The Effect of the European Excessive Deficit Procedure on Fiscal Spending Multipliers*

David A. Vespermann
Heidelberg University

Christopher Zuber
Heidelberg University

November 15, 2021

Abstract

The Excessive Deficit Procedure (EDP) is the corrective arm of European fiscal governance. Using a panel of European countries and state-dependent local projections, we document that cumulative fiscal spending multipliers are larger for countries in the EDP. This result is driven by lower interest rates and substantial crowding-in of private investment in response to a positive government spending shock. Multipliers in the EDP are even larger in times of weak fiscal position or recessionary episodes, indicating that the procedure is particularly effective. We show that the EDP is not simply a proxy for these times. In addition, we find that policy makers underestimate fiscal multipliers in real time. The results suggest that the EDP is functional and increases the effectiveness of fiscal stimulus.

Keywords: Fiscal policy, fiscal multipliers, government spending shocks, local projections, European Excessive Deficit Procedure

JEL-Codes: E32, E62, H60

*We would like to thank Christian Conrad, Jonas Dovern, Zeno Enders, Gernot Müller and the participants of the Macro-Econometrics seminar at Heidelberg University for helpful comments. Please address correspondence to christopher.zuber@awi.uni-heidelberg.de.

1 Introduction

The Stability and Growth Pact is one of the core elements of European fiscal governance. Implemented in 1997, it provides a rule-based framework for fiscal policy to ensure sound public finances and fiscal discipline in the European Union (EU). The Excessive Deficit Procedure (EDP) is the corrective arm of this framework. If a member state runs an excessive deficit or accumulates high levels of public debt, an EDP is launched. The adjustments prescribed by the EDP provisions aim to ensure appropriate fiscal policies and, as a consequence, sustainable public finances. These measures put an additional strain on public budgets and fiscal spending behavior. This may interfere with the impact of government spending on economic activity and, in particular, the effectiveness of fiscal stimulus in times of weak fiscal position or recessionary episodes. In this paper, we explore whether the EDP affects fiscal spending multipliers. Specifically, we estimate state-dependent impulse response functions and cumulative multipliers for a panel of European countries using local projections to identify how the EDP alters the transmission of government spending shocks.

In the literature on the determinants of fiscal multipliers, the recent debate has revolved around potential state dependence. In particular, a growing number of studies explores how economic conditions affect the output response to fiscal stimulus. Regarding the impact of the business cycle position, the empirical evidence is mixed. Auerbach and Gorodnichenko (2012, 2013) document that fiscal multipliers in recessions are larger than in expansions. By contrast, Owyang et al. (2013) and Ramey and Zubairy (2018) do not find evidence that public spending is more effective during times of slack measured by high levels of the unemployment rate.

Ilzetzki et al. (2013) evaluate how fiscal multipliers vary across different economic environments and find larger multipliers under fixed exchange rate regimes and negative multipliers in high-debt countries. The latter finding is confirmed by Nickel and Tudyka (2014), whereas, more recently, Banerjee and Zampolli (2019) report small but positive fiscal multipliers below one. Corsetti et al. (2012b) document negative government spending multipliers for countries with weak fiscal positions as measured by high debt levels and excessive deficits. This finding is driven by significant crowding-out of private investment. Huidrom et al. (2020) identify two channels through which high public debt has an effect on the size of fiscal multipliers. First, Ricardian households decrease consumption in response to fiscal stimulus because they anticipate adjustments in the future. Second, public spending raises sovereign credit risk and puts upward pressure on interest rates, which reduces private demand. Thus, lower multipliers in weak fiscal positions are rationalized by crowding-out of private consumption.

Surpisingly, the literature analyzing the effectiveness of the EDP is very limited. Hagen and Eichengreen (1996) suspected that the EDP is at best redundant but more likely detrimental as it limits policy makers' scope for reactions. De Jong and Gilbert (2020) show that countries in the EDP largely comply in terms of fiscal reactions with the recommendations of the European Commission. Górnicka et al. (2020) calculate implied fiscal multipliers in the EDP from these recommendations and document implied multipliers below one, which increase over time. So far, however, none has systematically estimated fiscal multipliers for countries in the EDP over several horizons, compared them to multipliers for countries not in the EDP or investigated the transmission mechanisms of the EDP on the economy. We aim to fill this gap.

We contribute to the literature by evaluating the specific impact of the EDP on fiscal spending multipliers. More precisely, we screen official documents provided by the European Commission to construct a dummy variable indicating whether a country is in the EDP or not. We use this dummy as state variable in the estimation of cumulative spending multipliers and the corresponding impulse response functions. We document that cumulative multipliers for countries in the EDP are significantly larger, suggesting that government spending in these countries is more effective. The inspection of the underlying transmission mechanisms reveals that this finding is mainly driven by a decrease in interest rates and substantial crowding-in of private investment in response to fiscal stimulus. At the same time, we observe a decrease in debt. We find that the EDP is especially effective in bad times, as indicated by larger multipliers for countries with weak fiscal positions, in recessions or during banking crises. In addition, the comparison with alternative state variables provides evidence that the EDP is not simply a proxy for such bad times. Finally, we show that fiscal multipliers are significantly understated if estimated in real time.

These results have some important policy implications. First, and most importantly, our findings are evidence that the EDP is functional. The output response to government spending is stronger and public debt decreases in response to a positive spending shock for countries in the procedure. Thus, fiscal stimulus could help these countries to return on a path towards more sustainable public finances. By contrast, fiscal consolidation would have substantial contractionary and therefore harmful effects in these economies. Second, the EDP is especially effective in bad times which is the purpose it is designed for. Third, lower-than-expected fiscal multipliers in real time mask the effect of fiscal stimulus for countries in the EDP. This is especially important for policy makers who should treat real-time multipliers cautiously.

The remainder of the paper is structured as follows. In Section 2, we describe the EDP and explain the construction of the EDP dummy variable. Section 3 introduces the empirical strategy and the data. We discuss the estimation results in Section 4 and compare these results with fiscal multipliers based on alternative state variables in Section 5. In Section 6, we present results for real-time fiscal multipliers. Section 7 concludes.

2 The Excessive Deficit Procedure

The EDP is a multi-step procedure with the objective to correct imbalances of public finances by urging the affected member states to reduce excessive deficits and/or debt levels. Despite the name of the procedure, the EDP always takes into account both criteria. The details of the EDP provisions are laid down in Article 126 of the *Treaty on the Functioning of the European Union* (European Union, 2008). An EDP can be launched in two cases. First, for countries with headline budget deficits exceeding, or being at risk to exceed, the limit of 3% of GDP. Second, for countries with public debt levels above 60% of GDP which do not decrease at a sufficient pace, defined as 1/20th of the gap between the actual level and the 60% reference value per year. The initial step is a report by the European Commission notifying the non-compliance of a member state with the requirements for deficit and/or debt. In the next step, the Commission proposes whether or not an excessive deficit should be declared, taking into account "all other relevant factors, including the medium-term economic and budgetary position" (European Union, 2008) of the country. Based on this proposal, the European Council officially decides on the existence of an excessive deficit. If the Council declares that an excessive deficit exists, this decision formally opens the EDP and triggers a series of actions. Commission and Council issue recommendations on how to correct the paths of aggregate deficit and debt and set a deadline. The measures undertaken by the member states are continuously monitored and, in case of non-effective action, the Council can impose sanctions. The first stage is a compulsory non-interest-bearing deposit of 0.2% of GDP, which can be converted into a fine in the second stage. In practice, sanctions have never been imposed so far. The EDP is closed when the Council, on a recommendation by the Commission, decides that the excessive deficit has been corrected.

For our analysis, we use the official documents on all ongoing and closed EDPs provided by the European Commission. Out of the 27 EU countries and the former member United Kingdom, 25 countries have been at least once in the procedure between 2000 and 2019. Estonia, Luxembourg and Sweden are the only countries which have never been in an EDP. Table 1 lists the EDPs that are included in our sample.[1] Twelve of these EDPs were opened during the Great Recession or in its aftermath. At the end of 2019, there was no ongoing EDP. The last procedure was closed in summer 2019 after ten years of monitoring Spain. This is also the longest EDP in our sample where the average EDP duration is five years. The table shows that Commission and Council act rather swiftly with an average duration of two months between

[1]We cannot include all EDPs which have been opened by the Council in the period 2000–2019 because of missing data for some European countries. The main obstacle is insufficient data for the construction of government spending shocks which we detail in Section 3.3.

Table 1: EDPs included in the sample

#	Country	1st neg. report	Comm. rec.	Start	End
1	Belgium	–	11 Nov 2009	2 Dec 2009	20 Jun 2014
2	Czech Republic	–	24 Jun 2004	5 Jul 2004	3 Jun 2008
3	Czech Republic	7 Oct 2009	11 Nov 2009	2 Dec 2009	20 Jun 2014
4	Denmark	12 May 2010	15 Jun 2010	13 Jul 2010	20 Jun 2014
5	Finland	12 May 2010	15 Jun 2010	13 Jul 2010	12 Jul 2011
6	France	2 Apr 2003	7 May 2003	3 Jun 2003	30 Jan 2007
7	France	18 Feb 2009	24 Mar 2009	27 Apr 2009	22 Jun 2018
8	Germany	19 Nov 2002	8 Jan 2003	21 Jan 2003	5 Jun 2007
9	Germany	7 Oct 2009	11 Nov 2009	2 Dec 2009	22 Jun 2012
10	Hungary	12 May 2004	24 Jun 2004	5 Jul 2004	21 Jun 2013
11	Ireland	18 Feb 2009	24 Mar 2009	27 Apr 2009	17 Jun 2016
12	Italy	7 Oct 2009	11 Nov 2009	2 Dec 2009	21 Jun 2013
13	Netherlands	28 Apr 2004	19 May 2004	2 Jun 2004	7 Jun 2005
14	Netherlands	7 Oct 2009	11 Nov 2009	2 Dec 2009	20 Jun 2014
15	Portugal	22 Jun 2005	20 Jul 2005	20 Sep 2005	3 Jun 2008
16	Portugal	7 Oct 2009	11 Nov 2009	2 Dec 2009	16 Jun 2017
17	Slovak Republic	7 Oct 2009	11 Nov 2009	2 Dec 2009	20 Jun 2014
18	Spain	18 Feb 2009	24 Mar 2009	27 Apr 2009	14 Jun 2019
19	United Kingdom	21 Sep 2005	11 Jan 2006	24 Jan 2006	9 Oct 2007
20	United Kingdom	11 Jun 2008	2 Jul 2008	8 Jul 2008	5 Dec 2017

Notes: "1st neg. report" refers to the date at which the European Commission filed the first negative report on the non-compliance of the country. "Comm. rec." refers to the date the Commission recommends to the European Council that the EDP should be declared. "Start" and "End" denote the date of the opening and closing of the EDP as decided by the Council. For EDPs no. 1 and 2, the Commission proposed the EDP without a first negative report.

the first negative report of the Commission, the following Commission's recommendation on the opening of the EDP and the decision of the Council on that matter. After the Commission recommends an opening, the Council opens the EDP usually within one month. There exists no recommendation which was not followed by an opening of an EDP.

We code a dummy variable that captures the periods in which a country was in an EDP. Since we use semi-annual data from OECD Economic Outlook (EO) editions, we have to make sure that changes in the dummy variable correspond to the EO editions. More specifically, the EDP dummy variable is set to 1 from period t onwards if the procedure was opened between the forecast cut-off date of EO edition $t-1$ and the forecast cut-off date of edition t. Accordingly, the dummy variable is set to 0 from period $t+1$ onwards if the procedure was closed between the forecast cut-off date of EO edition t and the forecast cut-off date of edition $t+1$. By doing so, we carefully account for the information set which is available at the time each EO edition is published. This approach provides us with country-specific EDP dummies which we use as state variables in the panel estimation of state-dependent fiscal multipliers and impulse response functions presented in the next section.

3 Methodology and Data

3.1 Empirical Strategy

We use the local projection method proposed by Jordà (2005) to estimate impulse response functions and fiscal multipliers directly. This method allows for a flexible specification of state dependence and does not implicitly restrict the model dynamics. In particular, we follow Auerbach and Gorodnichenko (2013) and estimate state-dependent local projections using panel data for European countries. The state-dependent responses $\beta_{0,h}$ and $\beta_{1,h}$ of the variable of interest $Z_{i,t+h}$ to an exogenous change in government spending for each horizon $h = 0, ..., H$ are estimated from the following regression model:

$$\begin{aligned} Z_{i,t+h} &= (1 - \mathcal{I}_{i,t-1}) \left[\alpha_{0,i,h} + \beta_{0,h} G_{i,t} + \Phi_{0,h}(L) X_{i,t-1}\right] \\ &\quad + \mathcal{I}_{i,t-1} \left[\alpha_{1,i,h} + \beta_{1,h} G_{i,t} + \Phi_{1,h}(L) X_{i,t-1}\right] + \sum_{k=1}^{2} \psi_k T_t^k + \varepsilon_{i,t+h}, \end{aligned} \tag{1}$$

where $\mathcal{I}_{i,t-1}$ indicates the state of the economy of country i in the period before the change in government spending, $\alpha_{\bullet,i,h}$ measures unobserved state-dependent fixed effects, $G_{i,t}$ represents government spending in the current period, L refers to the lag operator, $X_{i,t-1}$ is a vector of controls and the series of ψ_k captures a time trend. In our baseline specification, the state indicates whether a country is in an ongoing EDP. The set of control variables includes output, government spending, private consumption, private investment, the interest-rate spread, the marginal tax rate and the public debt level. The variable of interest $Z_{i,t+h}$ is a variable from this set. The units of all variables measured in levels (i.e., output, goverment spending, private consumption, private investment and the public debt level) are normalized by an estimate of trend GDP (Gordon and Krenn, 2010; Ramey and Zubairy, 2018) rather than lagged GDP (as done by, e.g., Hall, 2009; Barro and Redlick, 2011). The latter approach produces fiscal multipliers which vary over the business cycle. We obtain trend GDP from a polynomial of order 3.[2] In order to address potential endogeneity issues in our regressions, we use an instrumental variable approach: Normalized government spending $G_{i,t}$ is instrumented by the forecast error of government spending (Auerbach and Gorodnichenko, 2012, 2013). We discuss the details on the identification of government spending shocks in Section 3.3. Finally, we use robust standard errors that account for cross-sectional dependence and autocorrelation as proposed by Driscoll and Kraay (1998).

The impulse response function for each of the states is constructed using the sequence

[2] Ramey and Zubairy (2018) use a polynomial of order 6. We justify our choice by the smaller sample period which makes less turning points necessary.

of responses $\{\beta_{\bullet,h}\}_{h=0}^{H}$. The impulse responses trace the impact of the exogenous shock on the path of specific variables and reveal the underlying transmission mechanisms. Following Mountford and Uhlig (2009) and Ramey and Zubairy (2018), we argue that the policy-relevant measure for the aggregate effect of government spending shocks on the economy is given by the cumulative fiscal spending multiplier. The cumulative multiplier compares the cumulative output response (i.e, the integral of the output response) to the cumulative path (i.e., the integral) of government spending, thereby providing a measure for the impact of fiscal stimulus over time. Note that this definition is different from the one in Blanchard and Perotti (2002) and Auerbach and Gorodnichenko (2012, 2013) who report fiscal multipliers given by the peak response of output relative to the initial fiscal spending impulse. Peak multipliers, however, do not take into account the underlying response of government spending and therefore complicate the comparison across estimations.

Specifically, we follow Ramey and Zubairy (2018) and Bernardini and Peersman (2018) and estimate state-dependent cumulative fiscal multipliers for horizon h in one step using the following regression:[3]

$$\begin{aligned}\sum_{j=0}^{h} Y_{i,t+j} \;=\; & (1-\mathcal{I}_{i,t-1})\left[\alpha_{0,i,h} + M_{0,h}\sum_{j=0}^{h} G_{i,t+j} + \Phi_{0,h}(L)X_{i,t-1}\right] \\ & + \mathcal{I}_{i,t-1}\left[\alpha_{1,i,h} + M_{1,h}\sum_{j=0}^{h} G_{i,t+j} + \Phi_{1,h}(L)X_{i,t-1}\right] + \sum_{k=1}^{2}\psi_k T_t^k + \varepsilon_{i,t+h},\end{aligned} \tag{2}$$

where $Y_{i,t+j}$ denotes normalized output and all other variables are defined as explained above. We instrument $\sum_{j=0}^{h} G_{i,t+j}$ by the forecast error of government spending. Hence, the instrument is independent of the horizon h, see Ramey and Zubairy (2018). In this specification, the estimated coefficients $M_{0,h}$ and $M_{1,h}$ provide direct measures for the cumulative fiscal multipliers for each state at horizon h.

