Long-term Effects of Fiscal Stimulus and Austerity in Europe

Article *in* Oxford Bulletin of Economics and Statistics · December 2018

DOI: 10.1111/obes.12287

3 authors, including:

Christoph Paetz
Institut für Makroökonomie und Konjunkturforschung (IMK)
9 PUBLICATIONS **148** CITATIONS

SEE PROFILE

OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 0305–9049
doi: 10.1111/obes.12287

Long-term Effects of Fiscal Stimulus and Austerity in Europe*

Sebastian Gechert,† Gustav Horn† and Christoph Paetz†

†*Macroeconomic Policy Institute (IMK), Duesseldorf, 40476, Germany (e-mail: sebastian-gechert@boeckler.de, gustav-horn@boeckler.de, christoph-paetz@boeckler.de)*

Abstract

We analyze whether there are negative (positive) long-term effects of austerity measures (stimulus measures) on potential output growth. Based on the approach of Blanchard and Leigh (2013) and Fatás and Summers (2018) and using a novel data set of narratively identified fiscal policy shocks, we estimate the impact of these shocks on potential output. We robustly find a considerable underestimation of multiplier effects and their persistence for most European countries in the early years after the financial crisis and subsequent Euro Area crisis. We conclude that fiscal consolidation was badly timed and thus not only deepened the crisis but may have caused evitable hysteresis effects.

I. Introduction

Output in many European countries has long remained below precrisis potential. The recession took considerably longer and was much deeper compared to past downturns and the recovery was comparably weak. Forecasts by the European Commission (EC) or the International Monetary Fund (IMF) in the aftermath of the crisis assumed a quick recovery to previous trends, but had to be revised downwards several times. These revisions most strikingly concerned not only GDP but also potential GDP forecasts. Figures 1 and 2 show repeated over-optimism of GDP and potential output forecasts for the EU as a whole and Greece as an extreme example.[1]

The persistent and systematic forecast errors call into question the structure and assumptions of the forecasting models employed. Clearly, the financial crisis and the subsequent crisis of the Euro Area were extreme events, whose dynamics and channels of impact might be quite different from more tranquil times. A number of influential factors that unexpectedly drove the severity of the crisis have been discussed, among them the fragility

JEL Classification numbers: E62, H68.

*We thank Achim Truger, Antonio Fatas, Karel Havik, Katja Heinisch, Oana Furtuna, Rudolf Zwiener, Wouter van der Wielen and two anonymous referees for helpful discussions and data access. All remaining errors are our own.

[1] Apart from Germany in all other major European countries potential output growth rates decreased considerably and are now below precrisis figures. Potential output estimates were revised downwards both for forecasted and past values in most European countries, apart from Spain.

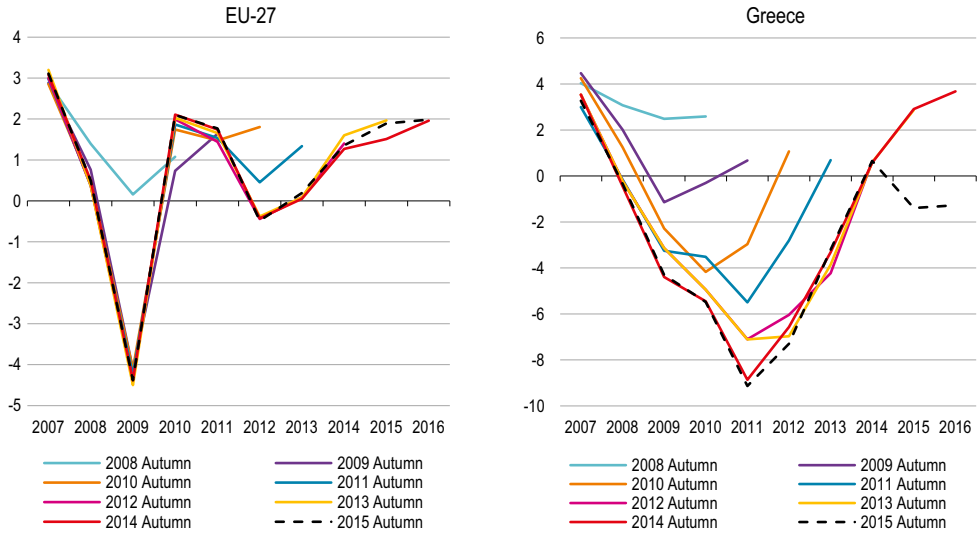

Figure 1 Vintages of GDP growth rate forecasts for the EU-27 and Greece, in %, 2007–16
Source: Ameco, Firstrun database 'A dataset of fiscal variables', own illustration.

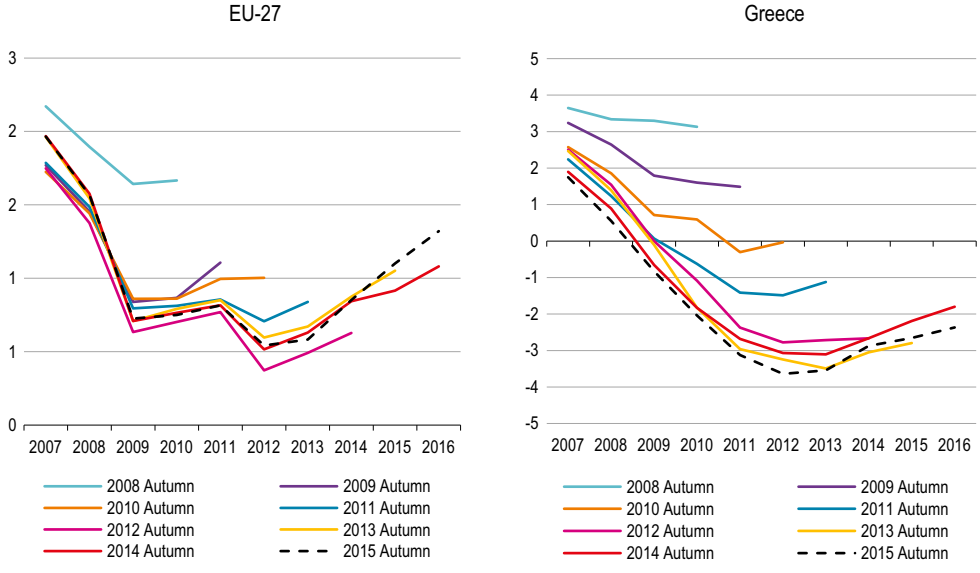

Figure 2 Vintages of potential GDP growth rate forecasts for the EU-27 and Greece, in %, 2007–16
Source: Ameco, Firstrun database 'A dataset of fiscal variables', own illustration.

of the financial system, private sector deleveraging, increased uncertainty of private agents, current account imbalances, monetary policy constraints, sustainability of public finances or the impact of discretionary fiscal policy.

In the present paper, we focus on fiscal policy, while we take into account the others. We ask whether the post-2009-shift towards fiscal consolidation had an unexpected substantial negative and persistent impact on GDP and potential output, in particular in the EU and

the Euro Area, which could be a major explanatory factor for the second recessionary dip that followed in due course and the persistent gap to precrisis GDP trend and unemployment levels. This is equivalent to asking whether there was an underestimation of fiscal multipliers and, more importantly, their persistence.

Since the crisis, there has been an intense debate and a growing literature on short-run fiscal multiplier effects (Hebous, 2011; Mineshima, Poplawski-Ribeiro and Weber, 2014; Gechert, 2015). Expansionary confidence effects of austerity have been discussed widely[2] (Alesina and Ardagna, 2010) but have been found to be rather special cases (Perotti, 2011). The general consensus among international institutions now seems to read that austerity reduces growth in the short run, can be particularly harmful during downturns and may even increase public debt-to-GDP ratios in the interim (Cottarelli and Jaramillo, 2012; Furman, 2016).

