INTERNATIONAL MONETARY FUND

Getting into the Nitty-Gritty of Fiscal Multipliers: Small Details, Big Impacts

José Federico Geli and Afonso S. Moura

WP/23/29

2023
FEB

WORKING PAPER

 WP/23/29

IMF Working Paper
Institute for Capacity Development

Getting into the Nitty-Gritty of Fiscal Multipliers: Small Details, Big Impacts

Prepared by José Federico Geli, Afonso S. Moura*

Authorized for distribution by Ali Alichi

Febraury 2023

ABSTRACT: The lack of a standardized framework to report fiscal multipliers limits comparisons across studies, budgetary items, or countries. Within a unified analytical framework (using a panel of 177 countries), we study how key methodological details affect the size and persistence of fiscal multipliers' estimates. Our baseline results are in line with the existing literature with average cumulative medium-term multipliers of -2.1 (-2.5) for taxes on personal income, 0.3 (1.7) for investment and, -0.5 (1.9) for consumption for advanced (emerging market) economies. However, we show that slight changes in the identification of shocks, based on forecast errors or in the definition of the fiscal multiplier, can artificially increase both the size and decrease the precision of estimates. We also emphasize the importance of accounting for the endogenous dynamic responses of fiscal variables to fiscal innovations by showing that multipliers calculated simply as the output response to fiscal shocks, as it is common in the literature, can potentially bias the results.

RECOMMENDED CITATION: Geli, José Federico and Afonso S. Moura, 2023, "Getting into the Nitty-Gritty of Fiscal Multipliers: Small Details, Big Impacts", IMF Working Papers 23/29.

JEL Classification Numbers:	E60, H30, H50
Keywords: Fiscal policy; Fiscal multipliers.	

Author's E-Mail Address:	José Federico Geli: jfgeli@imf.org; Afonso S. Moura: afonso.moura@novasbe.pt

* We are grateful to Ali Alichi, Marie-Pierre Aquino Acoste, Andrew Berg, Francesca Caselli, Antoine Cornevin, Nikolay Gueorguiev, Zeina Hasna, Koon Hui Tee, Ryota Nakatani, Martin Schlinder, Anna Ter-Martirosyan and participants at the ICD Departmental Seminar for very helpful comments and suggestions.

WORKING PAPERS

Getting into the Nitty-Gritty of Fiscal Multipliers: Small Details, Big Impacts

Prepared by José Federico Geli and Afonso S. Moura[1]

[1] We are grateful to Ali Alichi, Marie-Pierre Aquino Acoste, Andrew Berg, Francesca Caselli, Antoine Cornevin, Nikolay Gueorguiev, Zeina Hasna, Koon Hui Tee, Ryota Nakatani, Martin Schlinder, Anna Ter-Martirosyan and participants at the ICD Departmental Seminar for very helpful comments and suggestions.

Contents

1 Introduction **6**

2 Data **9**

3 Analytical framework **9**

3.1 Baseline specification 9

3.2 Identification of fiscal shocks 11

4 Results **13**

4.1 Baseline results 13

4.2 Alternative cleaning procedures of forecast errors 18

4.3 One-stage vs Two-stage Estimation 20

4.4 Different approach to accumulate responses 28

4.5 Alternative data sources 30

5 Conclusion and Policy Implications **31**

6 Appendix **38**

6.1 Descriptive Data 38

6.2 Summary of Main Estimations 45

6.3 Different Specifications 45

6.4 Uncertainty surrounding Forecast Errors 47

6.5 Instrument Relevance 49

List of Figures

1 Multiplier heterogeneity across different groups of countries. 16

2 Multiplier heterogeneity across fiscal tools. 18

3 Multipliers using different ways of cleaning forecast errors. 20

4 One vs Two-stage estimation. 23
5 Dynamic response of fiscal variable to the fiscal shock: Local projections of first stage regressions. 24
6 Cumulative response of government budget balance to public investment and public consumption shocks. 26
7 SNA vs GFM data. 31
8 Histogram of fiscal shocks for all sample, including AEs, EMEs and LICs. . . 38
9 Histogram of fiscal shocks for AEs. 39
10 Histogram of fiscal shocks for EMEs 39
11 Histogram of fiscal shocks for LICs. 40
12 GDP growth and Fiscal Shocks. 41

List of Tables

1 Multipliers when using FE of growth rates. 28
2 Different approaches to calculate cumulative multipliers. 29
3 Income classification of countries. 42
4 Descriptive Statistics. 43
5 Sample size: number of countries per year. 44
6 Summary of Results. 45
7 Local projections including Covid years in the estimation. 46
8 Different fixed-effect specifications. 46
9 Different lag specifications. 47
10 Different clustering levels of Standard Errors. 47
11 Estimated vs not estimated fiscal shock. 48
12 F-statistic of first stage regression for instrument relevance. 49

1 Introduction

In the past years, especially after the financial crisis, there has been a remarkable interest in studying fiscal multipliers. This has led to considerable innovations in both the empirical methodologies as well as in identifying theoretical determinants of their size[1]. One of the most important takeaways from this bulk of new research is how much multipliers are context-dependent. They can vary given the amount of slack in the economy (Auerbach and Gorodnichenko, 2012, Nakamura and Steinsson (2014), Hernández de Cos and Moral-Benito (2016)); the stance of monetary policy (Miyamoto, Nguyen and Sergeyev, 2018); the direction of the fiscal intervention (Barnichon, Debortoli and Matthes, 2022); the financing source (Kraay, 2012), among many other factors.

