)	(0.184)	(0.186)	(0.188)
Cross-section average of lagged debt to GDP	-0.068***	-0.072***	-0.068***	-0.062***	-0.062***	-0.059***
	(0.019)	(0.022)	(0.020)	(0.017)	(0.018)	(0.017)
Cross-section average of YVAR		-4.506***			-1.592*	
		(1.402)			(0.912)	
Cross-section average of GVAR		0.346***			0.391***	
		(0.124)			(0.076)	
Cross-section average Output gap			-0.123			-0.379
			(0.633)			(0.386)
Cross-section average of Expenditure gap			0.346***			0.402***
			(0.133)			(0.077)
Cross-section average of lagged Primary surplus to GDP				-0.505***	-0.455**	-0.454**
				(0.170)	(0.187)	(0.190)
Constant	-0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of observations	655	655	655	645	645	645

Notes: The dependent variable of the regression (1) to (12) is the primary surplus per GDP. Column (1) to (6) is estimated with a fixed effects regression. Driscoll-Kraay standard errors are reported in parentheses. It is controlled for year fixed effects, the time dummies are excluded from the output table. Column (7) to (12) is estimated with Pesaran (2006) CCEMG estimator. The estimator controls and corrects for CD, time-variant unobservables with heterogeneous impact across panel members.

We find evidence for statistically significant cross-sectional correlation of the estimators that impose slope homogeneity (see Table 3 above); we apply a CCEMG estimator that controls for common unobserved factors that conjointly influence the realization of the dependent variable. The estimator controls for CD and time-variant unobservables with heterogeneous impact across the Laender. This is achieved by the inclusion of cross-section averages of the dependent and independent variables that are commonly referred to as the "Mean Group" in the literature. However, the estimated coefficients on the cross-section-averaged variables are not interpretable in a meaningful way (only the sign and significance).

We show the long-run regressions in the specifications (7)-(12) in Table 4. We find a significant effect of common unobserved factors that jointly influence fiscal policy in the different West German Laender between 1950 and 2015. The factor, which is significant for the panel as a whole and in each cross-section, is the mean of the dependent variable. This is evidence in support of the CD of fiscal policy of the German Laender. We expected a positive and significant sign since we use post-fiscal-equalization data. In addition, the coefficient of the averaged lagged debt-to-GDP across all Laender is significant and negative. This is evidence in support of a surplus reducing effect for each Land if on average the last fiscal year saw an increase of debt across all West German Laender. At the same time, some West German Laender react significantly to increases of their own debt-to-GDP ratio by increasing surplus. Therefore, we have evidence to conclude that the fiscal reaction is not uniform across Laender.

If we compare the reaction coefficient of the West German Laender to increases of initial debt between 1950-2015 (Table 4) with the reaction coefficients in all German Laender from 1991-2015 (Table 5), we find higher coefficients after 1991 across 11 out of 12 specifications. This has several reasons.

First, the debt-to-GDP ratio has declined in the majority of the East German Laender as shown in Figure 1 (Thuringia, Saxony). Debt declines across the recent ten years to all-time lows since 1991 in the East, while some West German Laender (including Berlin and Bremen) share a positively sloped debt path and larger variances of surplus (Berlin, Bremen and Saarland). This is supported by empirical evidence from the state specific estimations (Table 2). If we further explore the CCEMG estimations, we find positive and significant reaction coefficient in most country-specific estimations across the East German Laender. The parameters are higher compared to the significant positive coefficients in, e.g., Bayern, Baden-Wuerttemberg and Hamburg. Note that we do not find

any significant reaction in most West German States (including Berlin, Bremen, Saarland and Rhineland-Palatinate) in the country specific estimations.