The long-term effects – although they are much more important in terms of welfare and sustainability of public finances – have attracted far less attention in the empirical literature and remain more controversial, except for the special case of public investment (Bom and Ligthart, 2014). Certainly, robust inference is much harder to achieve for longer horizons, which might explain the lack of evidence. For the few exceptions, the dominant reading seems to be that while austerity brings short-run pain, it provides long-term gain in terms of reduced tax distortions and debt risk (Rogoff, 2012; Born, Müller and Pfeiffer, 2015). DeLong and Summers (2012) on the other hand make the case for hysteresis effects where austerity in a deep slump would be self-defeating even in the long run.

The present paper builds on Blanchard and Leigh (2013) (BL hereafter) and Fatás and Summers (2018) (FS hereafter). BL exploit GDP growth forecast errors for European countries during the 2010–11 period to create a counterfactual of expected policy impact. They then regress these forecast errors on planned consolidation for the same sample in order to test whether the impact of consolidation was underestimated. They find a strong negative correlation between consolidation attempts and output revisions meaning that countries with bigger consolidation plans faced more severe growth disappointments – i.e. multipliers had been underestimated by forecasters. FS confirm the findings of BL with more recent data and extend their method by a second stage, where they regress longer-term potential output forecast errors on the GDP forecast errors that were arguably caused by the underestimation of multiplier effects. The coefficient of this second stage can be interpreted as a measure of persistence of these multiplier effects.

This paper provides two central innovations: (i) We argue that the measure of exogenous fiscal shocks employed by BL and FS, the change in the structural balance, may face endogeneity issues, as its calculation is based on potential output itself. We therefore opt for a narrative measure of the fiscal stance, the discretionary fiscal effort (DFE), as provided by the AMECO database (EC 2013). (ii) We rigorously test the robustness of our

[2] Indeed, official statements by leading policy makers at the time seemed to assume a strong confidence effect of fiscal consolidation that would imply expansionary effects, i.e. negative multipliers. 'My understanding is that an overwhelming majority of industrial countries are now in those uncharted waters, where confidence is potentially at stake. Consolidation is a must in such circumstances' (Trichet, 2010). 'All the eurozone governments need to demonstrate convincingly their own commitment to fiscal consolidation so as to restore the confidence of markets, not to speak of their own citizens' (Schäuble, 2010).

findings and those of FS in terms of omitted variable biases, outliers, alternative estimation techniques, data sources and sample periods.

We find a significant underestimation of fiscal multipliers of about 0.8 units on average, which is strong, but still somewhat less pronounced than in BL and FS. This would translate into a multiplier effect of about 1.3, given that forecasters likely assumed a multiplier effect of 0.5 in their forecasts. These effects have a permanent impact as measured by five-year-ahead forecasts, making a strong case for hysteresis effects of fiscal consolidations and expansions during a deep recession. Our findings are robust to a large set of perturbations. Yet, as a plausible qualification, we find a weakening of the effects in later crisis years, in line with the slowdown of consolidation, possible learning effects of forecasters (Górnicka *et al.*, 2018) or regime-dependent multiplier effects (Baum, Poplawski-Ribeiro and Weber, 2012; Auerbach and Gorodnichenko, 2012). Moreover, some Eastern European countries are influential outliers that weaken the relation to some extent. The effects seem to be stronger for spending than for revenue shocks. We conclude that the European austerity measures were more harmful than expected even in the longer-term, while countries with a more expansionary fiscal stance fared better as this may have prevented hysteresis effects. Accordingly, fiscal consolidation was badly timed which may have even had long-term negative consequences.[3]

The remainder of the paper is organized as follows. Section II explains our approach and data set. Section III presents the baseline results. Section IV checks the robustness of these findings. The final section concludes.

II. Method and data

First stage: Underestimation of fiscal multipliers

In line with BL, we regress the forecast error (*fe*) of cumulated GDP growth for the years of 2010 (= *t*) and 2011 for country *i* on planned (*f*) fiscal consolidation for the very same period:

$$\Delta Y^{fe}_{i,t:t+1} = \alpha + \beta \Delta F^{f}_{i,t:t+1|t} (+X_i\theta) + \varepsilon_{i,t:t+1|t} \tag{1}$$

where

$$\Delta Y^{fe}_{i,t:t+1} \equiv \Delta Y_{i,t:t+1} - \Delta Y^{f}_{i,t:t+1|t} \tag{2}$$

is the forecast error of GDP as given by the difference between current-vintage figures of the cumulated growth rate of GDP over 2010 and 2011 and its forecast in the vintage of spring 2010. This figure is negative for most countries during this period. $\Delta F^{f}_{i,t:t+1|t}$ is a measure of planned fiscal consolidation as a percentage of GDP over the same two-year period. X_i marks a set of control variables that are likely alternative explanations for the forecast errors besides consolidation. $\varepsilon_{i,t:t+1|t}$ is an *iid* error term. Two-year episodes are used to allow for lagged effects.

[3] Corsetti *et al.* (2013), in a New Keynesian model, point to an expectations channel according to which bringing down severe sovereign debt risk may outweigh the negative growth effects of spending cuts. Cottarelli and Jaramillo (2012) on the other hand make the point that in situations when multipliers are high, consolidation efforts do not succeed and might even be detrimental to lowering debt risk. In our analysis, we try to control for sovereign debt risk through various proxies.

The rationale is the following: Using the forecast error of GDP exploits the deviation of the actual data from a counterfactual scenario given by the expectations of forecasters, based on their information set, assumptions and model of the economy at the time, where channels work as expected by these experts. Regressing this forecast error on planned fiscal consolidation reveals, as to whether the impact of these consolidation plans was over- or underestimated. If the multiplier effect assumed in the forecasting model is correct, β should not deviate significantly from zero. The multiplier effect would be as expected.[4] A negative and significant β however, would imply that countries with a more ambitious consolidation plan had bigger growth disappointments during that period, and vice versa. The multiplier effect would have been underestimated

Second stage: Persistence of multiplier effects

With respect to welfare and sustainability of public finances, the long-term impact of the fiscal measures is key. In line with FS, we measure these long-term effects by 5-year-horizon forecast errors of cumulated potential output growth. For inference, we build a two-stage least squares (TSLS) framework, where the exercise of BL is considered as the first stage, measuring the growth disappointments as caused by the stronger than expected impact of fiscal consolidation:

$$\Delta \hat{Y}^{fe}_{i,t:t+1} = \alpha + \beta \Delta F^{f}_{i,t:t+1|t} \tag{3}$$

The fitted values of the first stage – interpreted as the unexpected GDP change due to a stronger than expected impact of fiscal consolidation – then enter the second stage, where the forecast error of potential output is regressed on these fitted values:

$$\Delta \mathrm{Pot} Y^{fe}_{i,t:t+5} = \gamma + \delta \Delta \hat{Y}^{fe}_{i,t:t+1} (+X_i \pi) + \omega_{i,t:t+1|t} \tag{4}$$

In a nutshell, persistent effects of cyclical changes in output are estimated by using fiscal policy shocks as instruments for these cyclical changes (Fatás and Summers, 2018). The relevant coefficient δ can therefore be interpreted as a measure of persistence of changes in output that are caused by changes in the fiscal stance. If $\delta = 1$, the multiplier effect would be fully persistent and growth disappointments would carry on one-to-one to the long-run. For a fiscal consolidation shock in a standard New Keynesian model, δ should be smaller than one and approach zero in the medium run, except for a cut in public investment that might drag down aggregate supply conditions. Of course, potential output figures usually follow persistent changes in GDP quite closely and might thus not be a perfect metric to investigate structural changes in output (Gechert, Rietzler and Tober, 2015).[5] However, a permanent effect on GDP after 5 years still runs counter to conventional assessments of

[4] The IMF (2010, p. 94) WEO of October 2010 has estimates of fiscal multipliers of 0.5 on average, based on the IMF's GIMF model that is likely to inform forecasters. BL points to some further evidence in this direction. According to the European Commission (2012, p. 41) European Economic Forecast, multipliers from the EC's QUEST model, which likely informs forecasters, range between 0.2 and 0.8 depending on the specific measure and are about 0.4 on average. Also the meta-analysis of Gechert and Rannenberg (2018) finds average multipliers of about 0.5 in their sample of precrisis studies. Of course, such averages mask likely heterogeneity of the various fiscal measures, but they may suffice for the broadly defined change in the fiscal surplus that we employ here.