However, estimates also diverge due to methodological differences, limiting comparison across studies. Results can vary due to the econometric approach, the identification strategy[2] and the multiplier definition. For this reason, comparing estimates across studies is not a straightforward task. Without a standardized framework, this comparison exercise may lead to wrong conclusions and ill-informed policy advice. This raises the question of what methods should policymakers rely on to assess the impact of fiscal decisions, a question that is specially relevant for Emerging Market economies (EMs) and Low Income Countries (LICs) for which the literature on the topic is still relatively scarce.

In this paper, we address this sources of heterogeneity among estimates in a consistent, uniform, and integrated framework, digging into the *nitty-gritty* of multipliers estimation and reporting. Although our baseline results are in line with the range of estimates Ramey (2019) reports, we show how some subtle methodological details can have a significant impact on the results. We do so by empirically[3] estimating multipliers across different horizons, country

[1]See, for example, Ramey (2019) for a survey on the topic.

[2]Caldara and Kamps (2017) derive an uniform framework and compare how different commonly used identification schemes implicitly or explicitly determine the size of the estimated fiscal multiplier.

[3]We focus on empirical methodologies instead of multipliers calculated from estimated and calibrated dynamic stochastic general equilibrium (DSGE) models. For a survey on the differences in the estimated multipliers across the two different approaches, see Ramey (2019).

groups and budgetary items, using the Jordà (2005) local projection methodology. We follow Ramey and Zubairy (2018) and estimate the local projections within a two-stage instrument variable framework using exogenously identified fiscal shocks as instrumental variables, and emphasize some of the most important advantages of this methodology.

We identify fiscal shocks by calculating forecast errors (FE) of fiscal variables using IMF's *World Economic Outlook* data.[4] Our analysis shows that obtaining valid and relevant exogenous fiscal shocks using forecast errors requires a careful filtering procedure. Otherwise, apparently good candidates for instruments could bias the results, usually increasing the estimated multiplier.

Another dimension we take into account is how fiscal variables endogenously respond to fiscal shocks throughout the horizon and how this impacts multipliers. We do this by comparing the results of the two-stage instrument variable (IV) approach with a (commonly used in the literature) one-step estimation that calculates the impact of exogenous shocks directly on output, without using them as IVs. Our results imply that neglecting this *dynamic* effect that shocks have on fiscal variables leads to incorrect fiscal multiplier estimates, both in terms of magnitude and persistence. In addition, we highlight how following the two-stage methodology can attenuate differences in results from using forecast errors calculated under alternative approaches. Finally, we also show how the estimation framework can easily be extended to assess the impact of the shocks on the governments financing needs.

Different data sources and reporting methodologies can also be a source of differences across estimates. Most expenditure-based estimates of fiscal multipliers use as input government expenditure data from the System of National Accounts (SNA). However, fiscal policy measures are usually designed using Government Finance Statistics (GFS) data. In many countries, in particular in EMs and LICs, one of the main differences between these two types of data is that SNA data is an accrued estimate of the cash-based fiscal accounts. Our results signal that using these two sources of data can also lead to substantial differences across re-

[4]We follow this identification approach throughout the paper because we aim to compare results in a uniform framework.

sults. Furthermore, different approaches to accumulate responses yield different multipliers estimates, even if the data source, identification strategy and methodological approach are the same.

An additional contribution to the literature is that we focus our analysis mainly on EMs and compare estimates of multipliers for both expenditure and revenue categories. Due to data availability or statistically quality reasons, most studies on fiscal multipliers have focused on advanced economies (AEs). However, studies that do focus on EMs can reach different conclusions not because of methodological differences, but due to different sample size and length considered[5]. A uniform framework is again important to overcome this comparison limitation and extend our understanding of multipliers in EMs and LICs. In addition, most authors who rely on forecast errors to identify fiscal shocks focus on expenditure multipliers[6]. To the best of our knowledge, we are the first ones to extend this identification approach to the revenue side. By reporting fiscal multiplier estimates within the same analytical framework, we can better single out differences among public investment, public consumption and personal income tax multipliers.

Finally, by relying on a uniform framework, our estimates can also be used to enhance tools such as the bucket approach (Batini, Eyraud and Weber, 2014), which is based on estimates from studies using different methodologies.