3.2 Data

The sample covers the period 2000H1–2019H1 and 17 European countries.[4] Our main data source is the OECD EO published on a biannual basis (spring and autumn of each year). We use this dataset for two reasons. First, because of its large coverage. It includes macroeconomic variables along with forecasts of up to two years ahead for most European countries. Hence,

[3] Ramey and Zubairy (2018) argue that the one-step estimation procedure is identical to a three-step procedure in which the sum of the output responses is divided by the sum of government spending responses if all responses are estimated on the same sample.

[4] Our sample includes Austria, Belgium, the Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Luxembourg, the Netherlands, Portugal, the Slovak Republic, Spain, Sweden and the United Kingdom.

this provides us with a consistent dataset. Second, this source additionally provides us with government spending forecasts from each EO edition which we use to identify government spending shocks, see Section 3.3 for details. Since the EO reports quarterly values only starting from edition 73 (published in spring 2003) and semiannual values before, we harmonize the frequency of all variables to consistently use semiannual data in the analysis.

We mainly use data from EO edition 106 published in autumn 2019. We take real data for output, government spending (government consumption plus—if available—government investment), private consumption and private investment.[5] We construct semiannual levels for these variables by aggregating the values from Q1 and Q2 (Q3 and Q4) for H1 (H2).[6] For the interest-rate spread, we take the 10-year government bond spread vis-à-vis Germany (in percent). The value for the semester is given by the average over its two quarters. In addition, we take the public debt level which is measured at the end of the year and reported at annual frequency only. Therefore, we use the annual value of year t for H2 in t and H1 in $t+1$.

Finally, we take the annual marginal personal income tax rate (in percent) from Table I.4 of the OECD Tax Database at an income level of 100% of the average wage. This rate includes the central government and sub-central income tax plus the employee social security contributions. We use the value of year t for H1 and H2 in t.

3.3 Shock Identification

We define government spending shocks as the forecast error of government spending growth. By doing so, goverment spending shocks aim to measure the unexpected change in government spending growth and can be used as an instrument for government spending. This approach was put forward by Auerbach and Gorodnichenko (2012, 2013) and has been widely used since then. The identification is based on the timing assumption of Blanchard and Perotti (2002). That is, unexpected changes in government spending growth are exogenous and therefore not a contemporaneous response to macroeconomic aggregates.

The approach can be divided into two steps. First, we calculate government spending growth from real government consumption plus—if available—real government investment for all countries and all EO editions between edition 67 (published in spring 2000) and edition

[5]Government and private investment are only reported for seven countries separately. For the other ten countries in our sample the sum of the two is reported. We therefore include government investment where it is available for these seven countries and proxy private investment by the sum of the annual shares of household and corporate investment multiplied by semiannual total investment. The latter approach is valid. The correlation between implied private investment and reported private investment is very high (above 0.85 for each of the seven countries).

[6]Before aggregation, we replace government investment for the United Kingdom in 2005Q2 by the average of the previous quarter and the following quarter because government investment was exceptionally negative due to the transfer of nuclear reactors to the government.

106 (published in autumn 2019). The inclusion of government investment is important because government investment multipliers are much smaller as shown by Boehm (2020). Solely using government consumption could consequently exaggerate the fiscal multipliers. We refer to the calculated growth rates by $g_{i,t}^{s}$. This is the semiannual government spending growth rate for country i between semesters $t-1$ and t based on data from the EO edition published in semester s. Second, we calculate the forecast error of government spending growth by:

$$FE_{i,t} = g_{i,t}^{2019H2} - g_{i,t}^{t-1}, \tag{3}$$

where $g_{i,t}^{2019H2}$ is the realized growth rate from EO edition 106 and $g_{i,t}^{t-1}$ is the one-step-ahead forecast for semester t published in semester $t-1$. We make sure that the growth rates are comparable in terms of the inclusion of government investment, i.e., either both growth rates contain government investment or both do not.

There are several alternative choices for the realized growth rate as the first release of the growth rate is obviously published in $t+1$ and one can consider realizations from $t+1$ to the latest available semester.[7] The OECD continuously revises government consumption and government investment also back to the past and we do not know which realization forecasters aimed to predict. Even if one arbitrarily chose the realizations after a specific fixed horizon h (i.e., the realization of t published in $t+h$), it would be uncertain whether forecasters consistently aimed to predict the realization published in $t+h$ for all t. We do not want to make a stand on which growth rate should be considered as the final release, i.e., the aim of the forecaster's prediction. By using the growth rate from the latest available semester, we are agnostic about which growth rate represents the final realization. The advantage of this approach is that our sample size increases considerably compared to the approach by Auerbach and Gorodnichenko (2013), while we will show that our results are very similar.

4 Fiscal Multipliers in the Excessive Deficit Procedure

In this section, we report the results of our baseline estimation, discuss the underlying mechanisms and evaluate the EDP multiplier in bad times.

4.1 Baseline Results

Figure 1 shows the estimates for the cumulative fiscal multipliers from our baseline specification given by Equation (2). The left panel displays the linear cumulative multipliers for the whole

[7] For example, Auerbach and Gorodnichenko (2013) use the realization published in $t+4$.

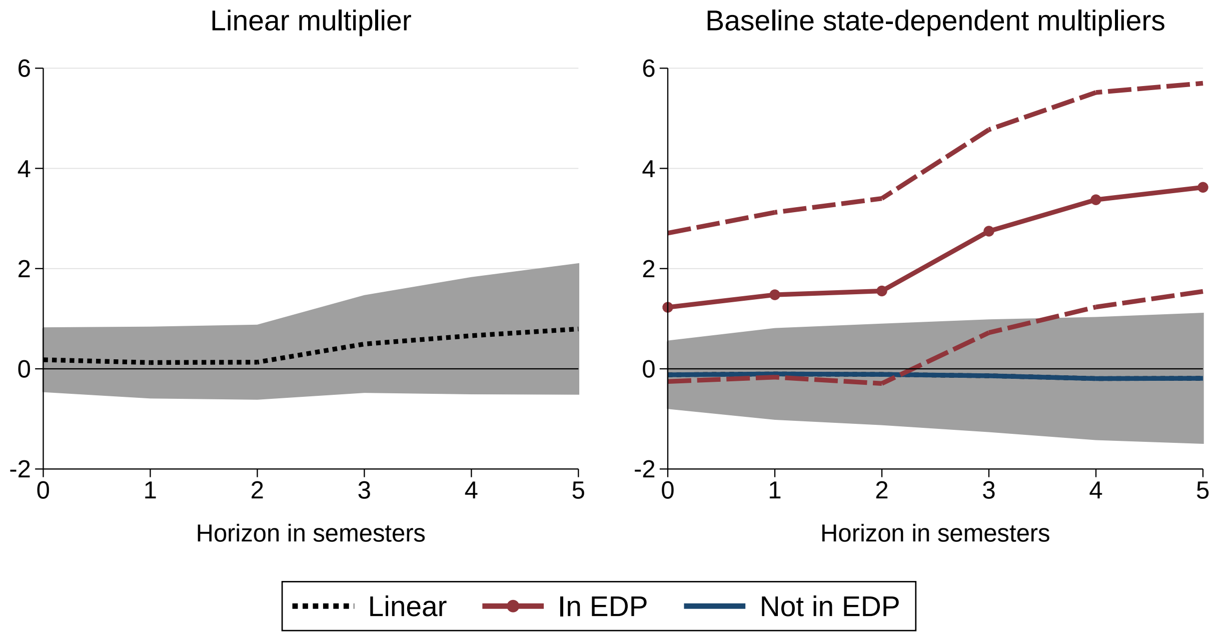

Figure 1: Linear and state-dependent cumulative fiscal multipliers

Notes: Estimates in both panels are based on the full sample with 463 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

sample, i.e., the multiplier estimates for the case without state dependence. We plot a horizon of five semesters and the shaded area refers to the 90% confidence intervals calculated from Driscoll-Kraay standard errors. The point estimates for the linear multipliers are positive and smaller than one across all horizons. The fiscal multiplier is 0.2 on impact and slightly increases over time reaching 0.7 after two years ($h = 4$), which is in the range of unconditional estimates for European countries commonly reported in the literature, see Mineshima et al. (2014) for an extensive survey.

The right panel of Figure 1 displays the cumulative state-dependent multipliers. The state indicates whether a country has been in an ongoing EDP in the period before the change in government spending. The red line refers to episodes in which countries are in an EDP, whereas the blue line is associated with countries which are not. The red dashed lines and the shaded area again indicate the 90% confidence intervals. The state-dependent multipliers for the EDP sample are positive and larger than one across all horizons. In other words, the cumulative GDP gain is larger than the underlying cumulative government spending following the impulse in period t. On impact, the fiscal multiplier is already 1.2 and further increases to 3.4 after two years. The multipliers for non-EDP episodes are essentially zero across all horizons.

In the first part of Table 2, we report the point estimates of the cumulative multipliers and the associated Driscoll-Kraay standard errors. In addition, we report the first-stage F-statistic (based on the test of Montiel Olea and Pflueger, 2013) which can be used to assess whether our

Table 2: Detailed results for the baseline specification

	Not in EDP			In EDP			Difference		
h	$M_{0,h}$	SE	F-stat.	$M_{1,h}$	SE	F-stat.	P. E.	p (DK)	p (AR)
Baseline (463 observations)									
0	−0.12	(0.40)	209.50	1.23	(0.90)	74.23	1.35	0.20	0.22
1	−0.10	(0.55)	163.86	1.48	(1.00)	61.13	1.58	0.22	0.24
2	−0.11	(0.61)	73.70	1.55	(1.12)	45.36	1.67	0.28	0.30
3	−0.14	(0.67)	54.67	2.75	(1.23)	33.30	2.89	0.08	0.12
4	−0.19	(0.74)	44.26	3.37	(1.30)	26.15	3.57	0.04	0.07
5	−0.19	(0.78)	42.13	3.62	(1.26)	24.10	3.81	0.02	0.05
Strict state definition (286 observations)									
0	0.08	(0.41)	114.64	0.68	(0.53)	133.75	0.60	0.24	0.31
1	0.02	(0.58)	94.87	1.00	(0.68)	125.83	0.99	0.17	0.24
2	−0.06	(0.60)	70.09	1.28	(0.78)	54.72	1.34	0.13	0.19
3	−0.24	(0.65)	51.05	2.78	(1.12)	39.71	3.02	0.01	0.05
4	−0.36	(0.69)	39.07	3.51	(1.19)	29.85	3.87	0.00	0.03
5	−0.32	(0.71)	37.55	3.60	(1.06)	25.86	3.92	0.00	0.02

Notes: We refer to the horizon by h. $M_{\bullet,h}$ denotes the point estimate of the multiplier in the respective state, "SE" the associated Driscoll-Kraay standard error, and "F-stat." the associated first-stage F-statistic. The critical values for the F-statistic are always 23.1 and 19.7 at the 5% and 10% significance level, respectively. We also report the point estimate of the difference between the two multipliers, "P. E.", and the associated Driscoll-Kraay (DK) and weak instrument robust Anderson-Rubin (AR) p-values.

instrumental variable is relevant in each state separately. Indeed, we find that the F-statistic is above the 5% critical value across all horizons and in each state. We therefore conclude that our instrument is relevant. Further, we report the point estimate for the difference and associated p-values which are heteroskedasticity and autocorrelation consistent (Driscoll-Kraay, DK) and which are robust in the presence of weak instruments (Anderson-Rubin, AR).[8] Both p-values confirm our visual observation from Figure 1. The difference is statistically significant at the 10% level from horizon three (four) onwards according to the DK (AR) p-values.

These results are not driven by (i) the method we use for the identification of government spending shocks, (ii) the definition of government spending, (iii) the trend GDP specification, (iv) the trend GDP timing, or (v) state-dependent time trends. First, we obtain very similar multipliers if we identify the shocks as originally proposed by Blanchard and Perotti (2002) or if the realization in Equation (3) is defined as in Auerbach and Gorodnichenko (2013). By using the latter identification approach, we show that the choice of the realization does not influence our results. Second, we confirm the finding of Boehm (2020) who reports that multipliers tend to be larger if one only uses government consumption and disregards government investment. This stresses the importance to incorporate government investment if it is consistently available

[8]We implement the AR test following Ramey and Zubairy (2018) and Bernardini and Peersman (2018).

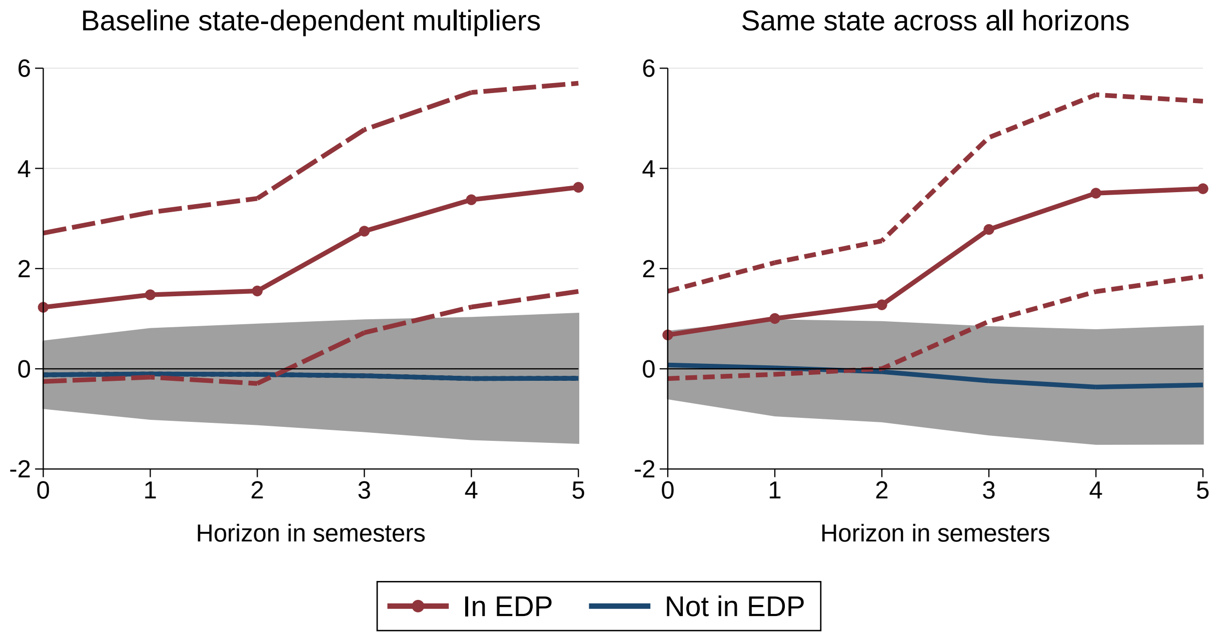

Figure 2: Cumulative fiscal multipliers being in the same state across all horizons

Notes: Estimates in the left panel are based on the full sample. Estimates in the right panel are based on 286 observations covering 15 countries. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

for a country. Third, we obtain very similar multipliers if we use alternative trend GDP series obtained from polynomials of order 2 or 4, the HP filter using the smoothing parameter $\lambda = 1600$ or the filter proposed by Hamilton (2018).[9] Fourth, multipliers of both states only slightly increase if we scale level variables in Equation (2) by different lags of trend GDP. This is indirect evidence that our baseline trend specification, a polynomial of order 3, does not vary much with the cycle. Finally, EDP multipliers are large if time trend coefficients can vary between states. We provide the results for these robustness checks in Appendix A.

The baseline multipliers shown in Figure 1 are based on our full sample and the estimation does not account for countries entering or leaving the EDP. Thus, the estimation does not necessarily include the same countries across all horizons. This could of course produce misleading estimates. For example, the GDP reaction could be always strongly positive just after the EDP ended. Figure 2 underlines that this is not the case. The left panel displays once again the baseline multipliers for comparison purposes. The right panel shows the multipliers for a specification with a strict definition of the state: We restrict the sample to countries remaining in the same state (EDP/non-EDP) across all horizons used in the estimation. Hence, we ensure that we only include countries and episodes for which we have observations in the semesters zero to five. The estimates are based on a smaller sample size, that is, 289 observations at each

[9]Based on Hamilton (2018), we set parameters to $h = 4$ and $p = 2$ (half of the parameters suggested for quarterly data).

horizon, but the result remains robust. Multipliers for EDP episodes are positive across all horizons and larger than one after one semester. The second part of Table 2 shows that the difference between EDP and non-EDP states becomes even slightly more pronounced and is now statistically significant at the 10% level for $h \geq 3$ according to the DK and AR p-values. The F-statistics for the state-dependent multipliers still show strong evidence that our instrumental variable is relevant.