[5] For a critical review see the conference contributions at http://ec.europa.eu/economy_finance/events/2015/20150928_workshop/index_en.htm

the persistence of demand shocks and is much more in line with theories and evidence of hysteresis (Fatás, 2000; Logeay and Tober, 2006; DeLong and Summers, 2012; Sturn, 2014).

Identification of consolidation shocks

When estimating the impact of fiscal policy, identification of exogenous fiscal shocks is crucial. Three main concerns are usually discussed in the literature: (a) Since the budget is highly sensitive to business cycle fluctuations via automatic stabilizers, estimation based on headline budgetary figures would be prone to an endogeneity bias. (b) Even discretionary measures may be immediate reactions to macroeconomic circumstances (e.g. countercyclical policies) and thus reverse causality may apply. (c) Agents may anticipate fiscal policy measures due to early announcement and hence react prior to implementation (e.g. in the case of a tax hike), outside the information set of the econometrician.

(a) BL and FS rely on changes in the structural balance (*SB*) which is an established measure of the fiscal stance. It is derived from the actual budget balance by subtracting a cyclical component, based on assumptions of automatic stabilizers and the output gap, as well as one-off events. We argue that the structural balance still faces a likely endogeneity bias when it comes to measuring its impact on potential GDP forecast errors. This is because the structural balance depends on the assessment of potential output itself. To see this, consider the situation in 2010 where potential GDP was forecasted too optimistic in a phase of severe slack. At first, there would be a large measured output gap and forecasting models would estimate the output gap to close with high speed under such circumstances as they include a closing rule effective within the forecasting horizon (Havik *et al.*, 2014). Any consolidation effort that improves the headline budget balance is then largely counted as cyclical, with only a smaller share left to be counted as structural consolidation. That is, if we consider two identical countries with the same true structural consolidation effort, and one country is hit by a stronger negative GDP shock than the other, the former would have a larger forecast error of potential GDP and a lower measured improvement in the SB. This will lead to inflated coefficients β and δ, measured with lower precision at that.[6]

In light of these issues, we opt for an alternative measure of the fiscal stance, namely the DFE as published by the AMECO database. It is available for EU27 countries on an annual basis since 2010.[7] The DFE is essentially a mixed method for determining the discretionary fiscal stance. Changes on the revenue side are entirely based on a narrative account of fiscal shocks where the expected budgetary impact of factual law changes and other measures is recorded. On the expenditure side, where substantial discretionary changes happen at all levels of government and a full narrative record would be too costly, the DFE is calculated as the gap between public spending growth and a smooth trend output growth, while excluding changes in cyclical spending components (in particular

[6] When potential growth turns out lower than expected and is revised downward, so would the structural share of the consolidation effort need to be revised upward. However, SB enters the regression without such revision. Note that such revisions would be required due to pure technical dependence of the calculation of the structural balance on potential output figures, and must not be confused with revisions due to truly more ambitious consolidation efforts.

[7] AMECO publishes the DFE in nominal terms of national currency. For our econometric analysis below we express the discretionary changes in percentage of potential GDP just as the structural balance.

unemployment spending). The DFE thus avoids the dependence on estimated potential output figures and uncertain budget elasticities.

The DFE shock series has been argued to be more robust in estimating fiscal multipliers.[8] In the next section, it will be shown that this is indeed the case for our exercise. In line with the arguments above, we find that the cumulated 2010–11 DFE is more positive on average than the respective change in SB ($\mu_{DFE} = 2.46\,pp$, $\mu_{SB} = .53\,pp$) and is moreover much more dispersed ($\sigma_{DFE} = 3.42$, $\sigma_{SB} = 1.68$), while the two are still highly correlated ($cor_{DFE,SB} = .74$). This could speak of an attenuation of the SB measure towards zero.

(b) Separating truly exogenous from endogenous legislations is an issue that is addressed by extensive country studies collecting data similar to the DFE, but also looking at the motivation of single law changes (Romer and Romer, 2010; Cloyne, 2013). We do not have enough information to make such a separation for the DFE. The narrative studies usually find that not controlling for endogenously motivated law changes tends to downward-bias the multiplier estimates (Mertens and Ravn, 2014). Thus, we regard our estimates as conservative in this regard. In any case, note that the SB approach does not address this issue either. Even detailed country-studies find it hard to give a clear judgment regarding the motivation of single law changes and the identified shocks may be susceptible to measurement error, which for example Mertens and Ravn (2014) try to address in a proxy SVAR framework. Generally, existing measures of fiscal shocks, as ours, are only second best proxies, but a first best solution is yet to be discovered in the literature (Caldara and Kamps, 2017).

(c) Anticipation bias may arise when econometricians draw inference from ex post realized data while agents may have had additional information from preannounced policies and reacted in advance. This should be less of a concern for our framework, since we create a counterfactual of realized data against expert forecasts that should be informed about policy announcements at least as well as the general public. In that sense, anticipation may only be of concern if agents on average were better informed about policy actions than forecasters, which is rather unlikely.

Further data

In our baseline, we stick to IMF World Economic Outlook (WEO) forecasts for GDP and potential output and use the vintage of spring 2016 *vis-à-vis* the spring 2010 forecast for the calculation of forecast errors.[9] Importantly, comparing data of different vintage years requires correction for changes of the base year, accounting rules and reassessments of past potential output figures.[10] The second stage of our model uses $t + 5$ forecasts for potential output, as given by unpublished vintages of the IMF WEO.[11] In the baseline sample, we focus on European countries, but due to missing data end up with 22/21

[8] See section 2 in Carnot and de Castro (2015) for an elaborate discussion on the DFE measure of identification.

[9] BL compare the IMF autumn 2012 forecast to the spring 2010 forecast.

[10] See Appendix A for a more detailed description of the computation of forecast errors for GDP and potential GDP.

[11] We are grateful to Antonio Fatás for providing us with the WEO data and files for replication of the FS results.

observations.[12] Due to this small sample, section IV includes a battery of robustness checks for the baseline estimates. First, we include various alternative explanatory factors to control for omitted variable biases. Data for sovereign CDS spreads, precrisis household debt-to-GDP ratios and precrisis current-account-to-GDP ratios are obtained from the BL data set.[13] Second, we also run our model using European Commission forecasts. The forecast vintages are obtained from a data set by the FIRSTRUN project,[14] which collects vintages of the AMECO data set; moreover, we use unpublished $t+4$ EC forecasts of potential output.[15] The EC data allows to extend the sample to the whole EU27 and thus some additional Eastern European countries that are absent from the IMF data set. The third class of robustness checks extends the time horizon by applying a moving window and panel data analysis to increase the number of observations, where we use different spring vintage sets from the IMF and the EC data, respectively, and compare them to the vintage of spring 2016 to obtain our forecast errors.