The rest of the paper is organized as follows. Section 2 describes the data used in the study. Section 3 explains the analytical framework. We show the baseline econometric specification and carefully explain the process to identify exogenous fiscal shocks. Section 4 explores the results. We show the baseline results and all the important changes that lead to

[5]A few examples on how much the sample size can vary are: Ilzetzki, Mendoza and Végh (2013) who use a panel of 44 countries, of which only 24 are developing economies; Carrière-Swallow, David and Leigh (2021) who present evidence for 14 Latin American and Caribbean countries; Restrepo (2020) who focus only on Latin America economies; and Honda, Miyamoto and Taniguchi (2020) who use data for 42 LICs which are not resource rich economies

[6]They do so either by looking at government consumption or government expenditure. Moreover, even though the evidence shows that multipliers for public consumption and investment can differ, a usual practice in the literature is to sum their shocks as a measure of total government expenditure innovations (IMF, 2018, Colombo et al., 2022).

significant differences in the estimates. We explore either different specifications compared to the baseline results, or use different data (either by slightly changing the fiscal shocks used as instruments or the source of the data being used). Section 5 concludes and analyzes the policy implications of the paper.

2 Data

We use an unbalanced panel of countries with annual data taken from the IMF's World Economic Outlook (WEO) database archives. For public consumption and public investment multipliers, the exercises include data spanning from 1995 until 2019. Personal income taxes series, however, is shorter. Because we need real-time vintage data when computing the fiscal forecast errors, for this variable available data starts only in 2010. All variables are expressed in real terms, deflated by the GDP deflator. Since forecast errors of fiscal variables or variations in output can be particularly large for the Covid-19 years, only data until 2019 was included in the exercises.[7] The baseline exercises include 38 AEs, 85 EMs, and 54 LICs. See appendix 6.1 for descriptive statistics of the data used.

3 Analytical framework

3.1 Baseline specification

To calculate[8] multipliers over a certain horizon after the fiscal shock takes place, we use the Jordà (2005) local projection methodology[9]. We specify the regressions such that the impulse response functions that the method yields can be directly interpreted as cumulative multipliers for each horizon h. As in Auerbach and Gorodnichenko (2013), we apply this method in a context of a panel data framework. Additionally, we follow Ramey and Zubairy (2018) and extend the local projections to an instrumental variable framework.

[7] Appendix 6.3 shows how results change when including the year of 2020 and 2021 in the sample.

[8] All codes were written using R and can be made available upon request.

[9] Local projection methods do not impose any dynamic restrictions to the model, as opposed to a VAR structure. Additionally, the local projection model can be more parsimonious than the VAR specification.

For the baseline model, we estimate the following regressions for each horizon *h=0,1,...,H*:

$$y_{i,t+h} = \alpha_{i,h} + \mu_{t,h} + \beta_h \hat{f}_{i,t+h} + \theta_h X_{i,t-1} + \varepsilon_{i,t+h} \tag{1}$$

$$f_{i,t+h} = \chi_{i,h} + \xi_{t,h} + \rho_h shock^f_{i,t} + \omega_h X_{i,t-1} + u_{i,t+h}, \tag{2}$$

where equation 1 is the second stage regression and equation 2 is the first stage regression in which an exogenous fiscal $shock^f$ variable is used as an instrument for the respective fiscal variable. We focus on three different types of fiscal tools, so f can be either government consumption, government investment or personal income taxes. Both regressions include country and time fixed effects. In the baseline specification, $X_{i,t-1}$ includes one lag of $y_{i,t}$, $f_{i,t}$ and $shock_{i,t}$[10]. The dependent variable $y_{i,t+h}$ is the cumulative change in GDP from t to $t+h$ and $f_{i,t+h}$ is the cumulative change in the fiscal variable of interest from t to $t+h$. We follow Canova and Pappa (2021) (henceforth, C&P) and express the cumulative variables, in order to calculate cumulative multipliers at each horizon h, as $y_{i,t+h} = \sum_{1}^{h} \frac{Y_{i,t+h-1} - h * Y_{i,t-1}}{Y_{i,t-1}}$ and $f_{i,t+h} = \sum_{1}^{h} \frac{F_{i,t+h-1} - h * F_{i,t-1}}{Y_{i,t-1}}$. All variables are normalized by the same period GDP (at $t-1$), and therefore can be interpreted as unit multipliers directly, instead of elasticities[11].

We define cumulative multipliers as in Ramey and Zubairy (2018), which slightly changes the Mountford and Uhlig (2009) definition of present value multipliers. That is, we acknowledge that in a dynamic environment, multipliers should be reported as the *integral of the output response divided by the integral government spending response after a fiscal shock taking place*[12]. Because of the IV approach and the way we accumulate both output and the

[10]Appendix 6.3 shows results using different econometric specifications. Especially for the case of the multiplier for personal income taxation, since the number of the years included in the exercise is smaller, the use of a within estimator when the model has a dynamic component (the lagged term $y_{i,t}$) can introduce some bias in the estimations. In the appendix, we report the results (for EMEs) when we drop the dynamic component of the model and results do not significantly change.