One could be concerned that higher state-dependent fiscal spending multipliers in EDP are driven by significantly less government spending in the EDP and/or an asymmetric distribution of government spending shocks across the two states. A simple approach to test the first concern is to regress the share of government spending over GDP on the EDP dummy including country-specific constants. We find little evidence that the share is substantially different in the two states. In addition, we find little evidence that the shock distributions for the two states are a possible driver of the higher multipliers. The mean of the shocks occurring in the EDP (0.19 percentage points) is smaller than the mean of the shocks outside the EDP (0.35 percentage points) while the standard deviation is roughly the same (1.25 in the EDP and 1.18 outside the EDP). Restricting the sample only to the negative (positive) shocks also shows that the shock mean is always smaller in the EDP. Negative shocks occur as often in the EDP (40%) as outside the EDP (38%). Finally, we directly estimate state-dependent multipliers from negative shocks only. The multipliers in EDP are not different from our baseline results, see Appendix A. Hence, higher state-dependent fiscal multipliers in EDP are not driven by negative government spending shocks during EDP episodes and a contractionary shock affects the fiscal multipliers like the average shock in our baseline specification.

The results obtained from our baseline specification show that the fiscal multipliers are significantly larger for countries that are in an ongoing EDP. The multipliers in these countries are strictly larger than one, implying that government spending is more effective in the sense that the cumulative GDP gain exceeds the underlying cumulative government spending. We explore the mechanisms behind our results in the next section.

4.2 Mechanisms

The size of fiscal spending multipliers is determined by many factors and depends in particular on the dynamics of other macroeconomic variables. Our baseline estimates suggest that government spending is more effective for countries in EDP. In order to rationalize this finding, we investigate the dynamics of the other variables included in our empirical model. Figure 3 shows the impulse response functions of these variables to a 1% increase in government spending, as estimated from Equation (1). The first row displays the responses of government spending,

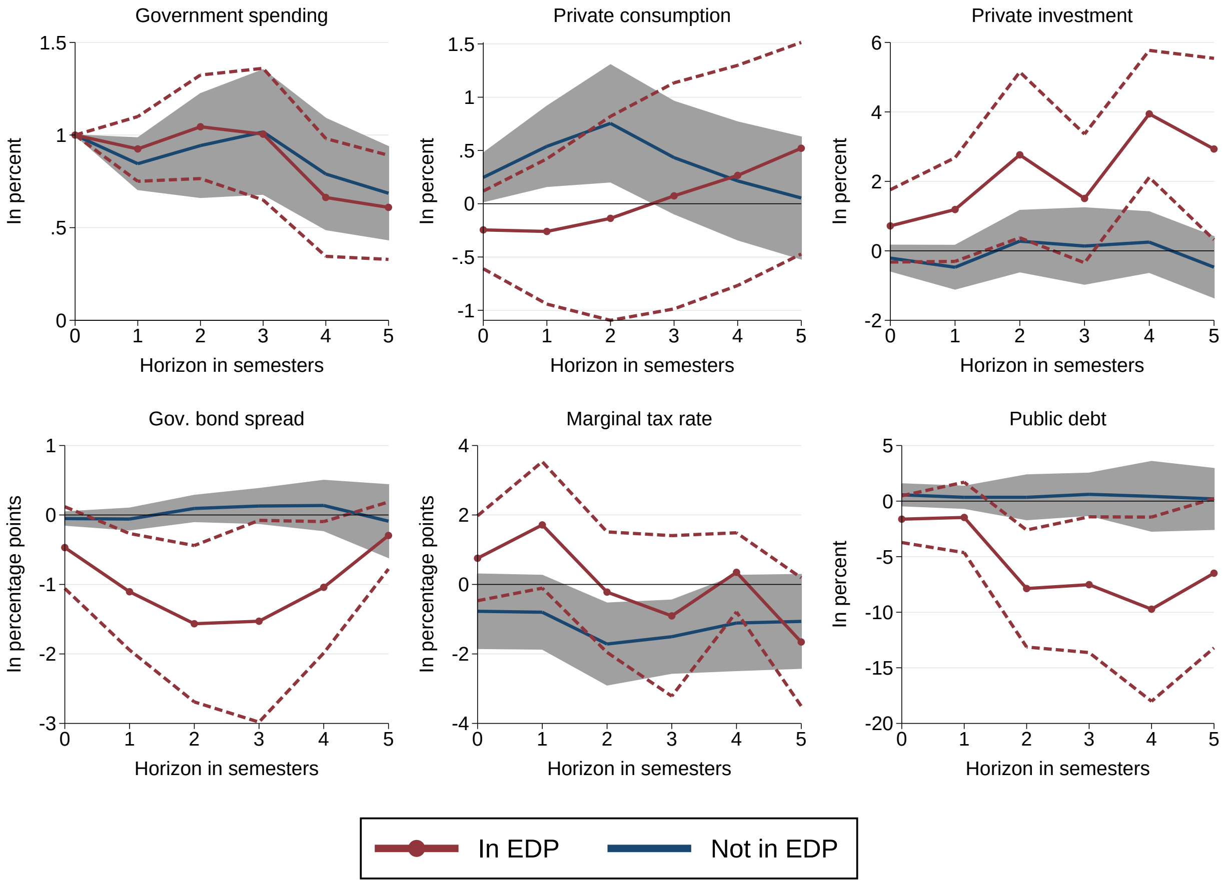

Figure 3: Impulse response functions

Notes: Estimates in each panel are based on the full sample. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

private consumption and private investment (measured in percent). The second row shows the responses of the government bond spread, the marginal tax rate (both measured in percentage points) and the public debt level (measured in percent).

The responses provide evidence that the difference in fiscal multipliers between the EDP and non-EDP samples is mainly driven by diverging investment dynamics. We observe similar investment dynamics in each of our robustness exercises (i)–(v). In the EDP sample, we observe a substantial (and mostly significant) positive response of private investment to a positive government spending shock. Private investment in the non-EDP sample, however, is essentially not reacting to the shock. The path of government spending is similar across states, while we observe a significant positive response of consumption in non-EDP episodes at shorter horizons.[10] The government bond spread (vis-à-vis Germany), as a measure for long-term interest rates, shows a clearly negative reaction for the EDP sample. This is in line with

[10]We do not observe government spending reversals, which is consistent with the reaction of the debt level in our sample, see Corsetti et al. (2012a).

the observed response of private investment: lower interest rates stimulate investment. Bond spreads do not respond to the government spending shock in non-EDP episodes. The initial rise in the marginal tax rate in the EDP sample reflects the reaction of the financing side of the government budget balance to the increase in government spending induced by the shock. Finally, we observe that countries significantly reduce their public debt level in response to the fiscal impulse during EDP episodes. This reaction is in accordance with the EDP's objective of encouraging countries to bring debt levels under control. There is no response of the public debt level in the non-EDP sample. Using the strict state definition from the previous section, we observe even more pronounced reactions of private investment, the bond spread and the public debt level during EDP episodes, see Appendix A.

Overall, we find evidence that fiscal multipliers are larger for countries in EDP because government spending provides a stronger stimulus to economic activity. In response to a positive spending shock, countries in EDP achieve a significant reduction of public debt and long-term interest rates decrease, signaling a stable fiscal outlook. This boosts private investment and gives rise to substantial crowding-in. One possible explanation is that being in the EDP demonstrates credible commitment to fiscal discipline and therefore leads to lower risk premia. Thus, the EDP fulfills its task as corrective arm of the EU fiscal framework, while at the same time ensuring the effectiveness of fiscal stimulus.

4.3 Fiscal Multipliers in Bad Times

The sample used in our baseline estimation includes a broad set of countries which might differ in many dimensions. For example, countries have varying fiscal positions or face different cyclical fluctuations over time. The EDP is designed to ensure stable public finances, which is more likely to be an issue in member countries going through bad times. Thus, the measures implemented by the EDP are supposed to be especially effective for these countries. Economic theory provides different explanations for the potentially different size of the fiscal multiplier in bad times. From a Keynesian point of view, multipliers are larger during times of slack because government spending is less likely to crowd out consumption and/or investment. From a neoclassical perspective, however, consumption decreases when government spending (and deficits) increase significantly since consumption depends on intertemporal optimization. In order to isolate the effect of the EDP on the multiplier in bad times, we estimate Equation (2) for different subsamples. For all subsamples, the relevant state variable is the EDP dummy.

We show the cumulative fiscal multipliers in Figure 4 and report detailed test results in Table 3. Supporting impulse responses can be found in Appendix B. The left panel of Figure 4 considers a subsample with episodes of countries which do not fulfill the Maastricht criteria for

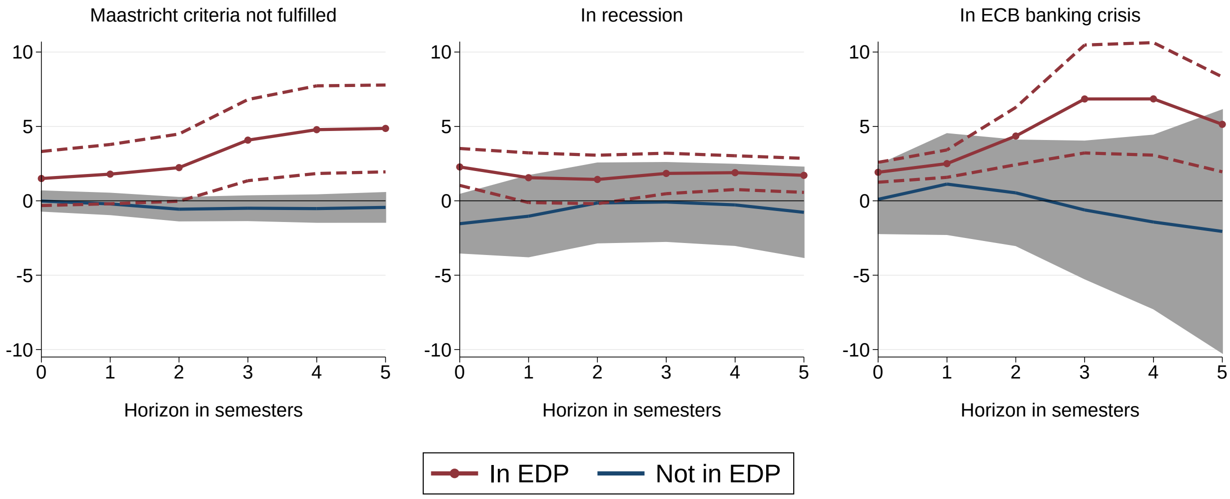

Figure 4: Cumulative fiscal multipliers in bad times

Notes: Estimates are based on the following number of observations and countries. Left: 249 observations and 14 countries. Middle: 92 observations and 15 countries. Right: 122 observations and 14 countries. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

either deficit only (15% of the observations) or debt only (68%) or both (17%). These countries have in common that they are in a weak fiscal position with a special emphasis on unsustainable debt levels as the debt exceeds the Maastricht criterion for 85% of the observations. The estimates of the fiscal multipliers indicate that the EDP multipliers are statistically different from the non-EDP multipliers at the 10% level for $h > 1$. This shows that the EDP indeed increases the effectiveness of government spending in this subsample. Multipliers for the EDP sample are even larger than in the specification using the full sample as shown in the right panel of Figure 1. This is due to a more pronounced reaction of interest rates and private investment. The first-stage F-statistic exceeds the 5% critical value at all horizons in the non-EDP state and the 10% critical value for $h \leq 4$ in the EDP state, suggesting that our results do not suffer from a weakly identified instrument. The results are similar if we condition the subsample only on the observations for which debt exceeds the Maastricht criterion.

The middle panel focuses on a subsample with recessionary episodes. We identify recessions using the simple and transparent algorithm proposed by Harding and Pagan (2002) based on quarterly real GDP growth from the OECD Main Economic Indicators 2020–01.[11] The EDP multipliers are significantly positive for countries in a recession, but considerably smaller than the baseline multipliers at longer horizons. This can be rationalized by a more persistent increase in government spending in response to the shock, along with a crowding-in of consumption and a slight crowding-out of private investment in this subsample. Note that the

[11] As originally suggested by Harding and Pagan (2002), we require that complete cycles have a length of at least five quarters and that a cycle phase lasts at least two quarters. We list the identified recessions which are included in the baseline sample in Appendix C.

Table 3: Detailed results for multipliers in bad times

	Not in EDP			In EDP			Difference		
h	$M_{0,h}$	SE	F-stat.	$M_{1,h}$	SE	F-stat.	P. E.	p-val (DK)	p-val (AR)
Maastricht criteria not fulfilled (249 observations)									
0	−0.01	(0.41)	375.84	1.50	(1.11)	58.97	1.51	0.19	0.20
1	−0.21	(0.43)	251.18	1.79	(1.21)	64.34	2.00	0.11	0.12
2	−0.56	(0.48)	74.40	2.23	(1.37)	42.58	2.80	0.05	0.05
3	−0.50	(0.50)	56.55	4.08	(1.66)	30.96	4.58	0.01	0.01
4	−0.52	(0.56)	43.84	4.78	(1.79)	22.74	5.31	0.00	0.01
5	−0.44	(0.60)	36.69	4.87	(1.77)	19.14	5.31	0.00	0.00
Debt criterion not fulfilled (210 observations)									
0	0.79	(0.60)	135.81	1.66	(0.96)	114.82	0.87	0.46	0.45
1	0.68	(0.72)	107.29	2.44	(1.27)	64.44	1.75	0.21	0.22
2	0.14	(0.90)	66.29	3.49	(1.70)	23.54	3.35	0.05	0.06
3	0.35	(1.11)	43.32	5.88	(2.26)	13.24	5.53	0.01	0.02
4	0.33	(1.20)	28.59	6.09	(2.33)	9.61	5.76	0.00	0.01
5	0.28	(1.15)	23.25	5.14	(1.87)	8.27	4.86	0.00	0.00
In recession (92 observations)									
0	−1.54	(1.19)	50.98	2.28	(0.75)	34.06	3.82	0.02	0.15
1	−1.03	(1.66)	41.95	1.55	(1.02)	22.55	2.59	0.25	0.33
2	−0.14	(1.63)	50.32	1.44	(0.99)	13.47	1.58	0.46	0.47
3	−0.08	(1.61)	28.28	1.84	(0.83)	9.76	1.92	0.31	0.31
4	−0.27	(1.65)	12.30	1.89	(0.69)	8.03	2.17	0.20	0.18
5	−0.77	(1.84)	6.03	1.71	(0.70)	7.52	2.48	0.13	0.11
In ECB banking crisis (122 observations)									
0	0.10	(1.40)	96.05	1.92	(0.41)	46.85	1.82	0.19	0.24
1	1.13	(2.06)	81.63	2.50	(0.56)	30.32	1.37	0.50	0.51
2	0.54	(2.15)	20.97	4.35	(1.17)	12.92	3.82	0.12	0.15
3	−0.61	(2.81)	13.93	6.85	(2.21)	11.16	7.46	0.06	0.07
4	−1.42	(3.55)	6.27	6.85	(2.30)	9.97	8.28	0.06	0.06
5	−2.06	(4.97)	2.00	5.14	(1.94)	9.38	7.20	0.14	0.10

Notes: We refer to the horizon by h. $M_{\bullet,h}$ denotes the point estimate of the multiplier in the respective state, "SE" the associated Driscoll-Kraay standard error, and "F-stat." the associated first-stage F-statistic. The critical values for the F-statistic are always 23.1 and 19.7 at the 5% and 10% significance level, respectively. We also report the point estimate of the difference between the two multipliers, "P. E.", and the associated Driscoll-Kraay (DK) and weak instrument robust Anderson-Rubin (AR) p-values.

EDP/non-EDP multipliers are not statistically different and the first-stage F-statistic suggests that the instrumental variable is less relevant at longer horizons.