III. Estimation results

First stage: Underestimation of fiscal multipliers

First, in Table 1, we replicate the BL results by using IMF WEO data and the change in the SB as our fiscal measure. In Table 2 we use the DFE instead.

Column (1) of Table 1 shows the result of the replication. The original finding of BL, a significant underestimation of fiscal multipliers by about 1.1, is even reinforced with $\beta \approx -1.3$ Is the latter effect driven by the assessment of spending or taxation? Such data are not directly available for structural balance components. In line with BL, in column (2), we split the structural balance into spending (G) and revenues (T), where $SB = T - G$. In terms of cyclical adjustment, we assume that government spending is insensitive and use its actual value G, while calculating cyclically adjusted revenues as the residual $T = SB + G$.[16] It turns out that the negative impact of government spending cuts was more strongly underestimated than the one from tax hikes. This is consistent with evidence from the meta regression of Gechert and Rannenberg (2018) who show that in particular spending, multipliers increase during downturns.

A natural objection to the validity of the effects in columns (1) and (2) is the small sample size and the likely dependence on influential outliers. Using a quantile regression instead, does only minimally alter the coefficient (column (6)). Likewise, in column (3), we exclude those countries in our sample that were under a bailout programme (Greece, Ireland

[12] For SB, the sample comprises Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Malta, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland and the United Kingdom. For DFE, we have Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden and the United Kingdom.

[13] Further control variables are taken from the respective spring 2010 forecast, in line with BL.

[14] http://www.firstrun.eu/research/data/

[15] Courtesy of European Commission forecasting staff.

[16] This assumption is questionable since, for example, government spending on unemployment benefits and old age benefits is sensitive to the business cycle (Price, Dang and Guillemette, 2014). DFE measures are published separately for spending and revenues, so we do not need such an assumption there.

TABLE 1

First stage: Underestimation of multipliers with structural balance (SB)

Endogenous: forecast error of GDP 2010–11						
OLS	*(1) SB*	*(2) GT*	*(3) NOPROG*	*(4) ADVA*	*(5) EURO*	*(6) QUANT*
β	−1.341**		−0.942***	−0.632	−1.534**	−1.272***
	(0.53)		(0.243)	(0.614)	(0.578)	(0.306)
β^G		−1.699***				
		(0.477)				
β^T		−0.967**				
		(0.371)				
Const	1.15***	1.223***	1.101***	0.919*	1.34***	0.856
	(0.402)	(0.36)	(0.374)	(0.493)	(0.393)	(0.62)
N	22	22	19	31	14	22
Adj. R^2	0.4755	0.6024	0.3307	0.0749	0.5763	

TABLE 2

First stage: Underestimation of multipliers with discretionary fiscal effort (DFE)

Endogenous: forecast error of GDP 2010–11							
OLS	*(1) DFE*	*(2) GT*	*(3) G*	*(4) T*	*(5) NOPROG*	*(6) EURO*	*(7) QUANT*
β	−0.861***				−0.934*	−0.875***	−0.874***
	(0.055)				(0.498)	(0.052)	(0.055)
β^G		−0.928	−1.906***				
		(0.761)	(0.209)				
β^T		−0.812		−1.462***			
		(0.566)		(0.122)			
Const	2.75***	2.761***	2.84***	2.552***	2.7***	2.835***	2.573***
	(0.489)	(0.516)	(0.553)	(0.522)	(0.713)	(0.573)	(0.669)
N	21	21	21	21	18	16	21
Adj. R^2	0.6983	0.6816	0.6639	0.6721	0.1306	0.7508	

and Portugal). The effect is somewhat muted but still economically and statistically highly significant. Column (4) shows the results for a widened sample of advanced countries. Interestingly, β is not statistically significant any more. Arguably, many European countries were in a deeper crisis and bound to common currency and monetary policy at its zero lower bound during this phase. In line with this assessment, narrowing the sample to Euro Area countries in column (5) even slightly increases the effects.

In Table 2, using the DFE, the qualitative results are confirmed. However, the effect is somewhat smaller. This is in line with our reasoning above: the effects as measured by SB might be somewhat upward biased due to its possible endogeneity with growth forecast errors. Separating expenditures and revenues, which are directly available for the DFE, in column (2) gives consistent, though insignificant results; but the wide standard errors may not be trusted due to multicollinearity: The correlation of the series is extremely high ($cor_{DFEG,DFET} = 0.92$). Moreover, as shown in columns (3) and (4), including G and T, one at a time, strongly inflates the coefficients. Of course, the coefficients of (3) and (4) must

TABLE 3

Second Stage: Persistence of multiplier effects (SB)

Endogenous: forecast error of potential GDP 2010–15						
TSLS	*(1) SB*	*(2) GT*	*(3) NOPROG*	*(4) EURO*	*(5) QUANT*	*(6) DIR*
δ	1.005**	1.046***	1.296**	1.065***	1.401**	
	(0.402)	(0.289)	(0.544)	(0.387)	(0.647)	
η						−1.348
						(1.013)
Const	−3.521**	−3.537***	−4.016**	−3.548***	−3.834**	−2.365**
	(0.869)	(0.819)	(0.861)	(1.114)	(1.356)	(1)
N	22	22	19	14	22	22
Adj. R^2	0.5813	0.5813	0.3346	0.6866		0.1218
1st stage F	6.3952	6.3621	15.036	7.0449		

not be trusted as they pick up the influence of the omitted counterpart of the budget, but they still show that the coefficients of column (2) could be significant if multicollinearity was absent.

In general, we can reconfirm the substantial underestimation of fiscal multipliers during the early stages of the Euro Area crisis as found by BL. Using a superior measure of the fiscal stance, the effect however, is more in a range of 0.8–0.9. Together with the well-documented assumption that IMF forecasters implicitly used a multiplier effect of around 0.5, actual multipliers for the crisis period under investigation should lie in a range of 1.3–1.4. This is in line, but on the lower end of findings of ZLB effects in standard macroeconomic models (Christiano, Eichenbaum and Rebelo, 2011; Eggertsson, 2011; Woodford, 2011).[17]

Persistence of fiscal multipliers

Investigating the persistence of multipliers, in Table 3 column (1), we first replicate the result of FS. The factor of persistence is close to one, which could be interpreted such that the GDP losses caused by fiscal consolidation became permanent, at least given the currently available information set on a five-year horizon.

Again, the results of the TSLS estimation are robust to the changes already discussed for the first stage in Tables 1 and 2. Splitting the structural balance in spending and revenue components only minimally changes the estimated persistence (column (2)). So does a sample based on Euro Area countries (columns (4)). Down-weighting outliers by excluding program countries or using quantile regressions even reinforces the persistence (columns (3) and (5)). Interestingly, the results are not robust to using a direct regression like

$$\Delta \text{Pot} Y^{fe}_{i,t:t+5} = \zeta + \eta \Delta F^{f}_{i,t:t+1|t} + \vartheta_{i,t:t+1|t}$$

of the $t+5$ potential output forecast error on the fiscal stance in column (6). The coefficient, which should be $\eta = \beta \times \delta$ has the expected sign and is large, yet is not statistically significant.

[17] Boneva, Braun and Waki (2016) question the robustness of the high multiplier effects in these models and show that they are very sensitive to the respective parameterization and solution method.