[11]See Ramey and Zubairy (2018) for a detailed explanation on the drawbacks of estimating elasticities and then convert the estimates into unit multipliers.

[12]Mountford and Uhlig (2009) calculate the present value multiplier, discounted by the average interest rate of their sample. However, as our estimation is done in a panel data framework, discounting the multipliers with the average of the interest rates across country and across time would not be a good exercise, given the

fiscal variable across each horizon, the estimates directly give dynamic cumulative multipliers at each horizon h. In section 4.3 we explore how results change when this *dynamic environment* is neglected.

3.2 Identification of fiscal shocks

The literature mostly relies on three different approaches to identify exogenous fiscal shocks: the narrative approach (Romer and Romer, 2010), recursive approach (Blanchard and Perotti, 2002) or the use of forecast errors (Auerbach and Gorodnichenko, 2013). As Colombo et al., 2022 points out, the economic or institutional heterogeneity that may be present in a large panel of countries can make the identification of shocks difficult using the first two approaches due to their information requirements. For EMs and LICs, in which publicly available official documents are scarce, the narrative identification becomes unfeasible. Additionally, the lack of macroeconomic statistics at a quarterly frequency also does not allow for a time restriction identification of fiscal shocks. To overcome this constraint, we construct forecast errors of fiscal variables and use it to derive a measure of fiscal shocks.

Forecast errors for a given variable are the difference between the projected and the realized value. By construction, they should be unexpected shocks to the economy[13]. In theory, forecast errors avoid the problem of *fiscal foresight*: if a shock happens at t but was previously announced at *t-1*, an econometrician might not be able to find any significant economic response because economic agents may have reacted before the shock actually took place. By capturing only purely unanticipated changes, forecast errors reduce this *fiscal foresight* bias.

We construct fiscal forecast errors using World Economic Outlook real-time vintage data, in the following way:

$$FE_{i,t}^{f} = f_{i,t} - f_{i,t|t-1}, \tag{3}$$

heterogeneity among the all the countries in our panel.

[13] As An et al. (2018) demonstrate, WEO fiscal projections are relatively precise and unbiased in comparison with other private forecasters.

where f is the fiscal variable of interest, $f_{i,t}$ represents its realized value, for a given country and period of time, and $f_{i,t|t-1}$ is the forecast for that same fiscal variable[14]. Both $f_{i,t}$ and $f_{i,t|t-1}$ are expressed as a share of its corresponding vintage contemporaneous GDP[15]. That is, $f_{i,t} = \frac{F_{i,t}}{Y_{i,t}}$ and $f_{i,t|t-1} = \frac{F_{i,t|t-1}}{Y_{i,t|t-1}}$.

However, even if forecast errors are unexpected shocks in theory, in practice they might not be completely orthogonal to past macroeconomic trends. If that would be the case, by exploiting information on past economic developments, agents could, at least partially, adjust their actions before the shock takes place. To avoid this effect, we regress the fiscal forecast errors on a set of lagged macroeconomic variables, such as GDP growth rate, total revenue and total expenditure as a percentage of GDP, exchange rate growth and inflation[16].

Another potential issue related to the use of fiscal forecast errors is that fiscal forecasts are usually based on a set of macroeconomic assumptions. If these assumptions change, even if no new fiscal measures are implemented, fiscal forecasts are probably going to diverge from final outcomes. For example, a lower-than-expected VAT collection can be simply the result of a lower-than-expected private consumption expenditure and not due to any additional tax reduction or exemption measure. Hence, it is important to guarantee that identified fiscal shocks are not correlated with contemporaneous economic conditions or other types of shocks affecting the economy. To tackle this issue, we also regress fiscal forecast errors on *contemporaneous* forecast errors of GDP growth, exchange rates growth and inflation (as measured by the percent change of the output deflator). A similar approach is followed by Abiad, Furceri and Topalova (2016), IMF (2017) and IMF (2018), for example, that either regress fiscal forecast errors on output and inflation forecast errors or test if regressing on other forecast errors (such as private consumption or investment) influences the results.

[14] $f_{i,t}$ is taken from the October WEO released at $t+1$, which contains the realized variable at t, and $f_{i,t|t-1}$ is taken from the October WEO released at t.

[15] This also reduces the risk of artificially having a high forecast error because statistical changes occurred between WEO vintages, for example, because of a currency re-denomination. If this is the case, then normalizing the level of each fiscal variable by its corresponding GDP using the same vintage should overcome this issue.

[16] See, for example, Colombo et al. (2022) and Cacciatore et al. (2021) for similar ways to filter out the *expected* component of forecast errors.

Additionally, we include in the regression forecast errors of exchange rate movements to (partially) capture the presence of external shocks.