Finally, the right panel shows the multipliers for periods in which countries are hit by a banking crisis, as defined in the European Financial Crises Database provided by the European Central Bank and the European Systemic Risk Board (Lo Duca et al., 2017).[12] Fiscal multipliers

[12] The most recent version of the database was published in 2017. We extend the data using the warnings issued by the European Systemic Risk Board. In fact, the European Systemic Risk Board did not identify any banking crises in 2018 and the first half of 2019.

in the EDP sample are significantly positive and even larger than in the baseline estimation at longer horizons due to strong crowding-in of private investment. At the same time, the instrumental variable is less relevant for these horizons. Multipliers for the non-EDP sample are not significantly different from zero across all horizons. However, the difference between the two states is only statistically significant at the 10% level for horizons three and four.

The cumulative fiscal multiplier estimates for these subsamples confirm that the EDP is successful in increasing effectiveness of government spending for countries in bad times. In particular, the EDP seems to be fully functional for countries with a weak fiscal position (as indicated by non-compliance with the Maastricht criteria). This indicates that the procedure fulfills the purpose it was designed for.

5 Alternative State Variables

We have so far implicitly assumed that the indicator whether a member state is in the EDP or not has explanatory power for the differences in fiscal multipliers across countries. However, countries in the EDP could tend to have weak fiscal positions or to experience recessions more frequently than other countries, both of which could in turn explain the variations in fiscal multipliers. Potentially, the EDP/non-EDP states could simply be a proxy for different underlying state variables. In Figure 5, we present fiscal multipliers for various alternative state variables. That is, we re-estimate Equation (2) using different indicator variables $\mathcal{I}_{i,t-1}$. The estimates refer to cumulative two-year multipliers (horizon $h = 4$) with the corresponding 90% confidence intervals. For comparison purposes, the first row shows the two-year multipliers from our baseline specification. The other rows report the multipliers if different combinations of non-compliance with the Maastricht criteria, recessions or banking crises are used as state variables. The recession and banking crisis dummy variables are constructed as described in the previous section. We provide corresponding impulse response functions in Appendix D.

Again, non-compliance with the Maastricht criteria is interpreted as a signal for a weak fiscal position. We observe that the corresponding multipliers are smaller than the EDP multipliers, but still positive. This stands in contrast to the findings reported in the literature. Corsetti et al. (2012b), Nickel and Tudyka (2014) and Banerjee and Zampolli (2019) find that crowding-out of private investment leads to lower and even negative multipliers in countries with weak fiscal positions, while Huidrom et al. (2020) explains the lower multipliers in these countries by a decrease in consumption. In this context, Banerjee and Zampolli (2019) and Huidrom et al. (2020) document the relevance of the transmission via the interest rate channel: Interest rates in high-debt countries rise in response to fiscal stimulus, reducing investment and consumption.

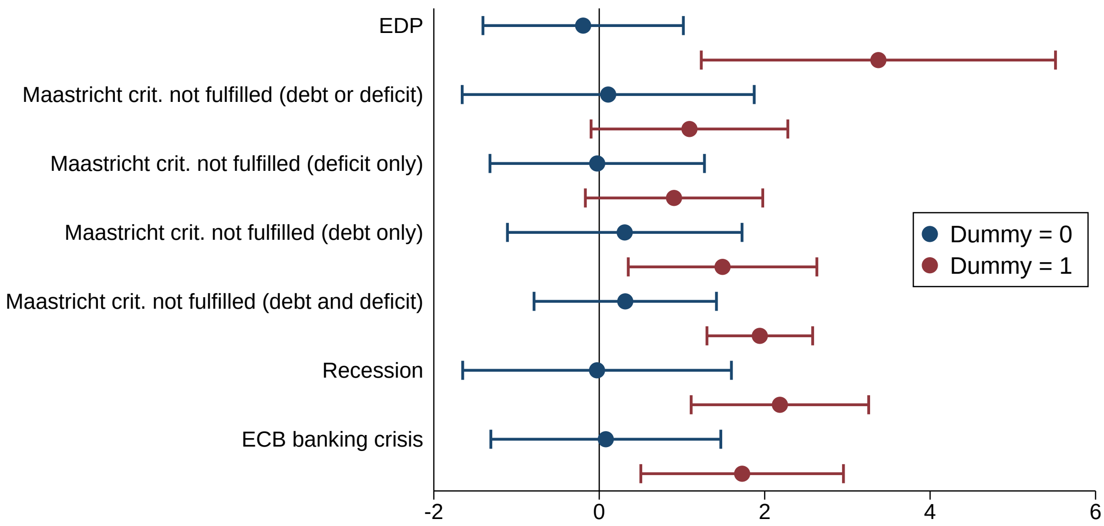

Figure 5: Cumulative two-year fiscal multipliers ($h = 4$)

Notes: Full sample for each state dependency. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

By contrast, we observe a decrease in interest rates after a positive government spending shock for non-compliers with the Maastricht criteria which leads to crowding-in of private investment, thereby rationalizing positive multipliers.

Our results from Section 4.3 suggest that the EDP can explain these dynamics. In fact, Figure 4 and corresponding impulse responses show that the multipliers are positive, interest rates are lower and investment is crowded in only if a country with a weak fiscal position is in the EDP. However, for countries with weak fiscal positions outside the EDP, the multipliers are close to zero and private investment is crowded out due to a slight increase in interest rates. One possible explanation is that being in the EDP is a positive signal to investors that sovereign credit risk is reduced. This would lower risk premia and interest rates and in turn boost investment.

In line with Auerbach and Gorodnichenko (2012, 2013), we find significantly positive multipliers in recessions driven by crowding-in of private consumption and investment. We also find positive multipliers during banking crises. As in Corsetti et al. (2012b), we observe a positive reaction of investment for these episodes. Note that the multipliers associated with these states are smaller than the EDP multipliers.

The fiscal multipliers for all these alternative states are positive, suggesting that these state variables play a role. However, the multipliers are not significantly different from the respective other state (except for our baseline, the EDP state variable) and the estimated magnitudes are at most half the size of the EDP multiplier. These results confirm that the EDP is not only a proxy for different underlying factors. The EDP seems to be a suitable state variable for explaining variations in fiscal multipliers.

6 Fiscal Multipliers in Real Time

Using the rich availability of forecasts in different EO editions, we want to explore in this section how fiscal spending multipliers are observed in real time. This is relevant because policy makers may plan government spending based on the size of future fiscal spending multipliers implied by the projected paths of real GDP and real government spending at a specific time. For example, policy makers expect that an additional unit of government spending pays off more if the forecasts of the multipliers are large.

There is little evidence on how state-dependent fiscal multipliers are observed in real time. Blanchard and Leigh (2013, 2014) and Górnicka et al. (2020) investigate fiscal multipliers for European countries in the period of the European sovereign debt crisis. Both find that multipliers tend to be larger than initially forecasted which they attribute to the learning of the forecasters during the crisis. While Blanchard and Leigh (2013, 2014) provide evidence that multipliers exceed one, Górnicka et al. (2020) cannot confirm this finding. To our knowledge, we are the first to present evidence on how fiscal multipliers are observed in real time in different states. In particular, we can thereby investigate whether forecasters learn in a similar fashion in each state.

We use a version of Equation (2) in which we estimate real-time multipliers $M^s_{0,h}$ and $M^s_{1,h}$ by including real-time data from EO edition s:

$$\begin{aligned}\sum_{j=0}^{h} Y^s_{i,t+j} \;=\; &(1-\mathcal{I}_{i,t-1})\left[\alpha^s_{0,i,h} + M^s_{0,h}\sum_{j=0}^{h} G^s_{i,t+j} + \Phi^s_{0,h}(L)X^s_{i,t-1}\right] \\ &+ \mathcal{I}_{i,t-1}\left[\alpha^s_{1,i,h} + M^s_{1,h}\sum_{j=0}^{h} G^s_{i,t+j} + \Phi^s_{1,h}(L)X^s_{i,t-1}\right] + \sum_{k=1}^{2}\psi^s_k T^k_t + \varepsilon^s_{i,t+h}.\end{aligned} \tag{4}$$

Our approach features three distinct dimensions which explicitly account for the available information set of the forecasters at time s. First, the variables $Y^s_{i,t}$ and $G^s_{i,t}$ are now forecasts (nowcasts) if $t > s$ ($t = s$). Variables, for which $t < s$ is true, are ex-post values. Second, the instrument should contain only the information available up to time s, too. We therefore compute the forecast error in Equation (3) by using the realized growth rate as of time s, $g^s_{i,t}$, instead of the realized growth rate reported in EO edition 106, $g^{2019H2}_{i,t}$. Hence, the forecast error comparing the forecast and the realization is well defined for $s = t + k$, with $k \geq 1$:

$$FE^s_{i,t} = g^s_{i,t} - g^{t-1}_{i,t}. \tag{5}$$

Third, several studies show that trend estimates are unreliable in real time (e.g., Orphanides and Norden, 2002; Orphanides, 2003). Given this uncertainty around the trend estimates,

Table 4: Detailed results for multipliers in real time

	Not in EDP			In EDP			Difference		
h	$M^s_{0,h}$	SE	F-stat.	$M^s_{1,h}$	SE	F-stat.	P. E.	p-val (DK)	p-val (AR)
Expected multipliers for $s = t + 1$									
0	−0.32	(0.53)	70.73	−0.36	(0.48)	112.26	−0.04	0.96	0.96
1	−0.37	(0.66)	40.05	−0.43	(0.62)	47.41	−0.06	0.94	0.94
2	−0.29	(0.75)	26.46	−0.49	(0.76)	26.79	−0.20	0.83	0.83
3	−0.09	(0.81)	21.23	−0.48	(0.80)	24.17	−0.40	0.70	0.70
4	0.09	(0.85)	18.58	−0.46	(0.78)	23.91	−0.55	0.60	0.60
Ex-post multipliers for $s = t + 5$									
0	0.46	(0.68)	358.18	0.73	(0.62)	114.06	0.27	0.80	0.80
1	0.56	(0.72)	285.62	1.31	(0.67)	55.67	0.75	0.48	0.50
2	0.67	(0.75)	200.73	1.87	(0.68)	48.25	1.20	0.28	0.29
3	0.66	(0.78)	146.49	2.16	(0.78)	39.65	1.50	0.21	0.21
4	0.53	(0.81)	104.98	2.35	(0.80)	33.66	1.82	0.13	0.13

Notes: We refer to the horizon by h. $M^s_{\bullet,h}$ denotes the point estimate of the multiplier in the respective state, "SE" the associated Driscoll-Kraay standard error, and "F-stat." the associated first-stage F-statistic. The critical values for the F-statistic are always 23.1 and 19.7 at the 5% and 10% significance level, respectively. We also report the point estimate of the difference between the two multipliers, "P. E.", and the associated Driscoll-Kraay (DK) and weak instrument robust Anderson-Rubin (AR) p-values.

especially at the sample end, it is sensitive to scale level variables by a lag of the estimated trend rather than by the contemporaneous estimated trend.[13]

Figure 6 shows the real-time fiscal multipliers for $s = t + 1$ (first column) and $s = t + 5$ (second column). We further include the impulse responses of government spending and GDP for each specification. Our choice of s implies that the vector of control variables $X^s_{i,t-1}$ always contains past values.[14] Multipliers to the left of the dashed line in the first column are based on past values, while the multipliers to the right of the dashed line are partially based on nowcasts ($h = 1$) or on nowcasts and forecasts ($h \geq 2$). In the right column, multipliers are entirely based on ex-post values. We report corresponding statistics in Table 4 and supporting impulse responses can be found in Appendix E.

Fiscal multipliers in the upper right panel are similar to the multipliers from Figure 1 because both rely on ex-post values. One possible reason for the smaller EDP multipliers in Figure 6 is that the levels of GDP and government spending at $h = 4$ are still based on the first release and can be subject to further revisions. The EDP multipliers are still larger than

[13]This approach is valid as long as the trend does not fluctuate with the business cycle. Indeed, our robustness checks in Appendix A document that our baseline results from Figure 1 remain if we re-estimate the multipliers with level variables which are scaled now by different lags of the estimated trend.

[14]As the marginal tax rate is not available in real time, we include observations from EO edition 106 and thereby ignore possible revisions of past values. Given the nature of the variable, we believe that these revisions are negligible.

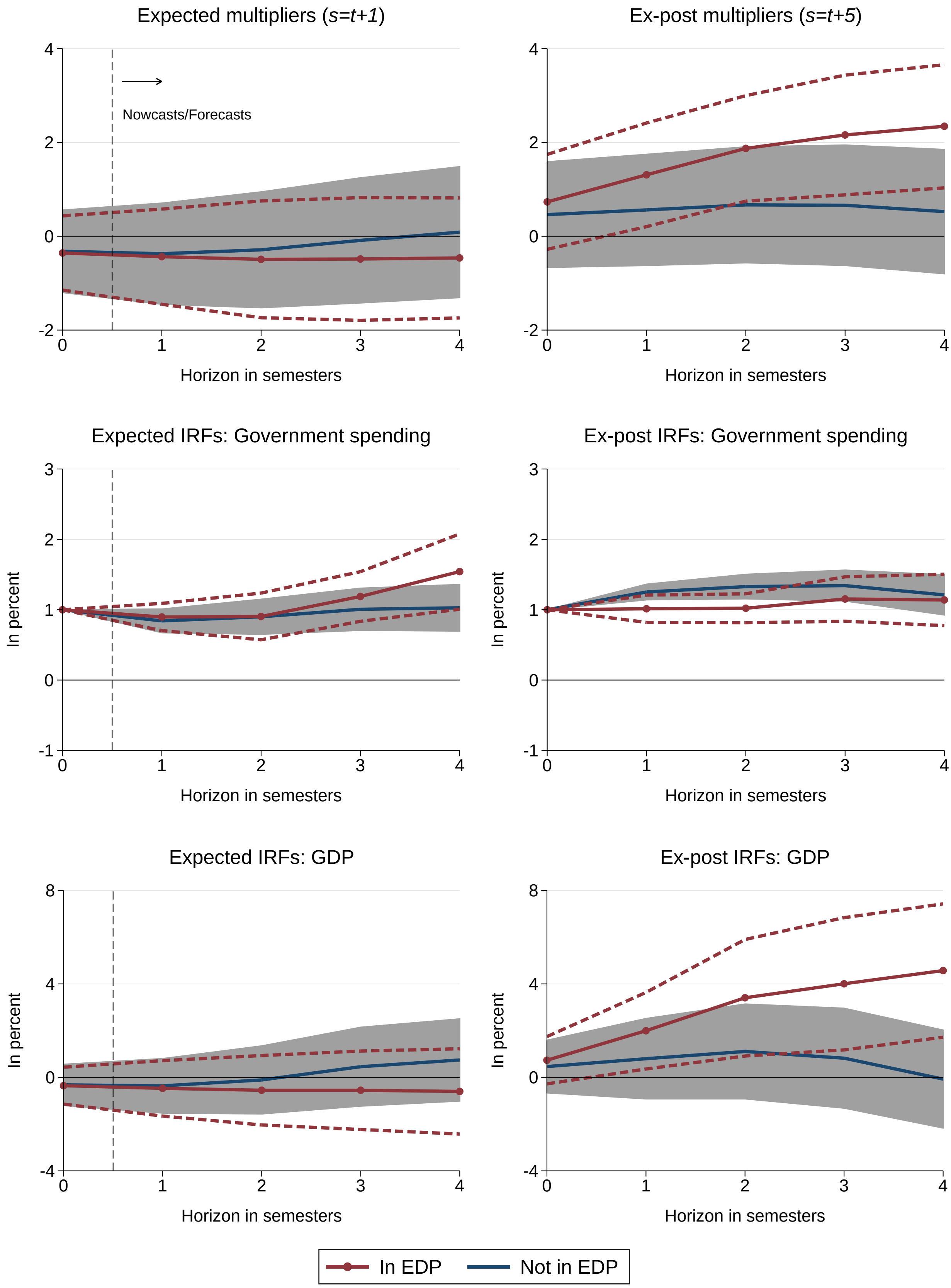

Figure 6: Multipliers and impulse responses in real time

Notes: Estimates in each panel are based on 373 observations covering 15 countries. In the left column, multipliers and impulse response functions (IRFs) to the left (right) of the vertical dashed line depend on ex-post real-time data (nowcasts/forecasts). In the right column, multipliers and IRFs always depend on ex-post real-time data. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

one and significantly positive after one semester. The non-EDP multipliers are smaller than one and not significantly different from zero again. In the upper left panel, fiscal multipliers at horizons $h \geq 1$ depend on nowcasts and forecasts for GDP and government spending. The state-dependent multipliers are not very different from each other, confidence intervals overlap and both are not significantly different from zero. The two very distinct EDP multipliers in the first row indicate that fiscal multipliers in the EDP are underestimated in real time. Hence, fiscal stimulus is expected to be less effective than it turns out ex post. The multipliers outside the EDP, however, do not depend on the time of the estimation.