Table 5: FRF Evidence for German Laender, 1991-2015

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Lagged debt to GDP	0.063*	0.070*	0.049	0.045**	0.050**	0.042**
	(0.035)	(0.036)	(0.031)	(0.017)	(0.018)	(0.016)
YVAR		-2.406**			-1.205	
		(1.116)			(0.873)	
GVAR		-0.173			-0.232**	
		(0.175)			(0.094)	
Output gap			1.233***			0.515***
			(0.277)			(0.179)
Expenditure gap			-0.370*			-0.275**
			(0.202)			(0.100)
Lagged Primary surplus to GDP				0.705***	0.696***	0.687***
				(0.041)	(0.037)	(0.036)
Number of observations	394	394	394	390	390	390
Within R-squared	0.2721	0.3042	0.3330	0.6598	0.6742	0.6725
Variables	(7)	(8)	(9)	(10)	(11)	(12)
Lagged debt to GDP	0.103***	0.104***	0.107***	0.099***	0.110***	0.092***
	(0.028)	(0.023)	(0.023)	(0.025)	(0.024)	(0.020)
YVAR		-0.231			0.464	
		(4.313)			(5.049)	
GVAR		-0.518***			-0.484***	
		(0.083)			(0.086)	
Output gap			-0.098			0.178
			(0.943)			(1.014)
Expenditure gap			-0.566***			-0.571***
			(0.097)			(0.090)
Lagged Primary surplus to GDP				0.165***	0.036	-0.011
				(0.064)	(0.052)	(0.061)
Cross-section average of primary surplus to GDP	0.996***	0.986***	0.975***	0.933***	0.979***	0.934***
	(0.186)	(0.230)	(0.203)	(0.221)	(0.317)	(0.250)
Cross-section average of lagged debt to GDP	-0.069	-0.129*	-0.130*	-0.043	-0.136	-0.102
	(0.046)	(0.072)	(0.071)	(0.036)	(0.087)	(0.074)
Cross-section average of YVAR		-0.606			-1.934	
		(5.537)			(7.857)	
Cross-section average of GVAR		0.467***			0.348**	
		(0.158)			(0.137)	
Cross-section average Output gap			0.089			-0.184
			(0.943)			(0.998)
Cross-section average of Expenditure gap			0.553***			0.408***
			(0.202)			(0.127)
Cross-section average of lagged Primary surplus to GDP				-0.136	-0.017	0.063
				(0.183)	(0.168)	(0.172)
Constant	-0.002	0.003	0.002	-0.010	0.000	-0.003
	(0.010)	(0.012)	(0.011)	(0.007)	(0.012)	(0.011)
Number of observations	394	394	394	390	390	390

Notes: See Table 4.

Second, economic growth rates have been smaller and with a lower variation after 1991. This can be inferred from the insignificant output gaps after 1991 in the cross-correlated effects panel

estimations (Table 5, columns 8, 11). Before 1991, both are significant and negative in the West German Laender mean group. Note that we find a negative effect of lagged debt-to-GDP in the mean group for the panel as a whole. This is evidence for a high persistence of debt dynamics in the panel. This factor has remained unobserved in other papers and underlines the importance of controlling for cross-correlated effects from mean group estimations.

To sum up, the results for the period from 1991-2015 show that the German Laender have significantly reacted to increasing debt levels. We do not find differences between OLS [(4)-(6)] and CCEMG [(7)-(12)] estimators with regard to the significance of the main explanatory variable, the reaction coefficient on lagged debt-to-GDP. We reject fiscal sustainability in only one specification. Due to omitted variable bias of this single specification, we do not cast doubt on this difference and conclude that on average the German Laender meet a significant and positive fiscal reaction to increased debt levels.

5 CONCLUSION

Sub-federal public finances are crucial for assessing fiscal sustainability of sovereign states. In the age of today's big governments, sub-federal government budgets contribute a notable share to general government finances. Moreover, fiscal equalization schemes interconnect public finances between the federal and state levels by horizontal and vertical transfer schemes. However, studies on fiscal sustainability have primarily analyzed general government or federal public finances.

Our paper instead focuses on the sustainability of public finances in the German federal states (Laender). We test whether politicians react to increasing debt levels by increasing primary surpluses over the long run from 1950 to 2015 for West Germany and for reunified Germany between 1991 and 2015.

Previous studies on fiscal sustainability of the German Laender have three notable shortcomings: First, test results have not been without ambiguity since so far the time-period covered by the univariate analyses is relatively short. Second, most univariate analyses do not specify their modeling in line with data and time series characteristics which is especially meaningful for the modeling of fiscal datasets in federal states. Third, most studies remain tacit about their econometric modeling of the different specifications of the FRF.

In contrast to most studies in the field, we document the procedure of the econometric modeling of FRF. We describe and document the characteristics of the fiscal time series for each cross-section and for the panel as a whole, respectively, and design the econometric model according to the data characteristics. From our point of view, we offer a transparent method to overcome model misspecification in the field of fiscal sustainability, as we inform the econometric modeling by evidence-based time series characteristics. The benefit of this strategy is twofold: First, we gain insights into fiscal policy of each Land (time series analysis) and in the dynamics of all Laender together before we choose and specify the model and estimate panel regressors. Second, we disclose every step of the modeling in a transparent manner.

To sum up, we find that fiscal policy on the level of the German Laender only partly meets fiscal sustainability criteria. We conclude that politicians have significantly reacted to increasing debt levels by increasing budget surpluses only since 1991. Long time-series evidence is not in support of a significant and positive reaction of increasing initial debt levels in the West German Laender panel between 1950 and 2015. Since a significant amount of total debt is owed by the Laender (with a share of about 40 percent of the total debt), we conclude that unsustainable public finances at the sub-federal level can seriously endanger general fiscal sustainability of the Bund – the most important sovereign of EMU. If the role of Germany as the most trusted debtor of EMU should be preserved, the Laender should keep up with their debt-brake requirements and balance their budgets until the year 2020 and beyond.