Before regressing each fiscal forecast errors on all of the above mentioned variables (in a panel data framework, including also time and country fixed effects), we first trim the sample, ruling out 1% of each tail of the distribution to remove (most likely) data outliers[17]. Additionally, we drop from our sample any forecast error that is exactly equal to zero. The residuals of this regression are used as the fiscal shocks[18].

The former filtering process warrants that identified shocks respect the properties Ramey (2016) points out as being necessary for a shock to be considered as such: uncorrelated with past conditions of the economy, with other shocks and, by construction, unanticipated. Appendix 6.1 plots the histograms of the shocks for the different tools by income groups. On section 4.2 we explore the impact of using alternative definitions of forecast errors.

4 Results

4.1 Baseline results

The first source of heterogeneity in estimates that we want to explore is the difference in output responses by country groups (AEs, EMs and LICs). In theory, estimated multipliers across different country groups can be heterogeneous. For example, they can vary due to differences in the exchange rate regime (Ilzetzki, Mendoza and Végh, 2013), on the quality of the institutions (Honda, Miyamoto and Taniguchi, 2020), the degree of informality (Colombo et al., 2022), among many other reasons. Figure 1 plots the cumulative multipliers, from impact until 2 periods-ahead of the fiscal shock taking place, for different country groups. We start by showing the impulse response functions (IRFs) when we use all the countries

[17]If outliers of forecast errors are not excluded before regressing them on current and past macroeconomic variables, they would bias the coefficients of such regression and the residuals would became artificially big.

[18]There is, however, one drawback of using the estimated residuals of this regression as our fiscal shocks. Because these variables will be estimates, instead of observed variables, there is some uncertainty associated to them that can impact the final estimates of fiscal multipliers, specially the standard errors of the estimated final coefficients. This concern has been overlooked, however, in the literature that follows a similar procedure. In the appendix we explore this problem and show the results of one specification that tries to overcome this problem.

in the sample. Then, we show the same IRFs by income group and focus on two different fiscal categories: personal income taxes on the revenue side and public investment on the expenditure side.

For AEs and EMs economies, although our estimates for personal income tax multipliers are relatively high, they are in line with what the empirical literature reports, with estimates usually around -2 and -3 (Ramey, 2019, Restrepo, 2020). In a similar fashion, public investment multipliers for these two country groups are also in accordance with the literature: impact multipliers bellow unit but higher afterwards. However, there is one important difference between AEs and EMs. Public investment multipliers are close to zero after two years for the first group, whereas the cumulative response of output is still increasing after two years in EMs. This suggests that, although the short-run multiplier for public investment in these two groups may be similar, the longer-term effects may differ considerably[19]. As Izquierdo et al. (2019) point out, the reason for this may hinge on the fact that countries with an initial higher stock of capital experience lower marginal productivity of additional units of capital, and may experience no crowding-in of private investment. That may explain why the positive impact in output from public investment in AEs is short-lived. Boehm (2020) finds similar results, estimating that, for OECD economies, government investment multipliers are close to zero while government consumption ones are close to one two years after the shock.

Unlike the previous cases, the results for LICs are considerably different both in terms of sign and precision. Estimates for personal income tax and public investment have an associated multiplier with the opposite sign one would expect. Still, the relatively high standard errors imply that estimates are not statistically different from zero for most of the horizons considered. In the case of public investment, a weaker institutional framework might signal a less efficient implementation of public projects (IMF, 2015) leading to lower (or even null) multipliers[20]. However, in the case of the personal income tax, the results are