With respect to the impulse responses of government spending and GDP in the EDP, we can trace the underestimation of the multipliers to a lower-than-expected government spending path and a larger-than-expected GDP path. The revisions of the GDP path seem larger than the revisions of the government spending path between the two specifications. Additionally, higher-than-expected private consumption and investment contribute to the change of the GDP response. Impulse responses for the interest spread are not significant but indicate that the spread is ex post lower than expected and therefore boosts private investment.

7 Conclusion

We estimate state-dependent fiscal spending multipliers to evaluate the effect of the EDP for 17 EU countries between 2000 and 2019. We show that fiscal multipliers in the EDP are larger than one and significantly different from the multipliers outside the EDP. The analysis of the underlying mechanisms shows that the higher multipliers are mainly driven by the crowding-in of investment which goes along with a significant reduction of public debt and a decrease of long-term interest rates in response to a positive government spending shock. The latter two reactions signal a stable fiscal outlook. Thus, the EDP fulfills its task as corrective arm of the EU fiscal framework, while at the same time ensuring the effectiveness of fiscal stimulus. In addition, we find that the EDP is especially successful in bad times. Furthermore, we show that it is not just a proxy for other underlying factors. Finally, we provide evidence that forecasters underestimate fiscal multipliers in the EDP in real time.

Our results have important policy implications. First, the EDP fulfills the function it was designed for. The output response to government spending is stronger for countries in the procedure. Second, the large EDP multipliers for a country in a weak fiscal position show that the EDP is more effective in bad times. Third, the underestimation of fiscal multipliers in real time masks the ex-post effect of fiscal stimulus. This could mislead policy makers who expect that a change in government spending does not have a substantial effect on the economy.

A Robustness of Results

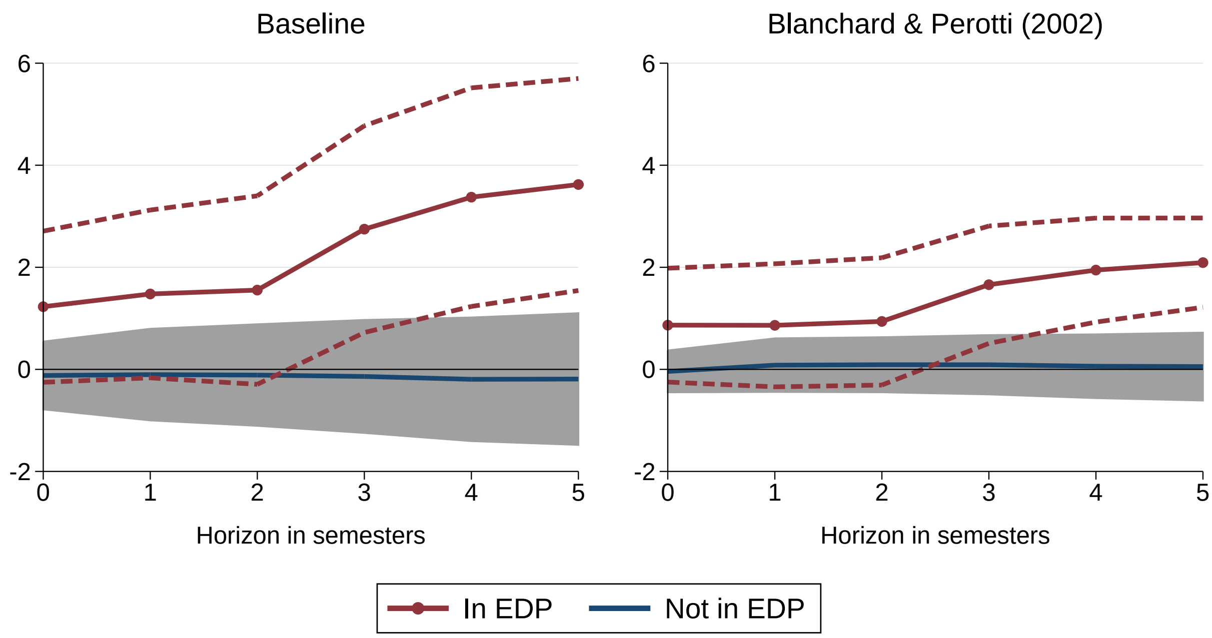

Figure A-1: Comparison of multipliers obtained from our baseline identification and from the identification of Blanchard and Perotti (2002)

Notes: The left panel repeats the multipliers from Figure 1. The right panel shows the multipliers using the shock identification of Blanchard and Perotti (2002). Estimates in both panels are based on the full sample with 463 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

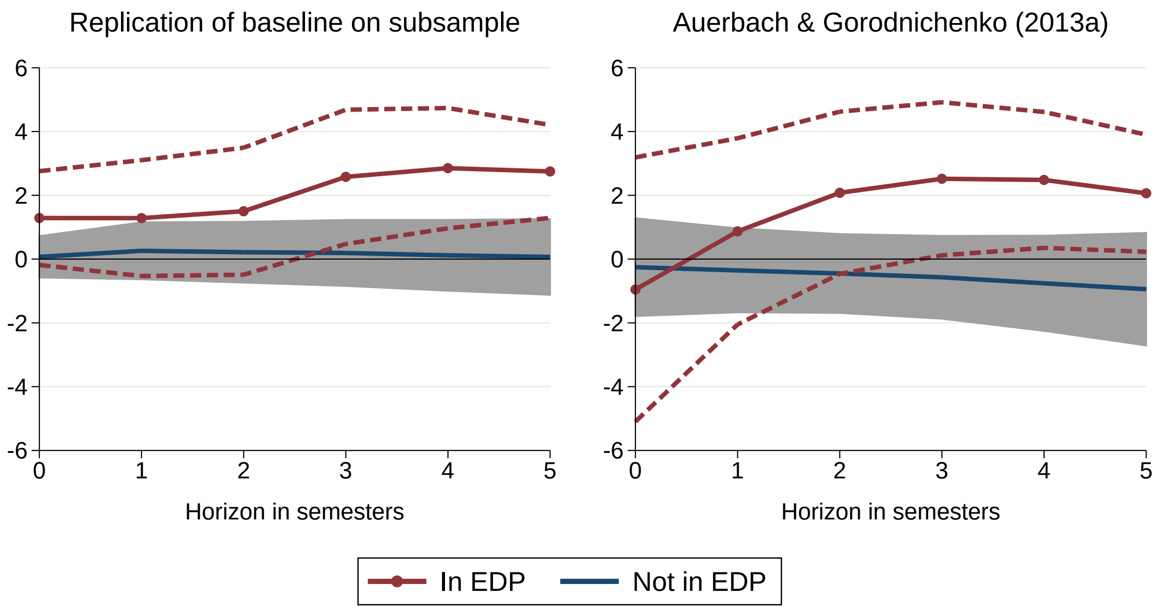

Figure A-2: Comparison of multipliers obtained from our baseline identification and from the identification of Auerbach and Gorodnichenko (2013)

Notes: The left panel displays the multipliers from Figure 1 based on the subsample which the left and right panel have in common. The right panel shows the multipliers using the shock identification of Auerbach and Gorodnichenko (2013). Estimates in both panels are based on 406 observations covering 15 countries. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

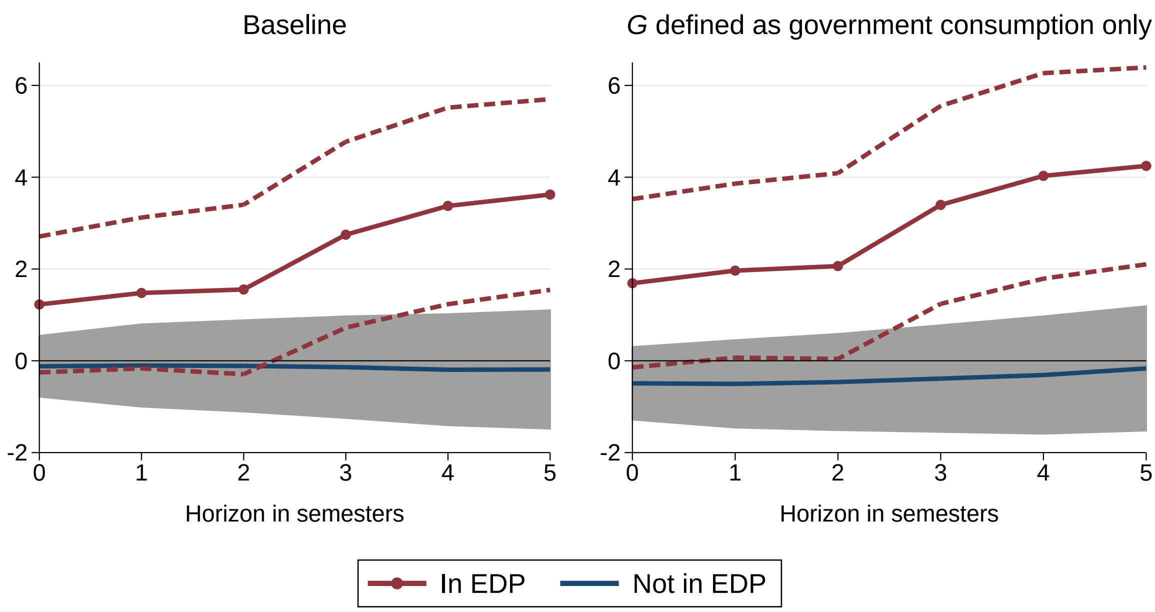

Figure A-3: Comparison of multipliers obtained from our baseline identification and from the definition of G as government consumption only

Notes: The left panel repeats the multipliers from Figure 1. The right panel shows the multipliers using G defined as government consumption only. Estimates in both panels are based on the full sample with 463 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

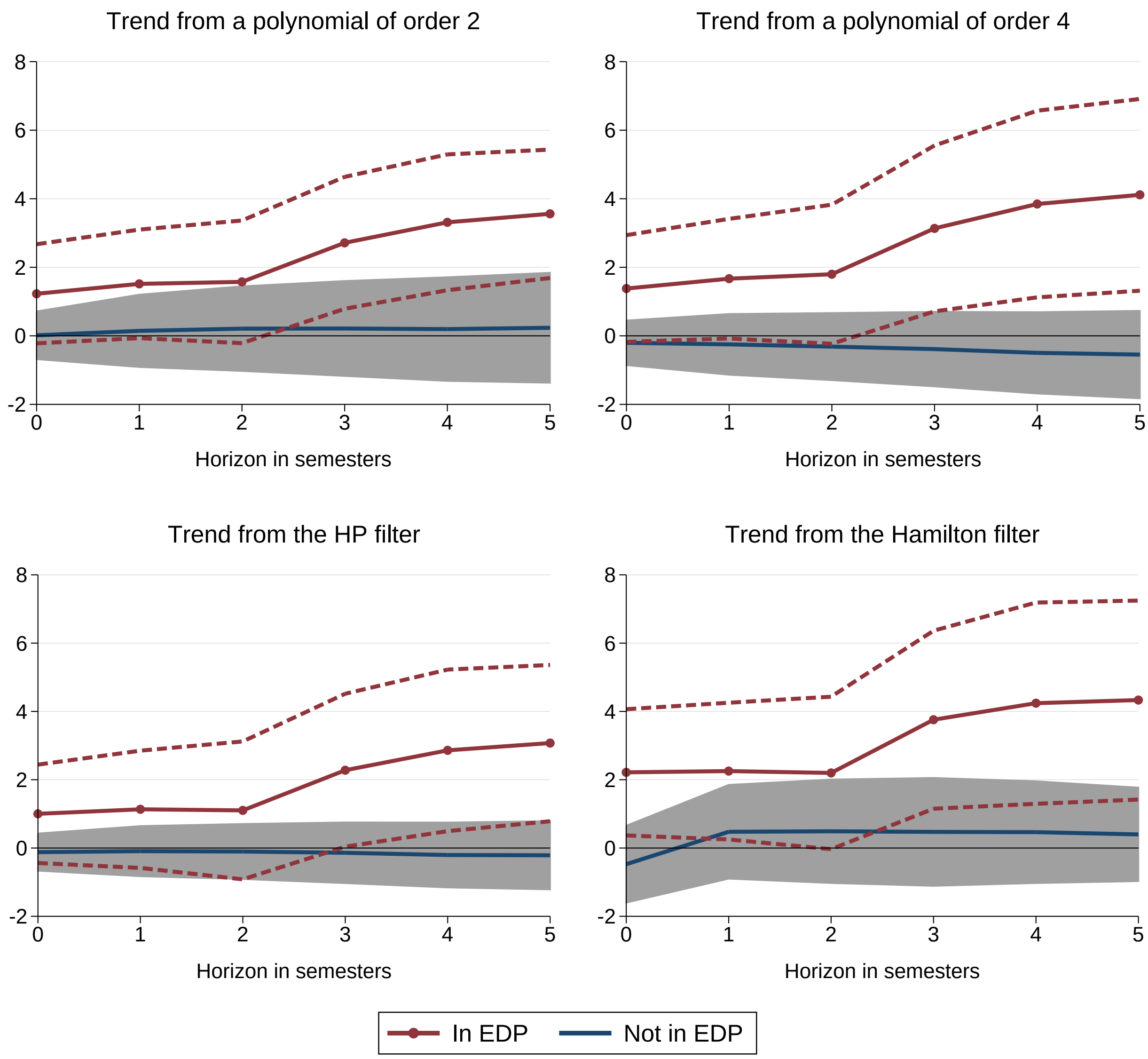

Figure A-4: Comparison of multipliers using different trend specifications

Notes: Each panel shows multipliers using different specifications to estimate trend GDP. Estimates in all panels are based on the full sample with 463 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

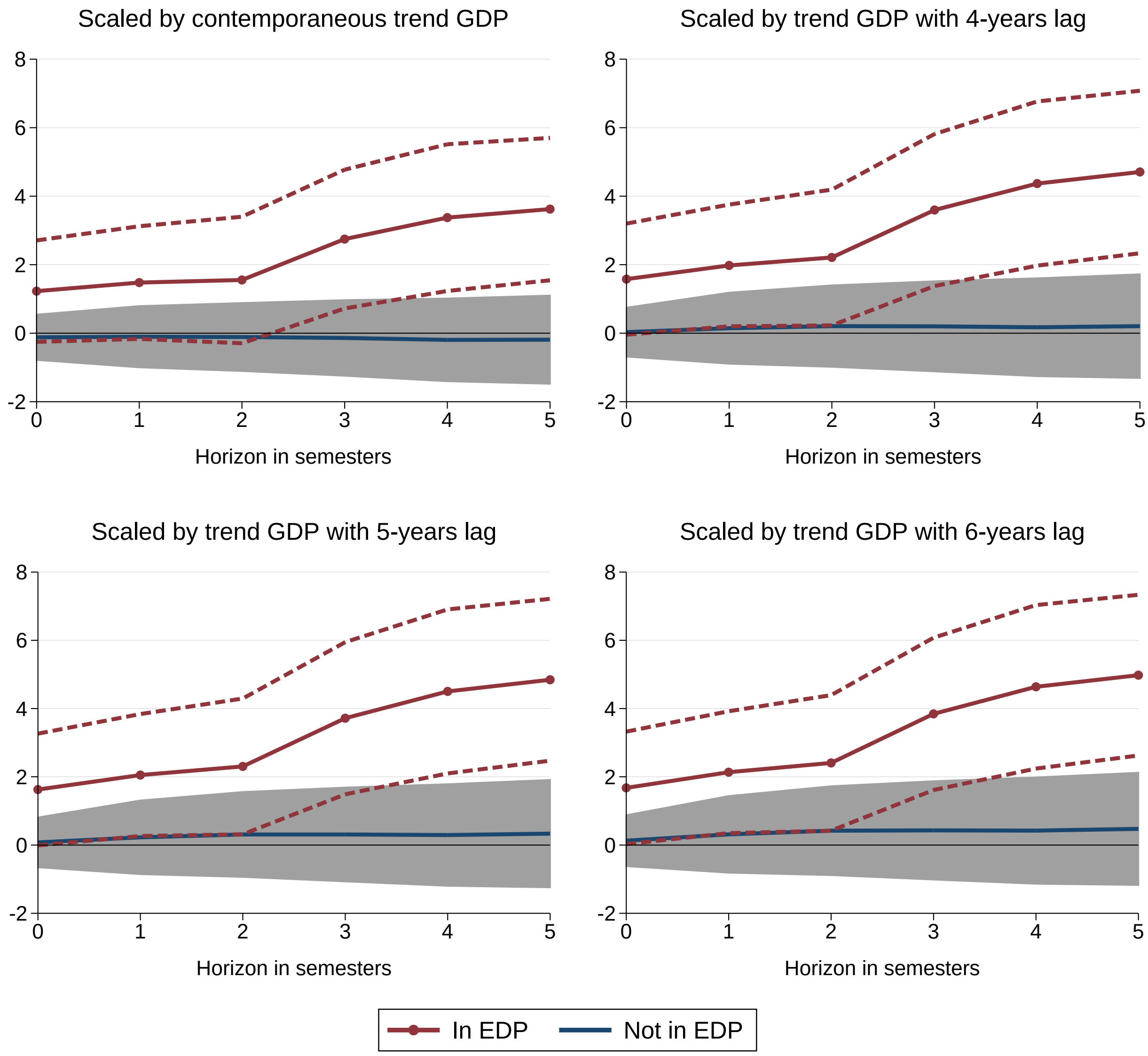

Figure A-5: Comparison of multipliers using different lags of trend GDP

Notes: The upper left panel repeats the multipliers from Figure 1. The other panels show the multipliers using different lags of trend GDP to scale level variables. Estimates in all panels are based on the full sample with 463 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