TABLE 4

Second Stage: Persistence of multiplier effects (DFE)

Endogenous: forecast error of potential GDP 2010–15						
TSLS	*(1) DFE*	*(2) GT*	*(3) NOPROG*	*(4) EURO*	*(5) QUANT*	*(6) DIR*
δ	1.236***	1.234***	1.319*	1.28***	1.216***	
	(0.072)	(0.073)	(0.689)	(0.086)	(0.233)	
η						−1.065***
						(0.1)
Const	−3.914***	−3.912***	−4.459*	−3.303***	−4.304***	−0.515
	(0.745)	(0.747)	(1.523)	(0.872)	(0.704)	(1.257)
N	21	21	18	16	21	21
Adj. R^2	0.6807	0.6807	0.6109	0.7303		0.3203
1st stage F	244.36	115.39	3.5254	279.43		

The results become more robust and persistence is even a little bit stronger when using the DFE measure of fiscal stance in Table 4. Moreover, the instrument seems quite strong judging from the first stage F statistics. In general, while the estimated multiplier effect is somewhat lower on impact when using the DFE, it is super-persistent and increases over the 5-year horizon by a factor of 1.25, or 1.05 per year. This time, the direct regression in column (6) is highly significant.

IV. Further robustness tests

Controlling for alternative explanations

As discussed in the introduction, there might be other factors at play that explain growth disappointments and that would lead to an omitted variable bias in our simple regressions. As a general note, it is vital to look at control variables that were already in the information set of forecasters to see if their impact was underestimated. Any later realizations of these variables that could have an influence on realized output growth would most likely be prone to reverse causality issues. For example, an increase in sovereign CDS spreads could cause lower growth but could as well be caused by growth disappointments (Cottarelli and Jaramillo, 2012).

Most basically, since we exploit cross-sectional variation, our findings could be challenged by varying optimism and pessimism of forecasts for specific country-year observations that could explain the variation in forecast errors after 2010.[18] Some earlier literature points to politically motivated over-optimism in growth forecasts by national authorities (Jonung and Larch, 2006; Frankel, 2011). If this is the case for our sample, there should be a positive correlation of precrisis and within-crisis forecast errors. However, the correlation coefficient for the average IMF WEO April forecast errors of vintages of 1997–2006, with the 2010 vintage forecast error is close to zero ($\rho = 0.03$). So there is no indication that forecasts for countries with a large negative forecast error in the relevant period 2010–11

[18] We thank an anonymous referee for pointing us to this concern.

TABLE 5

First stage with controls (SB)

Additional control X	β	θ	Const	n	Adj. R^2
(1) fe gdp0709	−1.316 (0.516)**	0.043 (0.143)	1.451 (0.915)	22	0.451
(2) forec gdp1011	−1.174 (0.442)**	0.242 (0.238)	0.409 (0.842)	22	0.473
(3) reces dummy	−1.018 (0.507)*	−2.567 (1.559)	1.437 (0.456)***	22	0.524
(4) fe dummy	−0.972 (0.632)	−2.323 (1.091)**	1.895 (0.604)***	22	0.550
(5) fe sb1011	−1.161 (0.427)**	−0.51 (0.365)	1.553 (0.582)**	22	0.501
(6) trade part cons	−1.402 (0.488)***	2.321 (1.695)	0.964 (0.419)**	22	0.509
(7) sov'n debt09	−1.29 (0.507)**	−0.008 (0.018)	1.632 (1.284)	22	0.451
(8) sb09	−1.095 (0.621)*	0.141 (0.248)	1.689 (1.131)	22	0.454
(9) sov'n cds10q1	−1.199 (0.587)*	−0.003 (0.006)	1.408 (0.639)**	22	0.457
(10) bankcrisis	−1.324 (0.515)**	−0.268 (0.881)	1.262 (0.481)**	22	0.450
(11) private debt 07	−1.312 (0.557)**	0 (0.008)	1.107 (0.955)	21	0.433
(12) ca07	−1.301 (0.685)*	0.013 (0.087)	1.143 (0.395)***	22	0.448

TABLE 6

First stage with controls (DFE)

Additional control X	β	θ	Const	n	Adj. R^2
(1) fe gdp0709	−0.861 (0.06)***	−0.06 (0.103)	2.186 (1.038)**	21	0.693
(2) forec gdp1011	−0.864 (0.139)***	−0.005 (0.305)	2.77 (1.195)**	21	0.682
(3) reces dummy	−0.794 (0.111)***	−0.902 (1.469)	2.733 (0.495)***	21	0.686
(4) fe dummy	−0.702 (0.081)***	−2.353 (0.639)***	3.19 (0.521)***	21	0.770
(5) fe sb1011	−0.906 (0.11)***	0.263 (0.326)	2.35 (0.473)***	19	0.733
(6) trade part cons	−0.857 (0.051)***	2.12 (0.686)***	2.557 (0.465)***	21	0.725
(7) sov'n debt09	−0.806 (0.084)***	−0.015 (0.018)	3.507 (1.183)***	21	0.695
(8) sb09	−0.803 (0.138)***	0.043 (0.198)	2.563 (0.812)***	19	0.725
(9) sov'n cds10q1	−0.966 (0.336)**	0.007 (0.018)	2.344 (0.949)**	20	0.680
(10) bankcrisis	−0.85 (0.053)***	−0.72 (0.85)	3.068 (0.686)***	21	0.694
(11) private debt 07	−0.876 (0.061)***	−0.004 (0.006)	3.17 (1.004)***	20	0.703
(12) ca07	−0.924 (0.125)***	−0.05 (0.094)	2.787 (0.525)***	21	0.690

where generally too optimistic in precrisis years. The same holds for AMECO spring forecast vintages where we have data from 2000 onwards ($\rho = -0.04$).

Tables 5–8 present regression results including various control variables using SB and DFE for the first-stage and second-stage regressions, respectively. Due to low degrees of freedom, we include these controls one at a time. Column β in Tables 5–6 and δ in Tables 7–8 show our parameters of interest, the effects of multiplier underestimation and persistence; column θ and π give the coefficients of the control variables.

In row (1), we ask whether the *under-prediction of the 2008–09 recession* might in fact predict the 2010–11 forecast error. The rationale would be that the persistence of the crisis was underestimated and that the double dip was inevitable though not forecasted. The effect of fiscal consolidation, however, remains intact and the financial crisis forecast error is not significant. This holds true for both SB and DFE for first and second stage. In a similar fashion, in row (2), we control for the *size of the forecasted GDP growth* during the 2010–11 period itself. Maybe countries with strongly negative forecast errors simply had a comparably large GDP growth forecast from the outset that was unrealistic. However,

TABLE 7

Second stage with controls (SB)

Additional control X	δ	π	Const	n	Adj. R^2	1st F
(1) fe gdp0709	0.845 (0.406)**	0.374 (0.263)	−0.715 (1.792)	22	0.575	3.327
(2) forec gdp1011	1.382 (0.403)***	−0.729 (0.275)***	−1.72 (1.102)	22	0.684	3.611
(3) reces dummy	0.93 (0.753)	−0.807 (4.141)	−3.344 (1.616)**	22	0.539	5.003
(4) fe dummy	0.579 (0.952)	−3.6 (4.355)	−1.877 (2.582)	22	0.472	9.202
(5) fe sb1011	1.043 (0.508)**	0.145 (0.636)	−3.679 (1.209)***	22	0.571	4.029
(6) trade part cons	1.135 (0.276)***	6.628 (1.59)***	−4.203 (0.598)***	22	0.708	5.004
(7) sov'n debt09	1.124 (0.472)**	0.025 (0.036)	−5.156 (2.78)*	22	0.590	3.252
(8) sb09	1.567 (0.699)**	−0.432 (0.54)	−5.816 (3.004)*	22	0.625	2.997
(9) sov'n cds10q1	0.962 (0.605)	−0.001 (0.01)	−3.364 (1.713)**	22	0.550	3.328
(10) bankcrisis	1.083 (0.37)***	1.625 (1.41)	−4.291 (1.199)***	22	0.595	3.620
(11) private debt 07	0.763 (0.456)*	0.027 (0.015)*	−7.345 (1.682)***	21	0.659	2.783
(12) ca07	0.744 (0.76)	0.111 (0.166)	−3.286 (0.986)***	22	0.505	3.541