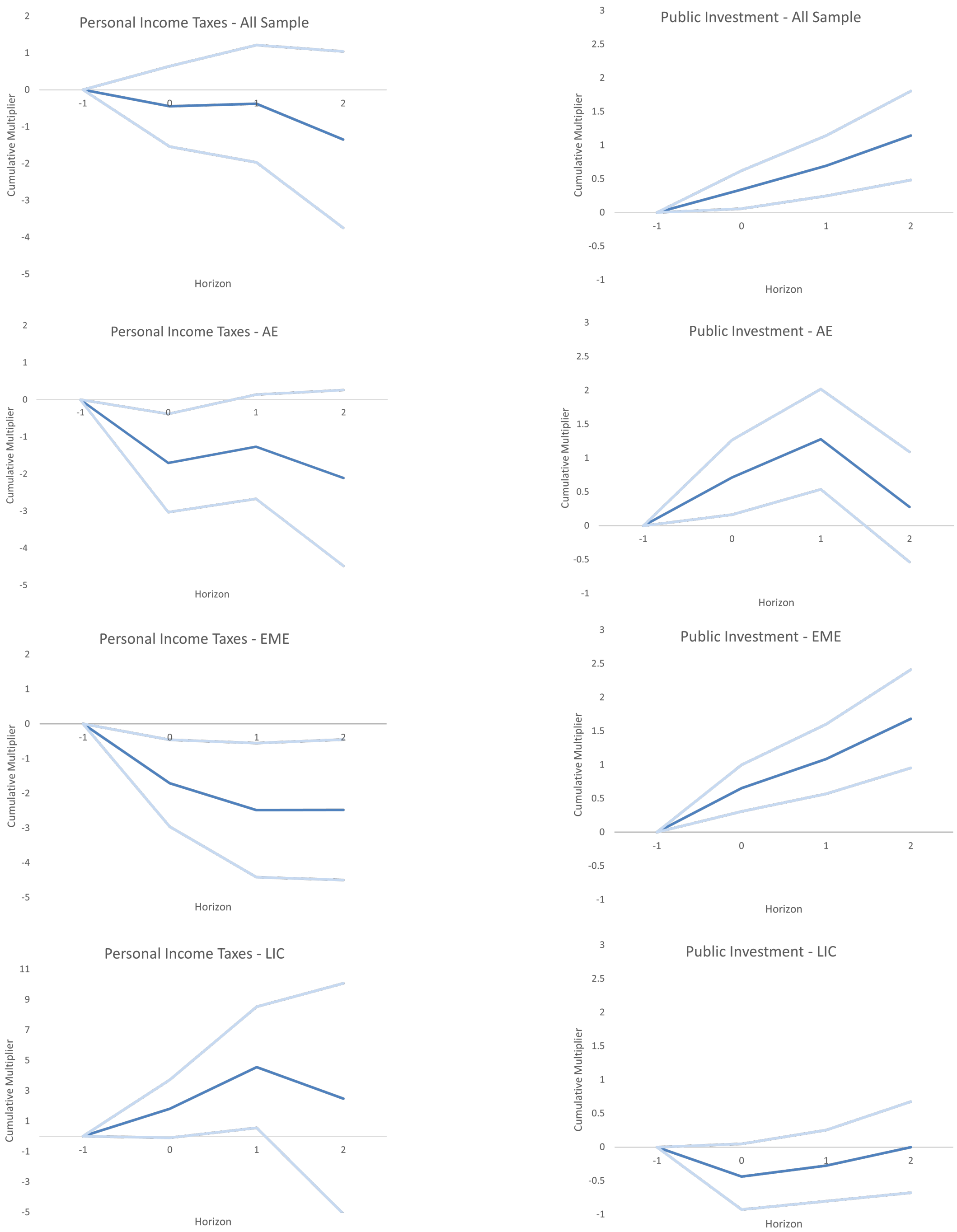
[19] See Raga (2022) for a review on short and long-term fiscal multiplier estimates across studies.

[20] See also Acemoglu, Johnson and Robinson (2001); Ederveen and Nahuis (2006); Rodríguez-Pose and

somehow puzzling. Since the same methodology is being followed for all the country groups, this results can be indicative of data quality issues, not properly reflecting the underlying economic developments, which bias the estimates.

Garcilazo (2015) or Avellán, Andrade and León-Díaz (2020) for a role of institutions in the size of expenditure multipliers.

Figure 1: Multiplier heterogeneity across different groups of countries.



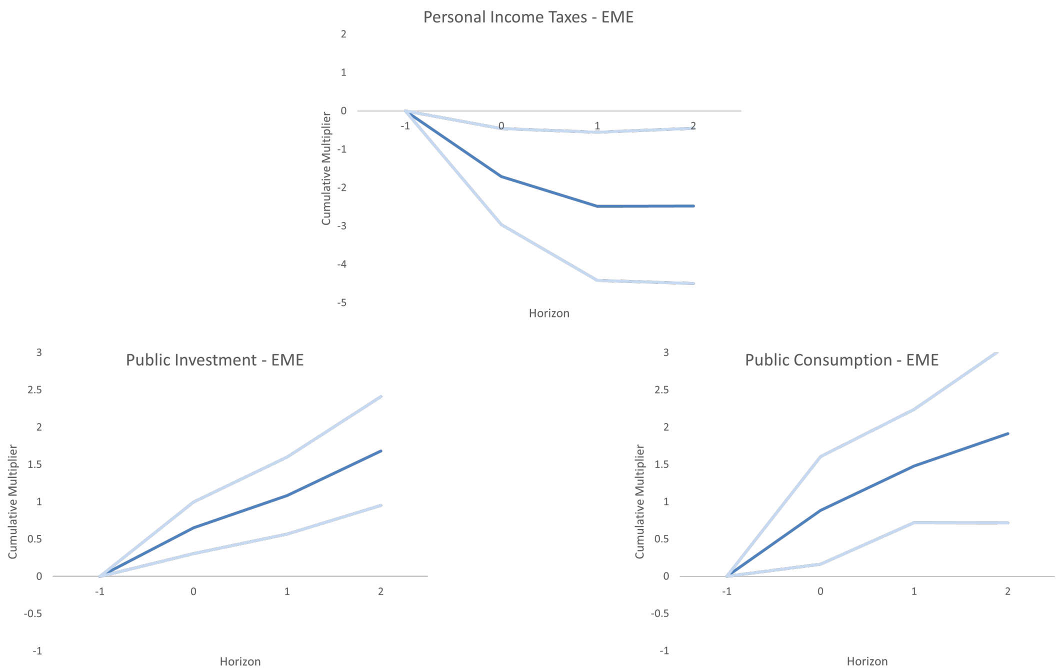
Note: C&P multipliers. One-standard deviation confidence bands using robust standard errors, clustered at the country level. Shocks to personal income taxes and public investment are expressed as the cumulative response of GDP to a unit increase in the respective cumulative fiscal variable, at a given horizon. Number of observations included in each estimation (for the $h = 0$), personal income taxes: all sample n=626; AEs n=111; EMs n=341; LICs n=174; public investment: all sample n=2202; AEs n=342; EMs n=1125; LICs n=721.