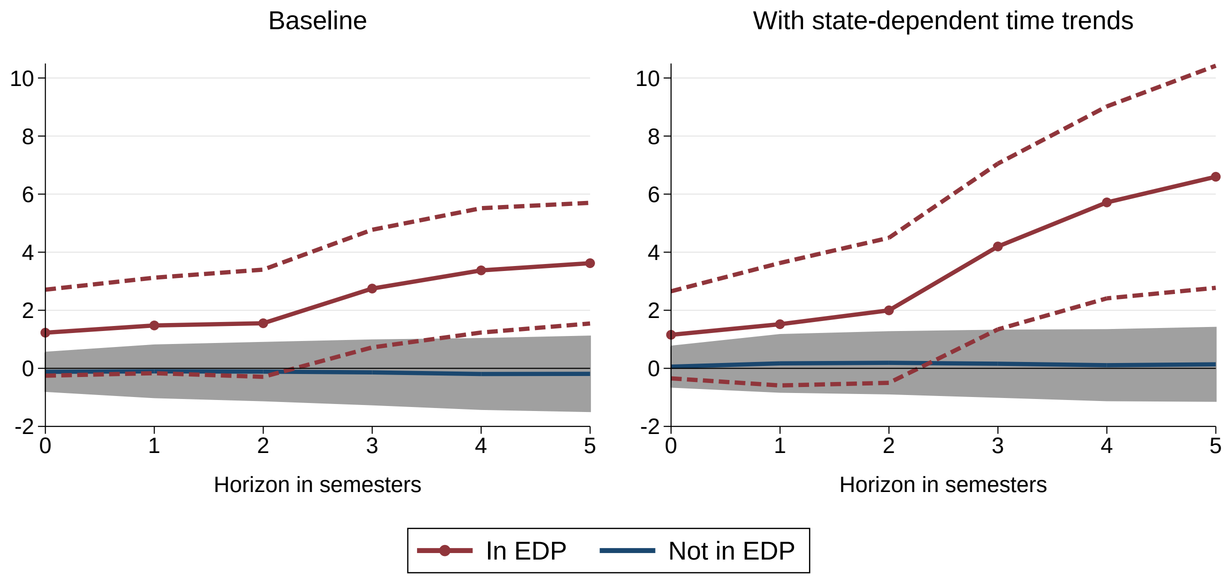

Figure A-6: Comparison of multipliers obtained from our baseline identification and an identification with state-dependent time trends

Notes: The left panel repeats the multipliers from Figure 1. The right panel shows the multipliers with state-dependent time trends. Estimates in both panels are based on the full sample with 463 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

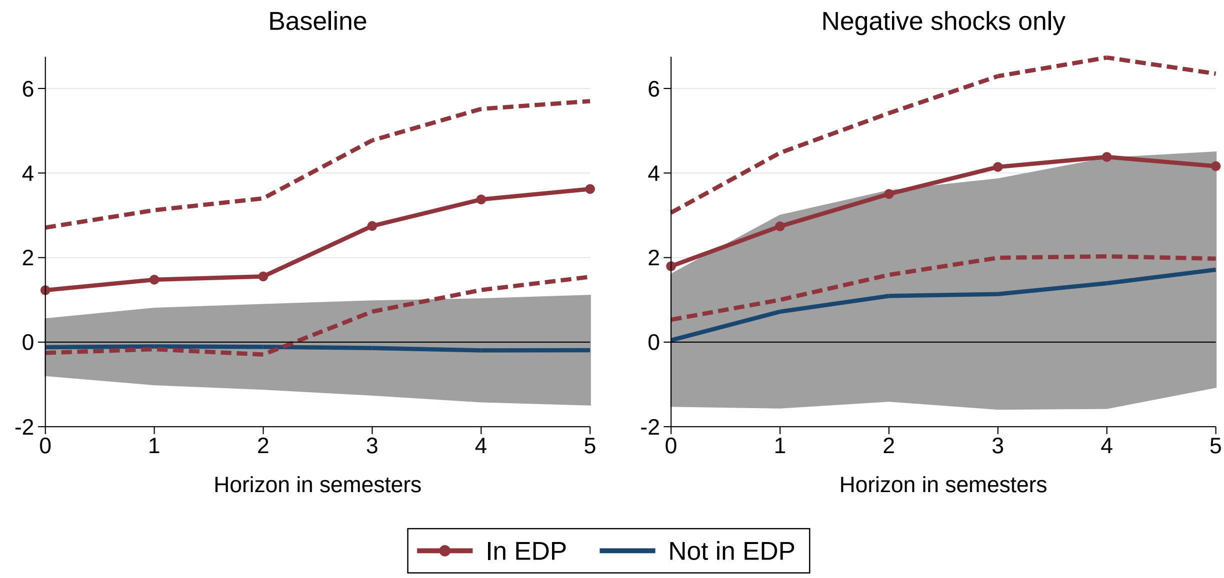

Figure A-7: Comparison of multipliers obtained from our baseline identification and from negative shocks only

Notes: The left panel repeats the multipliers from Figure 1. The right panel shows the multipliers from negative shocks only. Estimates in the left (right) panel are based on 463 (180) observations covering 17 countries. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

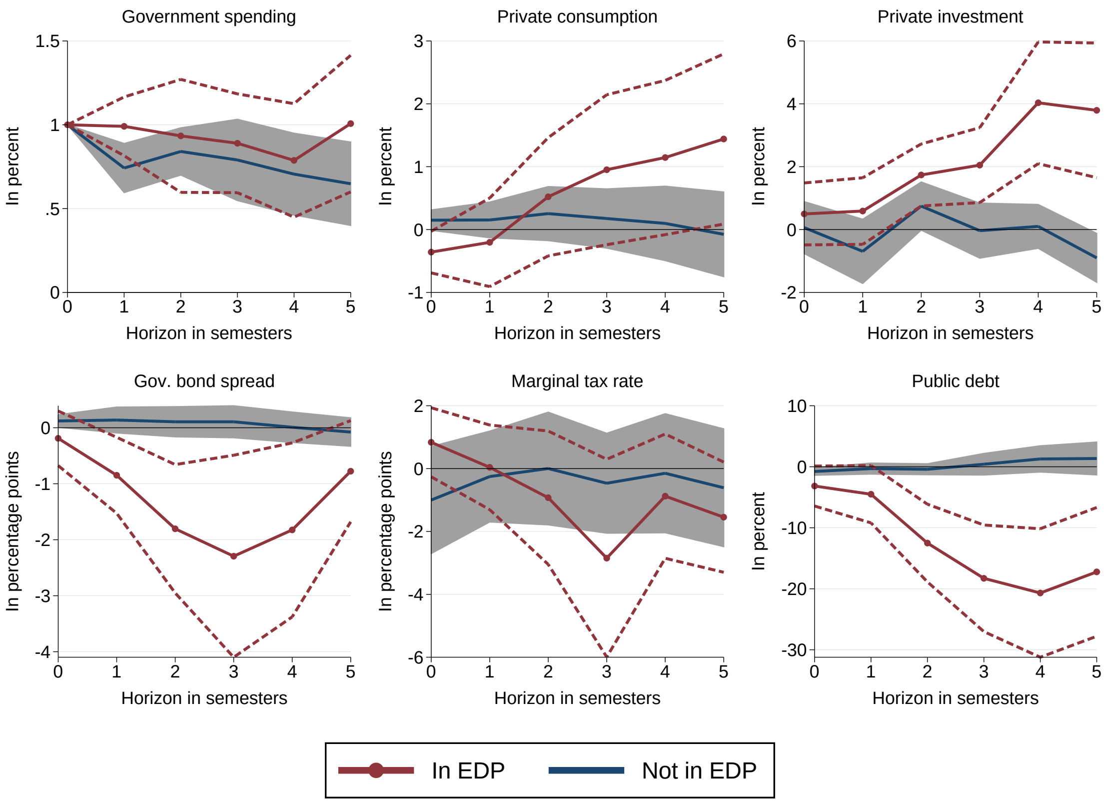

Figure A-8: IRFs for strict-state definition

Notes: Estimates in each panel are based on 286 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

B Detailed Results for Multipliers in Bad Times

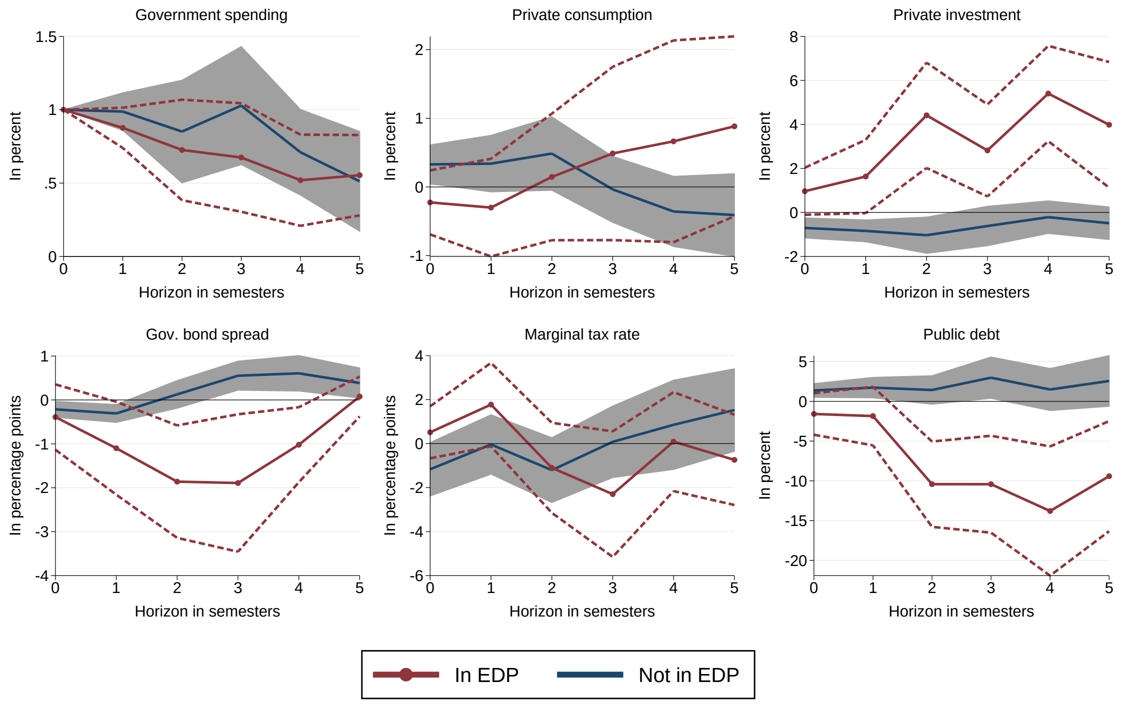

Figure B-1: IRFs in bad times – Maastricht criteria not fulfilled (debt or deficit)

Notes: Estimates in each panel are based on 249 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

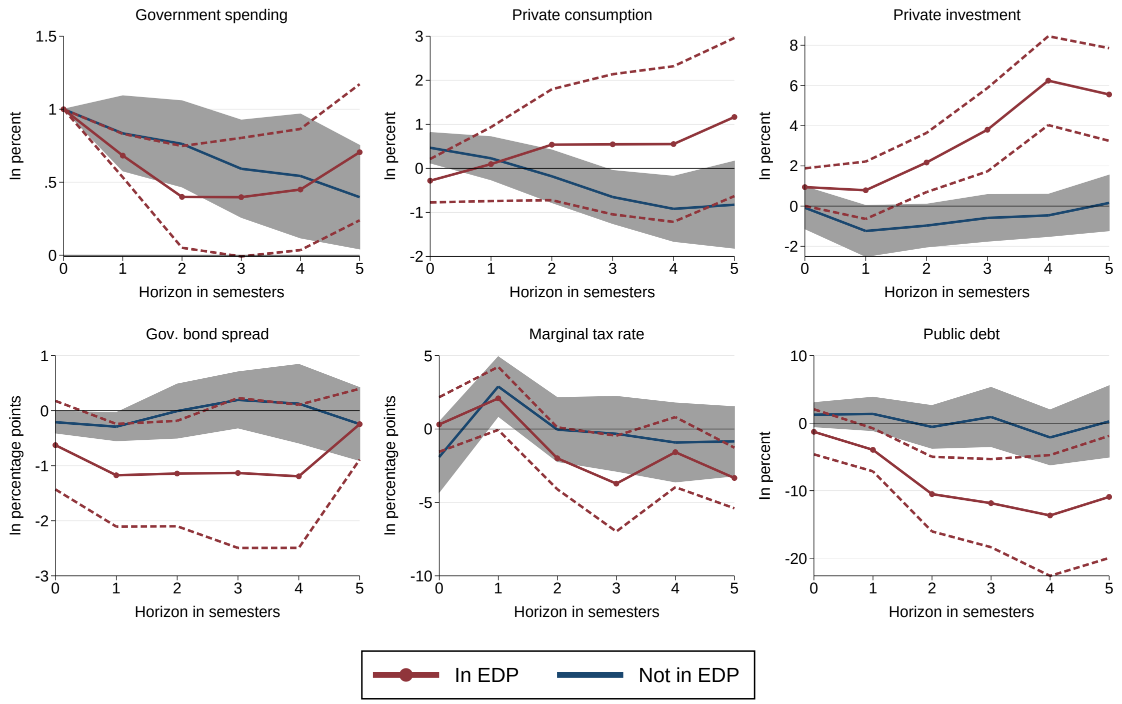

Figure B-2: IRFs in bad times – Debt criterion not fulfilled

Notes: Estimates in each panel are based on 210 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

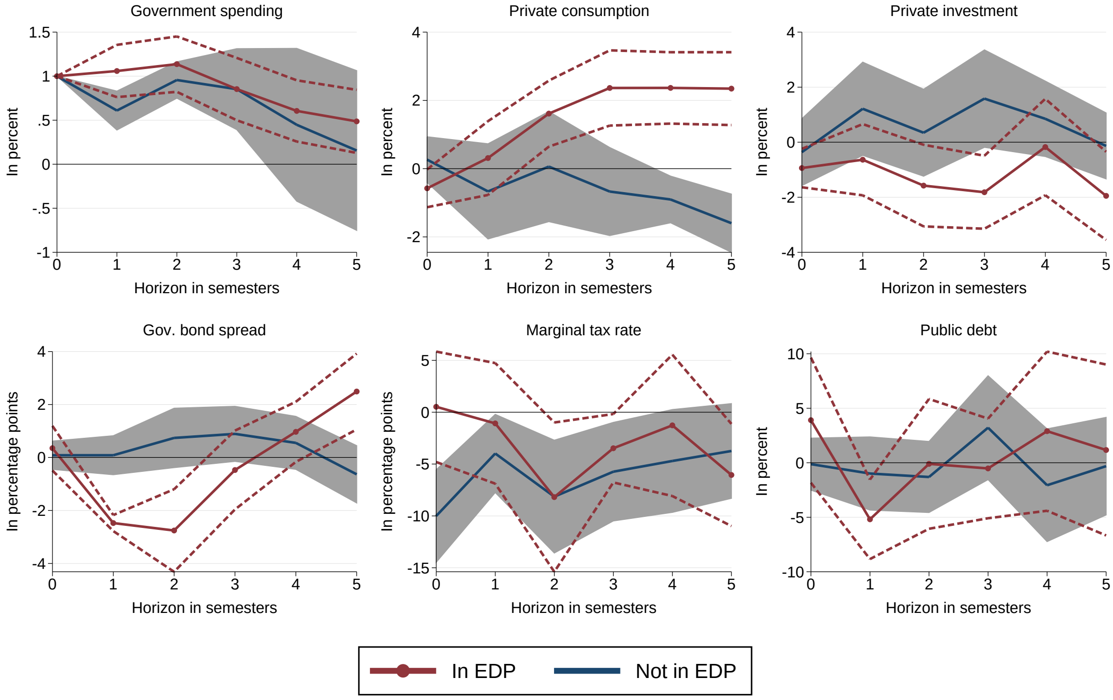

Figure B-3: IRFs in bad times – In recession

Notes: Estimates in each panel are based on 92 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