TABLE 8

Second stage with controls (DFE)

Additional control X	δ	π	Const	n	Adj. R^2	1st F
(1) fe gdp0709	1.236 (0.074)***	−0.023 (0.121)	−4.131 (1.44)***	21	0.663	103.67
(2) forec gdp1011	1.651 (0.201)***	−0.859 (0.369)**	−1.731 (1.251)	21	0.771	115.65
(3) reces dummy	1.289 (0.169)***	0.611 (1.948)	−4.048 (1.099)***	21	0.671	92.73
(4) fe dummy	0.974 (0.202)***	−3.319 (2.122)	−2.574 (1.311)**	21	0.659	49.25
(5) fe sb1011	1.312 (0.189)***	0.381 (0.537)	−4.582 (0.805)***	19	0.635	88.94
(6) trade part cons	1.22 (0.092)***	6.582 (1.605)***	−4.469 (0.548)***	21	0.780	143.76
(7) sov'n debt09	1.297 (0.144)***	0.014 (0.029)	−4.799 (2.018)**	21	0.677	128.32
(8) sb09	1.808 (0.263)***	−0.727 (0.297)**	−8.094 (1.738)***	19	0.755	91.09
(9) sov'n cds10q1	1.263 (0.737)*	0.003 (0.036)	−4.404 (3.956)	20	0.660	97.64
(10) bankcrisis	1.286 (0.099)***	2.709 (1.481)*	−5.249 (1.1)***	21	0.732	148.98
(11) private debt 07	1.12 (0.083)***	0.019 (0.016)	−6.542 (1.552)***	20	0.742	122.66
(12) ca07	1.331 (0.13)***	−0.065 (0.096)	−4.128 (0.707)***	21	0.678	130.57

including this variable does not affect the results qualitatively, even though the persistence parameter increases somewhat. The GDP forecast itself is negative and significant in the second stage. This is plausible, as higher expected GDP growth might have increased the potential output forecast and thus even made the potential output forecast error more negative

Dovern and Jannsen (2017) show that findings of low forecast errors mask substantial differences for periods of recessions and expansions with strongly too optimistic forecasts for periods that turn out to be recessions ex post. Since our data set includes observations with recessions and recoveries in 2010–11, the differential forecast errors might be driven by the *generally poor performance of forecasters in predicting recessions.*[19] We control for this possibility in row (3) by including a recession dummy that equals 1 if a country had a negative realized growth rate in 2010 or 2011. The coefficient θ has the expected negative sign and is large, but remains insignificant. The coefficients of the DFE variable

[19] We are grateful to an anonymous referee for pointing us to this issue.

remain largely unaltered but in the second-stage regression the SB turns insignificant. We also test a somewhat similar specification in row (4). There we use a dummy that equals 1 for an observation with a negative forecast error in 2010–11 as a control variable. Since such a variable should explain the big variation in forecast errors between countries, the fiscal policy variable can only take care of the more gradual differences between countries that are not due to underestimated recessions or recoveries As expected, the coefficient of the dummy is negative, large, and highly significant. β is somewhat reduced and becomes insignificant for the SB case, but remains highly significant in the DFE case. Second-stage results correspond to that.

Could it have been an underestimation of the sheer *size of consolidation* instead of the multiplier effect of consolidation that explains growth disappointments? In row (3), we add the forecast error of the change in the structural balance as an additional control. Again, the effects remain intact. Moreover, there seems to be no relevant underestimation of the consolidation effort during the 2010–11 years. The multiplier effect largely dominates the size effect in terms of forecast errors. What about the *consolidation effort of trading partners*, which could spill over to domestic growth? Adding in row (4) the trade-weighted consolidation effort of trading partners as measured by the change in their structural balance and scaled by the share of exports in GDP does not affect our coefficients of interest, even if the parameter itself becomes highly significant and large.

Another perturbing candidate could be ignoring the impact of the *soundness of domestic public finances*. Maybe forecasts were too optimistic because public finances were in bad shape and their influence on growth might have been underestimated. We test this possibility in rows (5)–(7) where we use as a proxy either the initial sovereign debt-to-GDP level of 2009, the initial structural balance of 2009 or the spread of sovereign credit default swaps as an average during the first quarter of 2010, respectively. The parameters belonging to the DFE measure are qualitatively unaffected. When controlling for the initial structural balance in 2009, the persistence of multipliers is even reinforced. The coefficient of the initial structural balance itself becomes significantly negative in the second stage of the DFE estimation, meaning that for countries with higher structural deficits on the outset, potential growth forecasts were comparably too pessimistic. The stabilizing role of expansionary fiscal policy seems to have been underestimated. In the case of sovereign CDS spreads, first-stage results do not change much, but the significance levels of the persistence parameter become lower in the DFE case and even insignificant when using the SB. Thus, there may be some counterbalancing effect of high sovereign risk as argued by Corsetti *et al.* (2013), but it does not seem to be very strong, at least when using our preferred measure of fiscal shocks.

What about the private sector and its likely underestimated impact on growth through bank stress or private debt overhang? Controlling for the indicator of Laeven and Valencia (2012), which signals whether a country is in a *banking crisis* in a certain year, does not affect our parameters of interest. Using *precrisis household debt-to-GDP ratios* of 2007 as a proxy for the pressure to deleverage does not affect the first-stage regressions, but lowers the significance level of the persistence parameter in the SB case. The DFE case again is much more robust. Finally, when controlling for the *precrisis current-account-to-GDP ratio* as a measure of external imbalances that might have stalled output growth more than expected, we again find our DFE estimation largely unaffected. For the SB case, the persistence coefficient turns smaller and insignificant.

Summing up, controlling for various alternative explanations does not affect our central findings at least when we rely on the narrative DFE measure, where also the F statistics still signal strong instruments. For the coefficients of the SB measure results remain robust in most instances but the instruments become even weaker.

Using European Commission data

Is the IMF forecast a special case? We test the European Commission's forecast as well, using the spring 2010 European Economic Forecast as well as $t+4$ forecasts of potential output. The EC data include the whole EU27 and thus some additional Eastern European countries, that are absent from the IMF data set. Repeating the previous regressions with EC data, most of the results are confirmed. Results are presented in an online Appendix B. Estimates using the structural balance are even more robust to the perturbations we tested for the IMF data. Concerning the DFE there are two interesting and plausible outliers: for the whole EU27, the coefficients of interest are somewhat weaker ($\beta = [0.5; 0.7]$, $\delta = [0.9; 1.1]$). Most notably, the relation completely diminishes when excluding the programme countries (Table B1(b), Column(3)), and the separate effects of spending and revenue shocks is turned upside down (Table B1(b), Column(2)). These findings are fully driven by the data of Latvia and Lithuania, countries that are absent in the IMF dataset and that witnessed a tremendous crash in 2009 with a cumulated GDP growth forecast error for the years 2008–09 as of the 2008 spring forecast of more than -20 pp each. It is not implausible that (potential) growth forecasts where more on the pessimistic side in the following years. Moreover, both countries are very small, very open economies that joined the EU only in 2004, which gave them a strong push to export growth. In such circumstances fiscal devaluation is considered less harmful (Perotti, 2011). When we exclude these special cases, the previous results of the IMF sample are reestablished in full (Series (c) of Tables B1-4, online Appendix B).