6 References

Auerbach, A.J. (2016). Long-Term Fiscal Sustainability in Advanced Economies. *Asia & the Pacific Policy Studies*, 3(2), 142-154.

Barro, R.J. (1981). Output Effects of Government Purchases. *Journal of Political Economy*, 89(6), 1086-1121.

Barro, R.J. (1986). US Deficits since World War I. *Scandinavian Journal of Economics*, 88, 195-222.

Bohn, H. (1995). The Sustainability of Budget Deficits in a Stochastic Economy. *Journal of Money, Credit and Banking*, 27(1), 257-271.

Bohn, H. (1998). The Behavior of US Public Debt and Deficits. *Quarterly Journal of Economics*, 113(3), 949-963.

Bohn, H. (2007). Are Stationarity and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint? *Journal of Monetary Economics*, 54(7), 1837-1847.

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

Bundesminsterium der Finanzen (BMF) (2017). EU Review of Germany's Public Finances: Targets Met, Challenges Remain. Monthly Report of the Federal Ministry of Finance April 2017, Available at: http://www.bundesfinanzministerium.de/Content/EN/Downloads/2017-04-27-stability-programme.pdf?__blob=publicationFile&v=3 .

Burret, H.T., Feld, L.P., and Köhler, E.A. (2016). (Un-) Sustainability of Public Finances in German Laender: A Panel Time Series Approach. *Economic Modelling*, 53, 254-265.

Burret, H.T., Feld, L.P., and Köhler, E.A. (2017). Fiscal Sustainability of the German Laender: Time-Series Evidence. *FinanzArchiv: Public Finance Analysis*, 73(1), 103-132.

Buti, M., and Carnot, M. (2016). Fiscal Space and Low Interest Rates: A Eurozone Perspective. *Vox CEPR's Policy Portal*, Available at: http://voxeu.org/article/fiscal-space-and-low-interest-rates-eurozone-perspective .

Checherita-Westphal, C., and Ždarek, V. (2017). Fiscal Reaction Function and Fiscal Fatigue: Evidence for the Euro Area. *ECB Working Paper Series* (No. 2036), Frankfurt.

Claeys, P. (2006). Policy Mix and Debt Sustainability: Evidence from Fiscal Policy Rules. *Empirica*, 33(2), 89-112.

Claeys, P., Ramos, R., and Suriñach, J. (2008). Fiscal Sustainability across Government Tiers. *International Economics and Economic Policy*, 5(1-2), 139-163.

D'Erasmo, P., Mendoza, E.G. and Zhang, J. (2016). What is a Sustainable Public Debt?, in: J.B. Taylor and H. Uhlig (ds.), *Handbook of Macroeconomics, Vol. 2*, Elsevier, Amsterdam, 2493-2597.

Durlauf, S.N., Navarro, S. and Rivers, D.A. (2016). Model Uncertainty and the Effect of Shall-Issue Right-to-Carry Laws on Crime. *European Economic Review*, 81, 32-67.

Fincke, B. and Greiner, A. (2011). Debt Sustainability in Germany: Empirical Evidence for Federal

States. *International Journal of Sustainable Economy*, 3, 235-254.

Herzog, B. (2010). Anwendung des Nachhaltigkeitsansatzes von Bohn zur Etablierung eines Frühindikators in den öffentlichen Finanzen. Beitrag zur aktuellen Debatte der Föderalismuskommission II. *Kredit und Kapital*, 43(2), 183-206.

Kitterer, W. (2007). Nachhaltige Finanz- und Investitionspolitik der Bundesländer. *Perspektiven der Wirtschaftspolitik*, 8(S1), 53-83.

Kitterer, W. and Finken, J. (2006). Zur Nachhaltigkeit der Länderhaushalte – eine empirische Analyse. *Finanzwissenschaftliche Diskussionsbeiträge / Finanzwissenschaftliches Forschungsinstitut an der Universität zu Köln*, (No. 06-7), Cologne.

Pesaran, M.H. (2004). General Diagnostic Tests for Cross Section Dependence in Panels. *CESifo Working Paper* (No. 1229), Munich.

Pesaran, M.H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4), 967-1012.

Potrafke, N. and Reischmann, M. (2015). Fiscal Transfers and Fiscal Sustainability. *Journal of Money, Credit and Banking*, 47(5), 975-1005.

Reinhart, C.M. and Rogoff, K.S. (2010). Growth in a Time of Debt. *American Economic Review*, 100(2), 573-578.

Schularick, M. and Taylor, A.M. (2012). Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008. *American Economic Review*, 102(2), 1029-1061.

Theofilakou, N. and Stournaras, Y. (2012). Government Solvency and Financial Markets: Dynamic Panel Estimates for the European Monetary Union. *Economics Letters*, 115(1), 130-133.

Westerlund, J. (2007). Testing for Error Correction in Panel Data. *Oxford Bulletin of Economics and Statistics*, 69, 709-748.