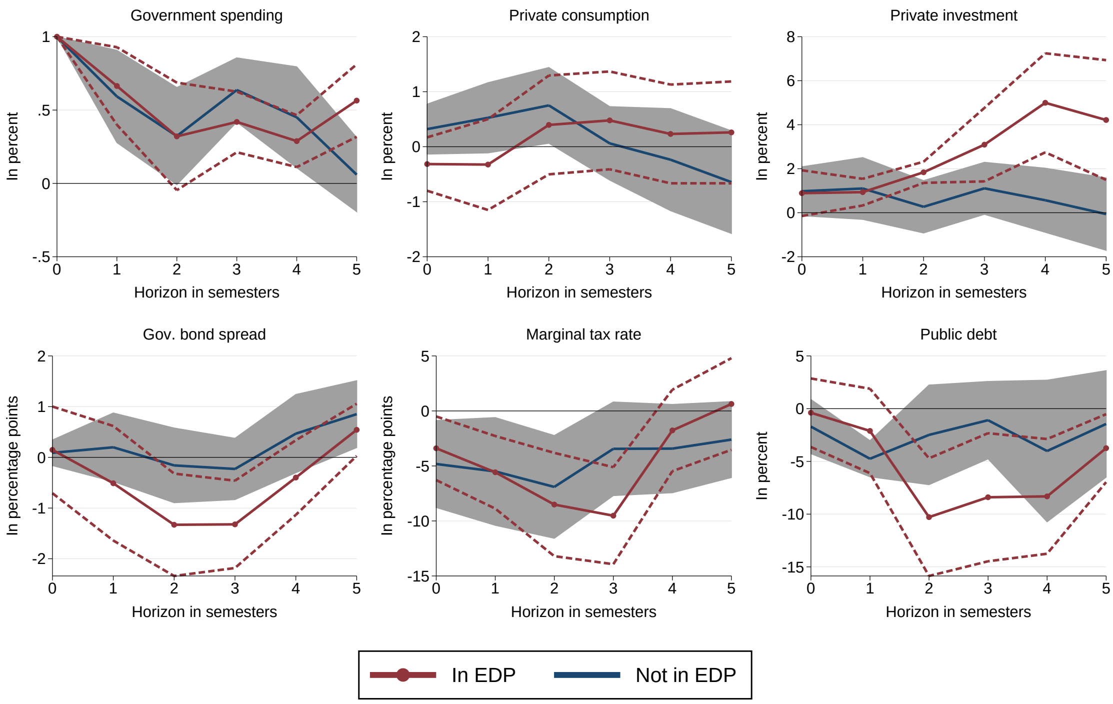

Figure B-4: IRFs in bad times – In ECB banking crisis

Notes: Estimates in each panel are based on 122 observations. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

C Identified Recessions

Table C-1: Identified recessions in our sample

#	Country	Start	Length	Depth	#	Country	Start	Length	Depth
1	AUT	2001q1	2	−0.21	20	FRA	2008q2	5	−3.87
2	AUT	2012q2	4	−1.10	21	FRA	2012q4	2	−0.12
3	BEL	2001q3	2	−0.29	22	GBR	2008q2	5	−6.04
4	BEL	2008q3	3	−3.69	23	IRL	2007q2	10	−10.87
5	BEL	2012q4	2	−0.32	24	IRL	2012q3	3	−1.80
6	CZE	2008q4	3	−5.86	25	ITA	2001q2	4	−0.65
7	DEU	2001q2	4	−0.85	26	ITA	2003q1	3	−0.57
8	DEU	2002q4	10	−0.47	27	ITA	2008q2	5	−7.46
9	DEU	2008q2	4	−7.03	28	ITA	2011q3	10	−5.33
10	DEU	2012q4	2	−0.88	29	LUX	2002q3	3	−2.40
11	DNK	2001q4	3	−0.24	30	LUX	2008q1	6	−8.00
12	DNK	2006q3	4	−1.05	31	LUX	2011q2	4	−1.57
13	DNK	2008q1	6	−7.07	32	NLD	2008q3	3	−4.35
14	DNK	2011q3	6	−0.50	33	NLD	2011q2	7	−1.97
15	ESP	2008q3	4	−4.36	34	PRT	2002q2	5	−2.43
16	ESP	2011q1	11	−5.28	35	PRT	2008q2	4	−4.33
17	FIN	2008q1	6	−9.49	36	PRT	2010q4	9	−7.87
18	FIN	2012q1	4	−2.49	37	SWE	2008q1	7	−5.86
19	FIN	2013q3	4	−0.96	38	SWE	2011q4	5	−1.59

Notes: "Start" refers to the first quarter of the recession, i.e., the quarter following the peak of the business cycle. "Length" states the duration of a recession in quarters. "Depth" refers to the deviation from the pre-recession peak level of output to the trough (in %).

D Detailed Results for Alternative States

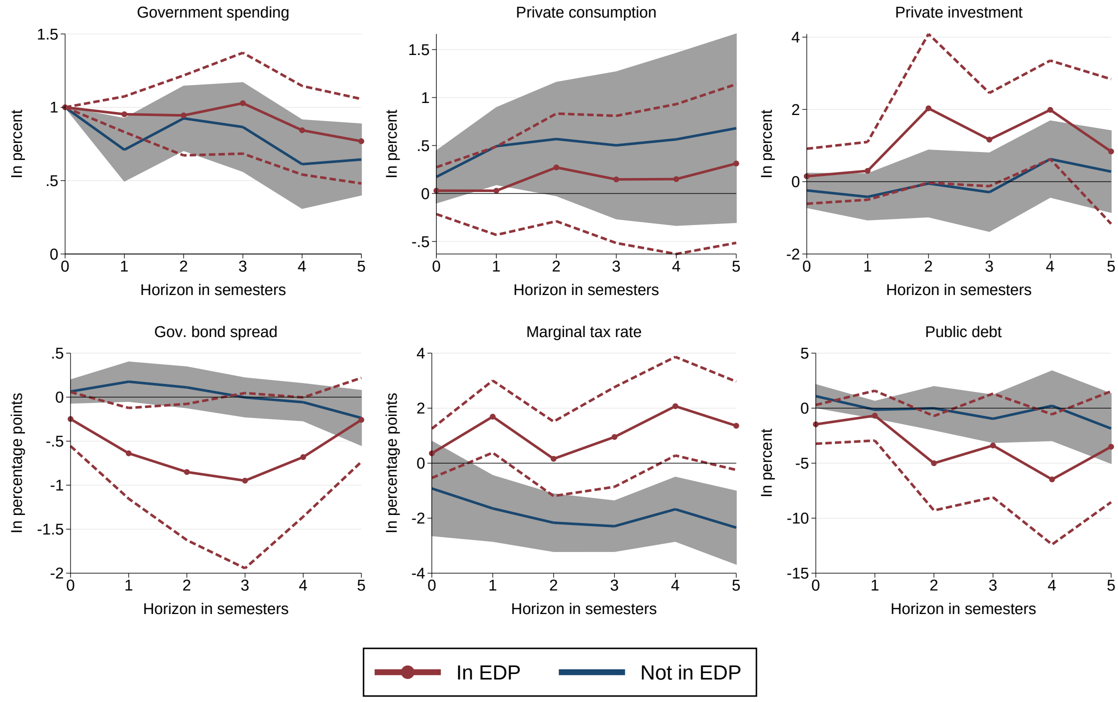

Figure D-1: IRFs for alternative states – Maastricht criteria not fulfilled (debt or deficit)

Notes: Estimates in each panel are based on the full sample. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

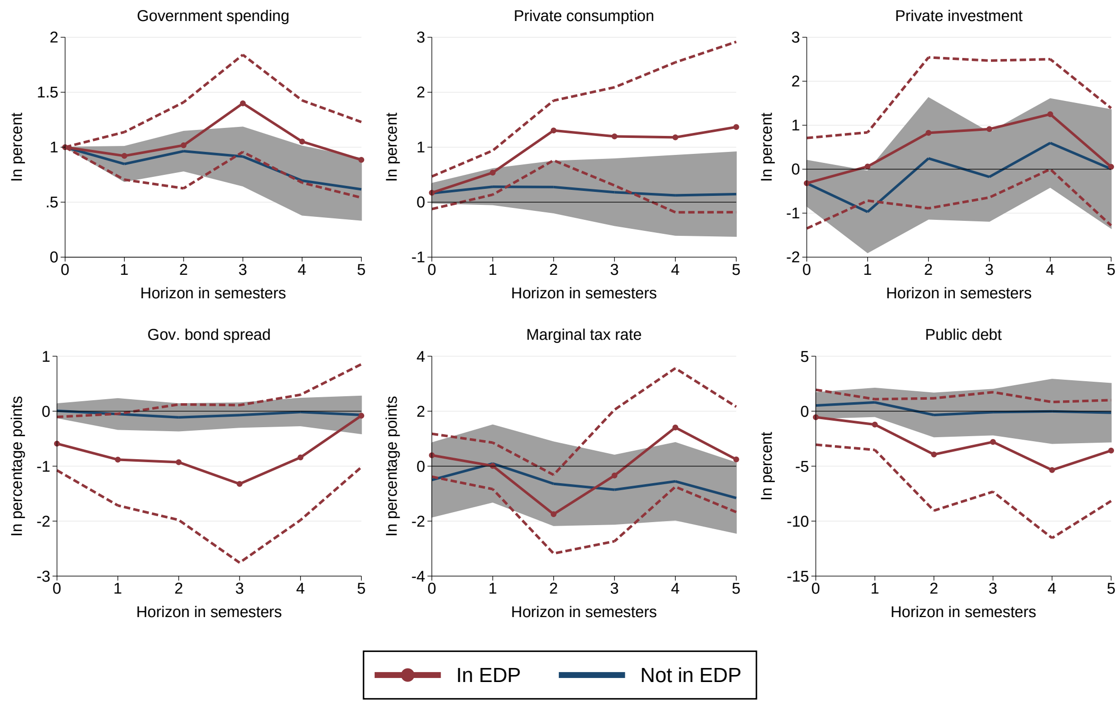

Figure D-2: IRFs for alternative states – Maastricht criteria not fulfilled (deficit only)

Notes: Estimates in each panel are based on the full sample. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

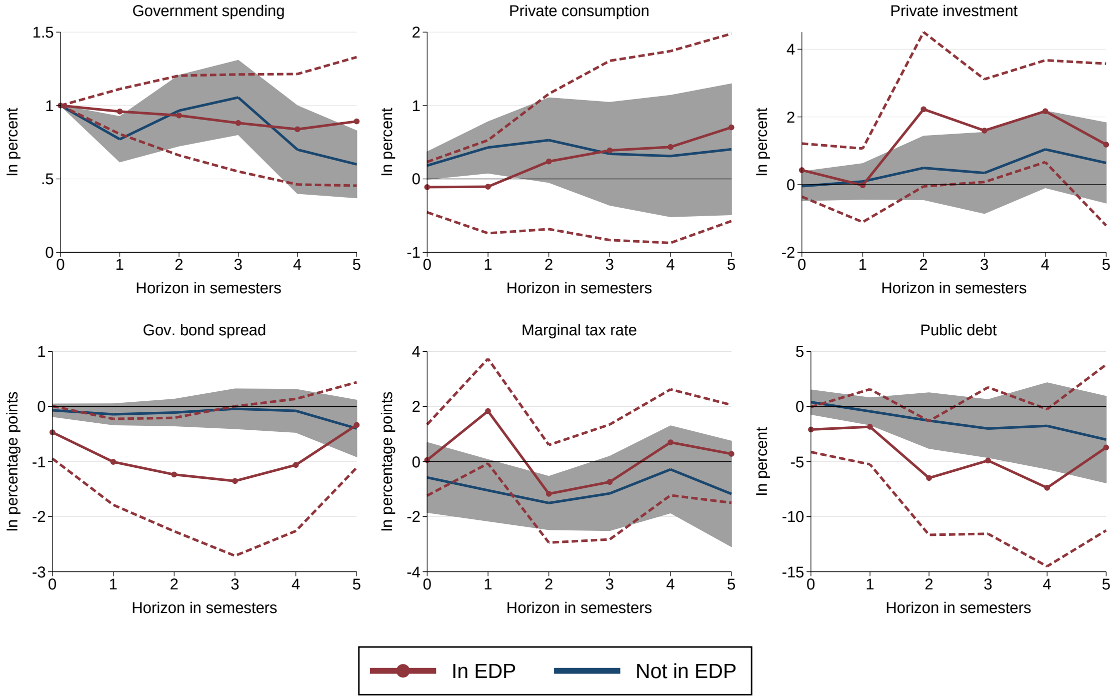

Figure D-3: IRFs for alternative states – Maastricht criteria not fulfilled (debt only)

Notes: Estimates in each panel are based on the full sample. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

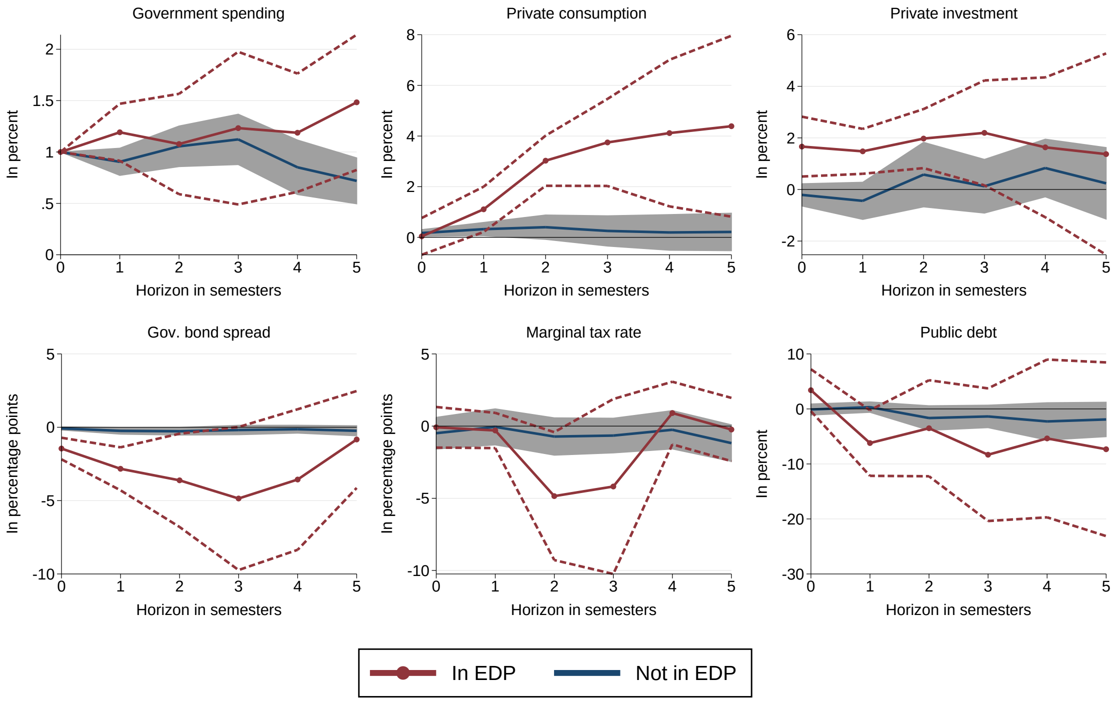

Figure D-4: IRFs for alternative states – Maastricht criteria not fulfilled (debt and deficit)