After analyzing the differences among country groups, we further explore the differences in multipliers when focusing on the impact of shocks in different budgetary items[21]. Figure 2 plots cumulative multipliers for the three variables analyzed, using the EMs sample. Besides differences on size (i.e larger multipliers, in absolute terms, for personal income tax), the main difference across results seems to be the persistence of responses. While personal income tax shocks have a considerable contemporaneous impact which becomes virtually zero after two years (i.e.: the cumulative effect remains stable), the marginal impact of both government consumption and investment shocks remains positive along the estimated horizon. In addition, the estimates for public investment multipliers and government investment multipliers are very similar at the horizons considered, although confidence intervals for the first are wider. A technical explanation for these differences lies in the endogenous response our shocks induce in the fiscal variables, as we explore in section 4.3. In that section we show that the output response to investment shocks is stronger than the response to public consumption shocks. However, our public investment shocks are more persistent - that is, they have a higher impact on public investment also in the years following the shock[22]. Therefore, by using the integral multiplier definition that we follow, estimates for the two types of expenditure become similar.

[21]For the remaining exercises of the paper we will focus only on EMs for the following reasons: 1) to ensure that results are comparable across different exercises, 2) because it is a less studied income group relatively to AEs and 3) due to the statistical quality concerns around LICs data that might bias the estimates, as figure 1 shows.

[22]One reason why public investment can show persistence years after a shock is implemented is due to the multi-annual nature of many investment projects. In addition, in many cases public contractors fail to comply with original deadlines and exceed the pre-agreed budget ceilings.

Figure 2: Multiplier heterogeneity across fiscal tools.



Note: C&P multipliers. One-standard deviation confidence bands using robust standard errors, clustered at the country level. Shocks to personal income taxes, public investment and public consumption are expressed as the cumulative response of GDP to a unit increase in the respective cumulative fiscal variable, at a given horizon.

4.2 Alternative cleaning procedures of forecast errors

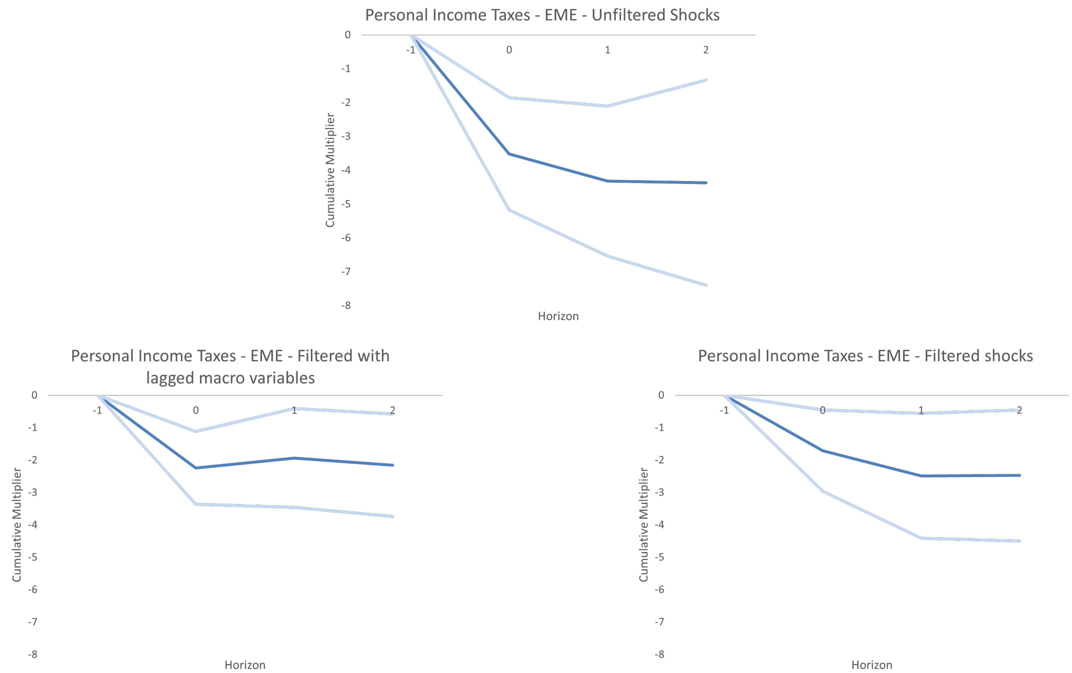
The identification of fiscal shocks using forecast errors, apart from capturing unexpected changes in the fiscal variables, also relies on the assumption that there is a lag associated to the implementation of fiscal policy measures. In other words, it relies on the quarterly time restriction (Blanchard and Perotti, 2002). That is because, presumably, when using the forecast of a fiscal variable taken from the October's WEO of that same year, the realized fiscal information until September is already incorporated in that forecast. Therefore, any forecast error must come from unexpected changes in the fiscal variable happening between October to December. However, we argue that, although this may be a reasonable assumption for the majority of the AEs, macroeconomic statistics for the third quarter of the year are most likely not yet available for the majority of the EMs and LICs at the time that the