Extending the time dimension

In our baseline, we derive forecast errors from the vintage of spring 2016 *vis-à-vis* 2010 and are therefore restricted to only 21/22 observations in the IMF case and 27 with EC data. Fiscal consolidation in many European countries has, however, continued after 2011. Also, it might be interesting to check the short and long run impact for late crisis years. Therefore, we test for forecasts in subsequent years and extend the time dimension of the estimation in two ways. First, we assess different forecast vintages individually in form of a moving window and second, jointly in a panel structure. As we only have limited access to IMF vintages with $t+5$ forecasts we concentrate in the main body of this paper on results with EC data for the moving window and panel model exercise. Online Appendix B presents limited samples with IMF data. Generally, the results for the first and second stage are robust to the exercise of extending the time dimension when using the DFE as fiscal shocks, while using SB produces rather inconsistent results. The model for the moving window is equivalent to section 2. The two-year fiscal shocks and growth forecast errors move along with the respective vintage year. Tables 9–12 show moving window regression results for vintages between 2010 and 2014 using SB and DFE for the first and second stage, respectively.

TABLE 9

Moving window first stage (SB)

Endogenous: forecast error of GDP $t:t+1$					
OLS	2010	2011	2012	2013	2014
β	−1.166**	0.065	−0.473	0.564	0.204
	(0.461)	(0.268)	(0.312)	(0.389)	(0.578)
Const	0.633	−2.916***	−0.876	0.584	0.649
	(0.500)	(0.847)	(0.554)	(0.571)	(0.526)
N	27	27	27	27	27
Adj. R^2	0.431	0.001	0.094	0.065	0.009

TABLE 10

Moving window first stage (DFE)

Endogenous: forecast error of GDP $t:t+1$					
OLS	2010	2011	2012	2013	2014
β	−0.586**	−0.609**	−0.530***	0.248	0.343
	(0.245)	(0.252)	(0.116)	(0.172)	(0.286)
Const	2.216***	−1.057	−0.029	0.474	0.417
	(0.552)	(0.851)	(0.460)	(0.570)	(0.516)
N	27	27	27	27	27
Adj. R^2	0.376	0.308	0.440	0.056	0.030

In the first stage, SB case baseline results are not confirmed by other vintage years, β becomes economically and statistically insignificant. However, using DFE provides robust results for the main period of European consolidation, vintage years 2010–12 with a multiplier underestimation between −0.5 and −0.6. Afterwards the effect vanishes which may be due to the slowdown of consolidation in general, the fact that forecasters learned from their mistakes or be interpreted in line with findings of regime-dependent multiplier effects (Baum, Schaffer and Stillman, 2010; Auerbach and Gorodnichenko, 2012) Turning to the second stage provides a similar picture. The baseline persistence is qualitatively confirmed for DFE while SB only yields mixed results. For the years 2010–12, persistence estimated using DFE is on a somewhat higher level compared to baseline, δ increases from 0.965 for 2010 to 1.394 for 2012, afterwards fiscal shocks show no significant persistence effect. Hence, we observe a weakening of the effects in late crisis years. Contrary to baseline estimates, the results for later vintages do not elementarily differ when excluding Latvia and Lithuania.

In a next step, we increase the number of observations by applying a panel structure with different sets of vintages, following BL in the case of short-term multipliers. For the sake of brevity, our panel regression results are discussed in detail in an online Appendix C. The panel results generally confirm the baseline estimates. Again, DFE proofs to be quite robust for alternative time dimensions, while SB remains ambiguous The coefficient β stays within the range of 0.4–0.6. Including late crisis years somewhat lowers β but the structural underestimation does not vanish. Coefficient δ on the contrary increases with time, from 1.0 (for years 2010/11) to 1.2 (2010/14). Note that further specifications with

TABLE 11

Moving window second stage (SB)

Endogenous: forecast error of potential GDP $t:t+4$					
TSLS	2010	2011	2012	2013	2014
δ	1.271***	−0.087	2.877***	1.379	1.445
	(0.222)	(5.688)	(1.095)	(1.279)	(2.223)
Const	−3.909***	−3.359	3.921**	0.601	0.401
	(0.690)	(15.840)	(1.836)	(0.978)	(1.518)
N	27	27	27	27	27
Adj. R^2	0.543	−0.126	0.186	0.658	0.729
1st stage F	6.408	0.0594	2.290	2.101	0.124

TABLE 12

Moving window second stage (DFE)

Endogenous: forecast error of potential GDP $t:t+4$					
TSLS	2010	2011	2012	2013	2014
δ	0.965***	0.983***	1.394***	0.893	−0.242
	(0.257)	(0.190)	(0.298)	(1.245)	(1.152)
Const	−3.675***	−0.312	1.731*	1.031	1.449
	(0.690)	(0.913)	(0.965)	(1.159)	(1.342)
N	27	27	27	27	27
Adj. R^2	0.514	0.522	0.486	0.504	−0.305
1st stage F	5.708	5.826	20.75	2.077	1.439

different panel dimensions[20] after the crisis for both the SB and DFE case do not alter the general picture drawn so far – quite robust estimates with general weakening of the baseline effects in later crisis years, in line with the slowdown of consolidation potential learning effects of forecasters (Górnicka *et al.*, 2018) or the end of the downturn regime

Exploring different time dimensions has shown that the severe underestimation of consolidation effects on output in the short-run and the subsequent persistence is restricted to the period of the European debt crisis where we observed significant contractions. Accordingly, the results could speak to the empirical literature on regime-dependent multipliers showing significantly increases of effect sizes in recessionary periods of the business cycle and on the lower end of findings under the ZLB in standard macroeconomic models (Christiano *et al.*, 2011; Eggertsson, 2011).

V. Concluding remarks

By exploiting forecast errors of output and long-term potential output growth in the spirit of Blanchard and Leigh (2013) and Fatás and Summers (2018), but using a superior, narratively identified measure of the fiscal stance, we have investigated as to whether the size and persistence of fiscal multipliers was underestimated for the austerity measures that were implemented in Europe after 2009. In line with these earlier papers, we find that multipliers were strongly underestimated by about 0.7–1.0 units. This would translate

[20] The panel sample might start later or be shortened, e.g. 11/12–11/14, 12/13 etc.

into a multiplier range of 1.2–1.5, given that forecasters of the IMF and the European Commission on average assumed a multiplier of 0.5, a claim, for which we presented some evidence. Most interestingly, fiscal policy seems to have had a permanent effect in the 2010–11 period and beyond. These results hold up to a battery of perturbations and particularly so when relying on our improved identification strategy. Interestingly, it turns out that the effects weaken for measures in late crisis years after 2013 and when including very small very open economies.

For our European sample, we find evidence for strong hysteresis effects as opposed to the short-run pain, long-term gain consensus that emerged after the early crisis years. That is, the turn to belt tightening was badly timed and therefore much more costly in terms of long-term output loss than a more gradual, backloaded consolidation.

Final Manuscript Received: November 2018

References

Alesina, A. and Ardagna, S. (2010). 'Large changes in fiscal policy: taxes versus spending', *NBER/Tax Policy & the Economy*, Vol. 24, pp. 35–68.

Auerbach, A. J. and Gorodnichenko, Y. (2012). 'Measuring the output responses to fiscal policy', *American Economic Journal: Economic Policy*, Vol. 4, pp. 1–27.