Notes: Estimates in each panel are based on the full sample. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

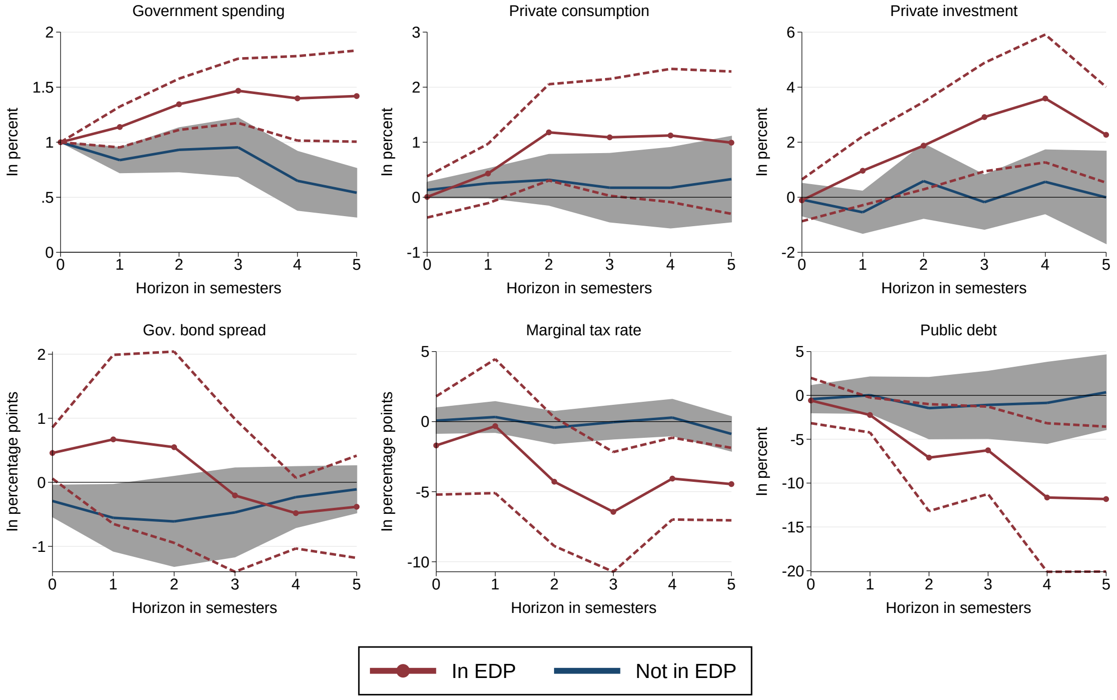

Figure D-5: IRFs for alternative states – Recession

Notes: Estimates in each panel are based on the full sample. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

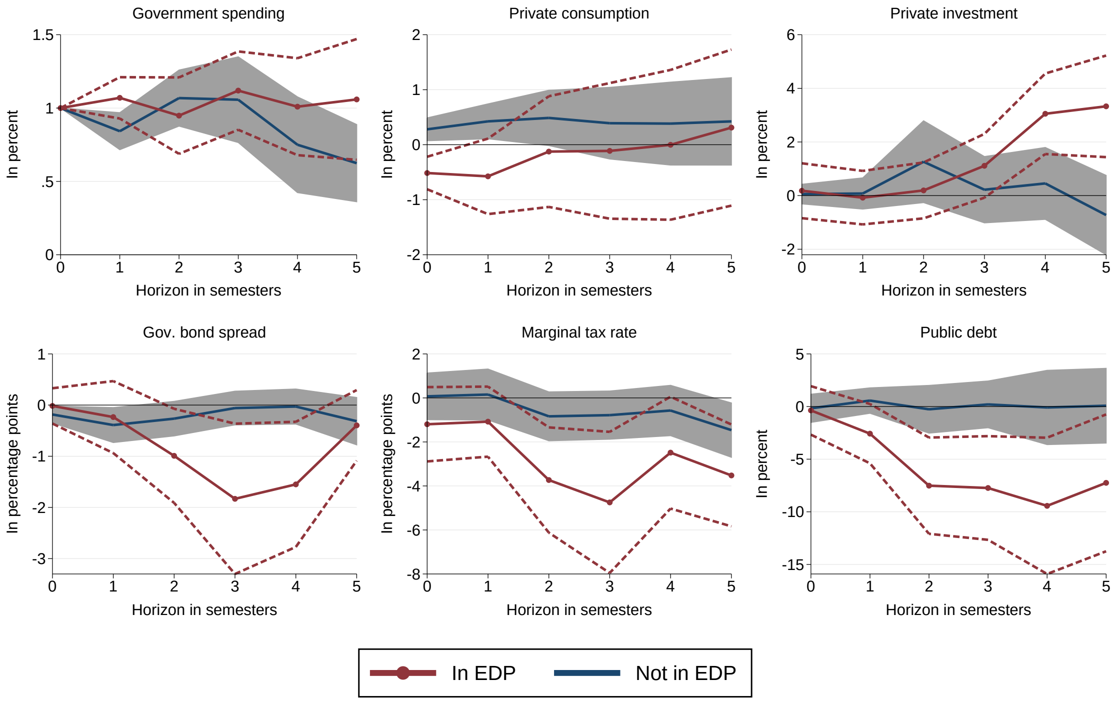

Figure D-6: IRFs for alternative states – ECB banking crisis

Notes: Estimates in each panel are based on the full sample. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

E Detailed Results for Real-time Multipliers

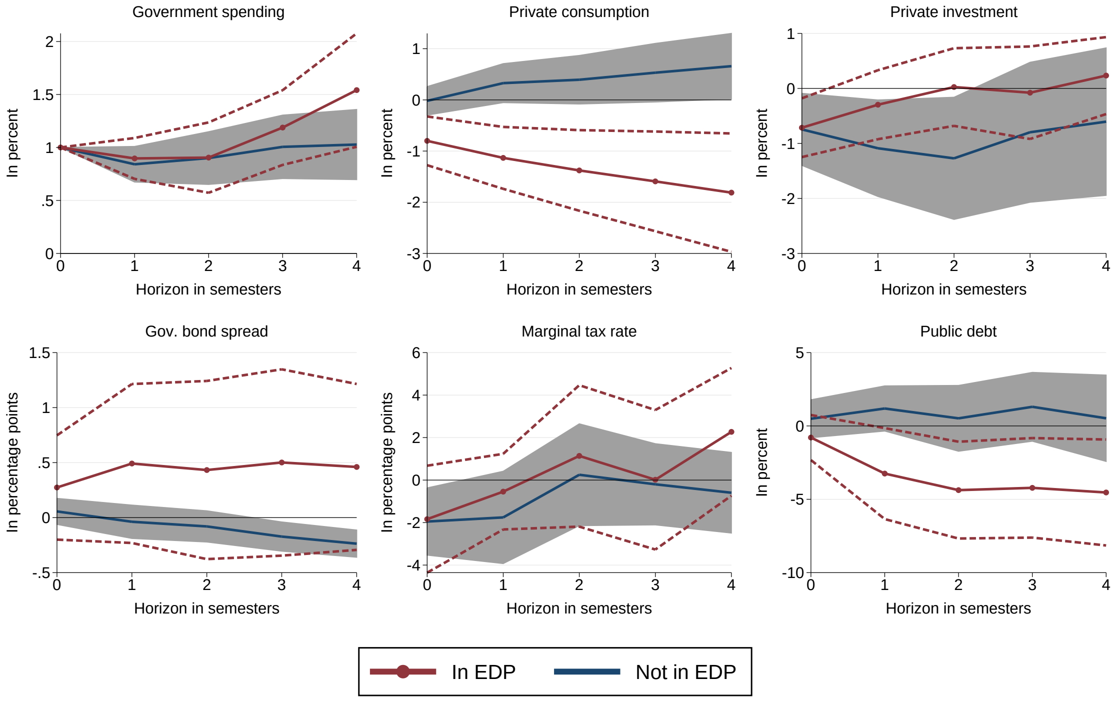

Figure E-1: IRFs for real-time exercise $s = t + 1$

Notes: Estimates in each panel are based on 373 observations covering 15 countries. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

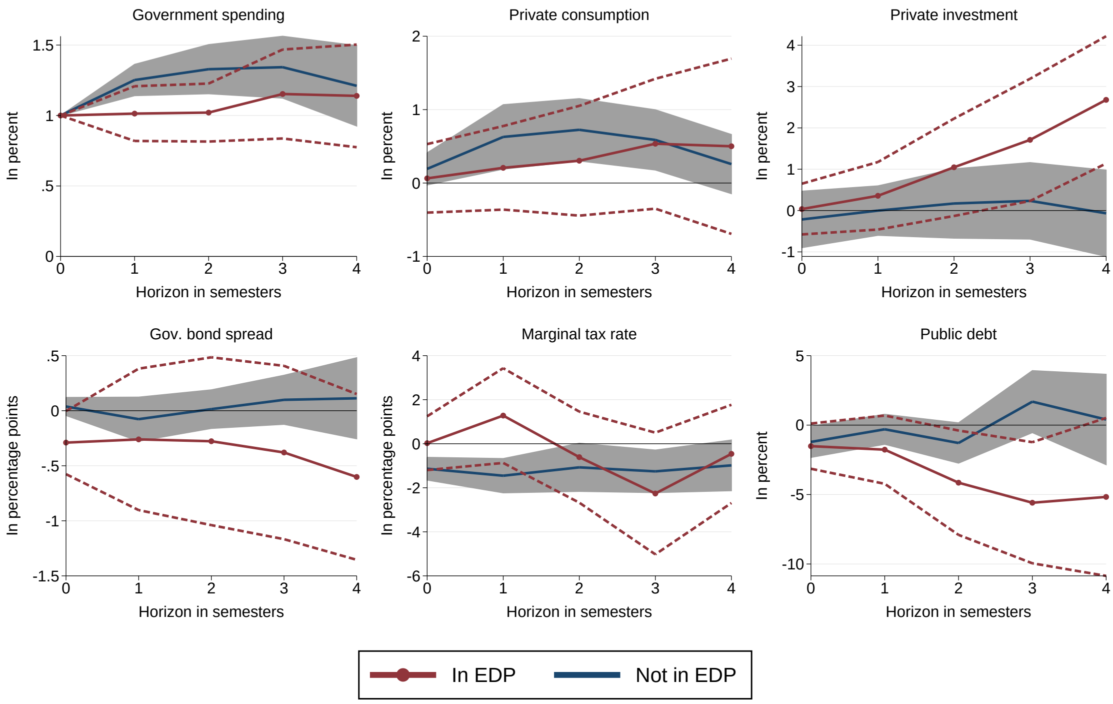

Figure E-2: IRFs for real-time exercise $s = t + 5$

Notes: Estimates in each panel are based on 373 observations covering 15 countries. 90% confidence intervals are calculated from Driscoll-Kraay standard errors.

References

Auerbach, Alan J. and Yuriy Gorodnichenko (2012). "Measuring the Output Responses to Fiscal Policy." *American Economic Journal: Economic Policy* 4 (2), 1–27.

— (2013). "Fiscal Multipliers in Recession and Expansion." In: *Fiscal Policy after the Financial Crisis*. Ed. by Alberto Alesina and Francesco Giavazzi. University of Chicago Press, 63–98.

Banerjee, Ryan and Fabrizio Zampolli (2019). "What drives the short-run costs of fiscal consolidation? Evidence from OECD countries." *Economic Modelling* 82, 420–436.

Barro, Robert J. and Charles J. Redlick (2011). "Macroeconomic Effects From Government Purchases and Taxes." *Quarterly Journal of Economics* 126 (1), 51–102.

Bernardini, Marco and Gert Peersman (2018). "Private debt overhang and the government spending multiplier: Evidence for the United States." *Journal of Applied Econometrics* 33 (4), 485–508.

Blanchard, Olivier J. and Daniel Leigh (2013). "Growth Forecast Errors and Fiscal Multipliers." *American Economic Review* 103 (3), 117–120.

— (2014). "Learning about Fiscal Multipliers from Growth Forecast Errors." *IMF Economic Review* 62 (2), 179–212.

Blanchard, Olivier J. and Roberto Perotti (2002). "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics* 117 (4), 1329–1368.

Boehm, Christoph E. (2020). "Government consumption and investment: Does the composition of purchases affect the multiplier?" *Journal of Monetary Economics* 115, 80–93.

Corsetti, Giancarlo, André Meier, and Gernot J. Müller (2012a). "Fiscal stimulus with spending reversals." *Review of Economics and Statistics* 94 (4), 878–895.

— (2012b). "What determines government spending multipliers?" *Economic Policy* 27 (72), 521–565.

De Jong, Jasper F. M. and Niels D. Gilbert (2020). "Fiscal discipline in EMU? Testing the effectiveness of the Excessive Deficit Procedure." *European Journal of Political Economy* 61, 101822.

Driscoll, John C. and Aart C. Kraay (1998). "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics* 80 (4), 549–560.

European Union (2008). *Treaty on the Functioning of the European Union*. URL: `https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A12008E126` (visited on 03/31/2021).

Gordon, Robert and Robert Krenn (2010). *The End of the Great Depression 1939-41: Policy Contributions and Fiscal Multipliers*. Working Paper 16380. Cambridge, MA: National Bureau of Economic Research.

Górnicka, Lucyna, Christophe Kamps, Gerrit Koester, and Nadine Leiner-Killinger (2020). "Learning about fiscal multipliers during the European sovereign debt crisis: evidence from a quasi-natural experiment." *Economic Policy* 35 (101), 5–40.

Hagen, Jürgen von and Barry Eichengreen (1996). "Federalism, Fiscal Restraints, and European Monetary Union." *The American Economic Review* 86 (2), 134–138.

Hall, Robert E. (2009). "By How Much Does GDP Rise If the Government Buys More Output?" *Brookings Papers on Economic Activity* (2), 183–231.

Hamilton, James D. (2018). "Why You Should Never Use the Hodrick-Prescott Filter." *The Review of Economics and Statistics* 100 (5), 831–843.

Harding, Don and Adrian Pagan (2002). "Dissecting the cycle: a methodological investigation." *Journal of Monetary Economics* 49 (2), 365–381.

Huidrom, Raju, M. Ayhan Kose, Jamus J. Lim, and Franziska L. Ohnsorge (2020). "Why do fiscal multipliers depend on fiscal Positions?" *Journal of Monetary Economics* 114, 109–125.

Ilzetzki, Ethan, Enrique G. Mendoza, and Carlos A. Végh (2013). "How big (small?) are fiscal multipliers?" *Journal of Monetary Economics* 60 (2), 239–254.

Jordà, Òscar (2005). "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1), 161–182.

Lo Duca, Marco, Anne Koban, Marisa Basten, Elias Bengtsson, Benjamin Klaus, Piotr Kusmierczyk, Jan Hannes Lang, Carsten Detken, and Tuomas A. Peltonen (2017). *A new database for financial crises in European countries.* Occasional Paper Series 194. European Central Bank.

Mineshima, Aiko, Marcos Poplawski-Ribeiro, and Anke Weber (2014). "Size of Fiscal Multipliers." In: *Post-crisis Fiscal Policy.* Ed. by Carlo Cottarelli, Philip Gerson, and Abdelhak Senhadji. MIT Press, 315–372.

Montiel Olea, José Luis and Carolin Pflueger (2013). "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics* 31 (3), 358–369.

Mountford, Andrew and Harald Uhlig (2009). "What are the effects of fiscal policy shocks?" *Journal of Applied Econometrics* 24 (6), 960–992.

Nickel, Christiane and Andreas Tudyka (2014). "Fiscal Stimulus in Times of High Debt: Reconsidering Multipliers and Twin Deficits." *Journal of Money, Credit and Banking* 46 (7), 1313–1344.

Orphanides, Athanasios (2003). "Monetary policy evaluation with noisy information." *Journal of Monetary Economics* 50 (3), 605–631.

Orphanides, Athanasios and Simon van Norden (2002). "The Unreliability of Output-Gap Estimates in Real Time." *Review of Economics and Statistics* 84 (4), 569–583.

Owyang, Michael T., Valerie A. Ramey, and Sarah Zubairy (2013). “Are Government Spending Multipliers Greater during Periods of Slack? Evidence from Twentieth-Century Historical Data.” *American Economic Review* 103 (3), 129–134.

Ramey, Valerie A. and Sarah Zubairy (2018). “Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data.” *Journal of Political Economy* 126 (2), 850–901.