Baum, A., Poplawski-Ribeiro, M. and Weber, A. (2012). *Fiscal Multipliers and the State of the Economy*, IMF Working Paper No. WP/12/286.

Baum, C. F., Schaffer, M. E. and Stillman, S. (2010). *ivreg2: Stata Module for Extended Instrumental Variables/2SLS, GMM and AC/HAC, LIML and k-class Regression*, http://ideas.repec.org/c/boc/bocode/s425401.html

Blanchard, O. J. and Leigh, D. (2013). *Growth Forecast Errors and Fiscal Multipliers*, NBER Working Paper No. 18779.

Bom, P. R. and Ligthart, J. E. (2014). 'What have we learned from three decades of research on the productivity of public capital?', *Journal of Economic Surveys*, 28, pp. 889–916.

Boneva, L. M., Braun, R. A. and Waki, Y. (2016). 'Some unpleasant properties of loglinearized solutions when the nominal rate is zero', *Journal of Monetary Economics*, Vol. 84, pp. 216–232.

Born, B., Müller, G. J. and Pfeiffer, J. (2015). *Does Austerity Pay Off?* CEPR Discussion Papers No. 10425.

Caldara, D. and Kamps, C. (2017). 'The analytics of SVARs: a unified framework to measure fiscal multipliers', *Review of Economic Studies*, Vol. 84, pp. 1015–1040.

Carnot, N. and de Castro, F. (2015). *The Discretionary Fiscal Effort: An Assessment of Fiscal Policy and Its Output Effect*, European Economy Economic Papers No. 543.

Christiano, L., Eichenbaum, M. and Rebelo, S. (2011). 'When is the government spending multiplier large?', *Journal of Political Economy*, Vol. 119, pp. 78–121.

Cloyne, J. (2013). 'Discretionary tax changes and the macroeconomy: new narrative evidence from the United Kingdom', *American Economic Review*, Vol. 103, pp. 1507–1528.

Corsetti, G., Kuester, K., Meier, A. and Müller, G. J. (2013). 'Sovereign risk, fiscal policy, and macroeconomic stability', *The Economic Journal*, Vol. 123, pp. F99–F132.

Cottarelli, C. and Jaramillo, L. (2012). *Walking Hand in Hand: Fiscal Policy and Growth in Advanced Economies*, IMF Working Paper No. WP/12/137.

DeLong, J. B. and Summers, L. H. (2012). 'Fiscal policy in a depressed economy', *Brookings Papers on Economic Activity*, Vol. 2012, pp. 233–274.

Dovern, J. and Jannsen, N. (2017). 'Systematic errors in growth expectations over the business cycle', *International Journal of Forecasting*, Vol. 33, pp. 760–769.

Eggertsson, G. B. (2011). 'What fiscal policy is effective at zero interest rates?', *NBER Macroeconomics Annual*, Vol. 25, pp. 59–112.

European Commission (2012). *European Economic Forecast. Autumn 2012*, European Economy 7/2012.
European Commission (2013). *Report on Public finances in EMU*, European Economy 4/2013.
Fatás, A. (2000). 'Do business cycles cast long shadows? Short-run persistence and economic growth', *Journal of Economic Growth*, Vol. 5, pp. 147–162.
Fatás, A. and Summers, L. H. (2018). 'The permanent effects of fiscal consolidations', *Journal of International Economics*, Vol. 112, pp. 238–250.
Firstrun Project (2016). *A Dataset of Fiscal Variables*, A vintage of the Ameco database of the European Commission.
Frankel, J. (2011). 'Over-optimism in forecasts by official budget agencies and its implications', *Oxford Review of Economic Policy*, Vol. 27, pp. 536–562.
Furman, J. (2016). *The New View of Fiscal Policy and Its Application*, Delivery for Conference: Global Implications of Europe's Redesign. October 5, 2016. New York.
Gechert, S. (2015). 'What fiscal policy is most effective? A meta-regression analysis', *Oxford Economic Papers*, Vol. 67, pp. 553–580.
Gechert, S. and Rannenberg, A. (2018). 'Which fiscal multipliers are regime-dependent? A meta-regression analysis', *Journal of Economic Surveys*, Vol. 32, pp. 1160–1182.
Gechert, S., Rietzler, K. and Tober, S. (2015). 'The European Commission's new NAIRU: does it deliver?', *Applied Economics Letters*, Vol. 23, pp. 6–10.
Górnicka, L., Kamps, C., Koester, G. and Leiner-Killinger, N. (2018). *Learning about Fiscal Multipliers During the European Sovereign Debt Crisis: Evidence from a Quasi-Natural Experiment*. ECB Working Paper No. 2154.
Havik, K., Mc Morrow, K., Orlandi, F., Planas, C., Raciborski, R., Roeger, W., Rossi, A., Thum-Thysen, A. and Vandermeulen, V. (2014). *The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps*, European Economy Economic Papers No. 535.
Hebous, S. (2011). 'The effects of discretionary fiscal policy on macroeconomic aggregates: a reappraisal', *Journal of Economic Surveys*, Vol. 25, pp. 674–707.
International Monetary Fund (2010). *World Economic Outlook. Recovery, risk, and rebalancing*. Washington, DC: WEO.
Jonung, L. and Larch, M. (2006). 'Improving fiscal policy in the EU: the case for independent forecasts', *Economic Policy*, Vol. 21, pp. 491–534.
Laeven, L. and Valencia, F. (2012). *Systemic Banking Crises Database: An Update*, IMF Working Paper No. 12/163.
Logeay, C. and Tober, S. (2006). 'Hysteresis and the NAIRU in the Euro area', *Scottish Journal of Political Economy*, 53, pp. 409–429.
Mertens, K. and Ravn, M. O. (2014). 'A reconciliation of SVAR and narrative estimates of tax multipliers', *Journal of Monetary Economics*, Vol. 68(S), pp. S1–S19.
Mineshima, A., Poplawski-Ribeiro, M., and Weber, A. (2014). 'Size of fiscal multipliers', in Cottarelli C., Gerson P. and Senhadji (Hrsg.) A. (eds), *Post-Crisis Fiscal Policy*, Cambridge, MA: MIT Press, pp. 315–372.
Perotti, R. (2011). *The 'Austerity Myth': Gain Without Pain?* NBER Working Paper No. 17571.
Price, R., Dang, T. and Guillemette, Y. (2014). *New Tax and Expenditure Elasticity Estimates for EU Budget Surveillance*, OECD Economics Department Working Papers No. 1174.
Rogoff, K. S. (2012). *Austerity and Debt Realism*, Project Syndicate, June 1, 2012.
Romer, C. D. and Romer, D. H. (2010). 'The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks', *American Economic Review*, Vol. 100, pp. 763–801.
Schäuble, W. (2010). A plan to tackle Europe's debt mountain. *Europe's World*. October 1, 2010. http://europesworld.org/2010/10/01/a-plan-to-tackle-europes-debt-mountain/
Sturn, S. (2014). 'Macroeconomic policy in recessions and unemployment hysteresis', *Applied Economics Letters*, Vol. 21, pp. 914–917.
Trichet, J.-C. (2010). Stimulate no more – it is now time for all to tighten. *Financial Times*. July 22, 2010. http://www.ft.com/cms/s/0/1b3ae97e-95c6-11df-b5ad-00144feab49a.html
Woodford, M. (2011). 'Simple analytics of the government expenditure multiplier', *American Economic Journal: Macroeconomics*, Vol. 3, pp. 1–35.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix A. Computing forecast errors.
Appendix B. Different data and time dimension.
Appendix C. Panel estimations.