October projections for the year are constructed. In practice, the October WEO projections only incorporate actual data from the first two quarters (or sometimes even one) of the contemporaneous year, which leads the time restriction assumption to be unreliable. Therefore, there is still the need to guarantee that the shocks are not responding to contemporaneous economic conditions, as we explain in section 3.2.

This section shows the importance of guaranteeing that forecast errors are purely exogenous innovations. For this exercise we focus on fiscal multipliers for personal income in EMs, but using different shocks as the $shock^f$ variable in equation 2. We plot the IRFs when using as IVs: i) uncleaned FE, ii) FE filtered from past conditions only, and iii) additionally making sure FE are orthogonal to several contemporaneous shocks impacting the economy.

Using uncleaned FEs as the IVs leads to multipliers which are almost twice as big as the estimates we get when using our cleaning approach, for every horizon considered, as figure 3 shows. In addition, unlike the case of multipliers obtained with FEs filtered only from lagged macro variables, the marginal multiplier (i.e. the period difference in cumulative multipliers) is always decreasing if estimated with FEs filtered from contemporaneous shocks. These results suggest that assuming that the forecast errors are unpredictable by past conditions of the economy, as some studies do, may invalidate the exogeneity assumption and therefore bias the results.

Figure 3: Multipliers using different ways of cleaning forecast errors.



Note: C&P multipliers. One-standard deviation confidence bands using robust standard errors, clustered at the country level. Unfiltered shocks are forecast errors which were not cleaned from any predictable component or any other contemporaneous shock to the economy. Filtered with lagged macro variables only cleans the shocks from predictable past variables. Filtered shocks make sure the shocks are orthogonal to other contemporaneous shocks, as explained in section 3.2.

4.3 One-stage vs Two-stage Estimation

The dynamic impact of fiscal shocks

Fiscal policy shocks have a dynamic impact both on GDP and on the fiscal variable itself[23]. Therefore, estimates that report fiscal multipliers as simply the impact on output from an initial fiscal shock ignore the dynamic endogenous response that fiscal variables may also have to that initial fiscal shock. Ramey and Zubairy (2018) emphasize the importance of considering this effect by showing that the higher response of output to fiscal shocks

[23]In the long run, authorities have control on the evolution of public consumption and investment. However, due to implementation lags or political feasibility, for example, some government expenditure items are not fully controlled by the authorities in the short term. The quintessential example being the quantity of government employees, which typically shows a substantial degree of inertia.

in periods of economic slack is accompanied also with a stronger response of government spending to fiscal shocks in those periods. Therefore, although output response is higher, these findings suggest that the implied fiscal multiplier does not significantly differ according to the state of the economy. This is at odds with the results of Auerbach and Gorodnichenko (2012), which do not take into consideration the dynamic behavior of the fiscal variables and therefore conclude that multipliers are higher during recessions.

This section analyzes how important is to account for this dimension even in a linear framework (that is, without considering state-dependencies). Therefore, we show how this can account for considerable differences in the multipliers reported across studies. Figure 4 plots the estimated multipliers calculated using the following one-stage regression, for each horizon $h=0,1,...,H$:

$$y_{i,t+h} = \alpha_{i,h} + \mu_{t,h} + \beta_h shock^f_{i,t} + \theta_h X_{i,t-1} + \varepsilon_{i,t+h} \tag{4}$$

and compares the IRFs to the multipliers obtained using a two-stage IV approach reported previously[24]. IRFs obtained by equation 4 focus only on the output response to a fiscal shock at time t and are commonly reported by the literature as the estimates for the fiscal multiplier.

For personal income taxation, the estimate for the impact multiplier is almost twice higher when using the two-stage estimation than the results obtained when using the one-stage approach. Similar differences are also reported for the public investment multiplier, with the second-stage approach yielding a 50% higher multiplier 2 periods after the shock than the one found when using the one-stage approach. The lower short-term impact on output we find when we use the one-stage approach is in line with, for example, estimates obtained by Furceri and Li (2017) that use a similar methodology to equation 4.

Perhaps the most striking difference in estimates corresponds to public consumption, with

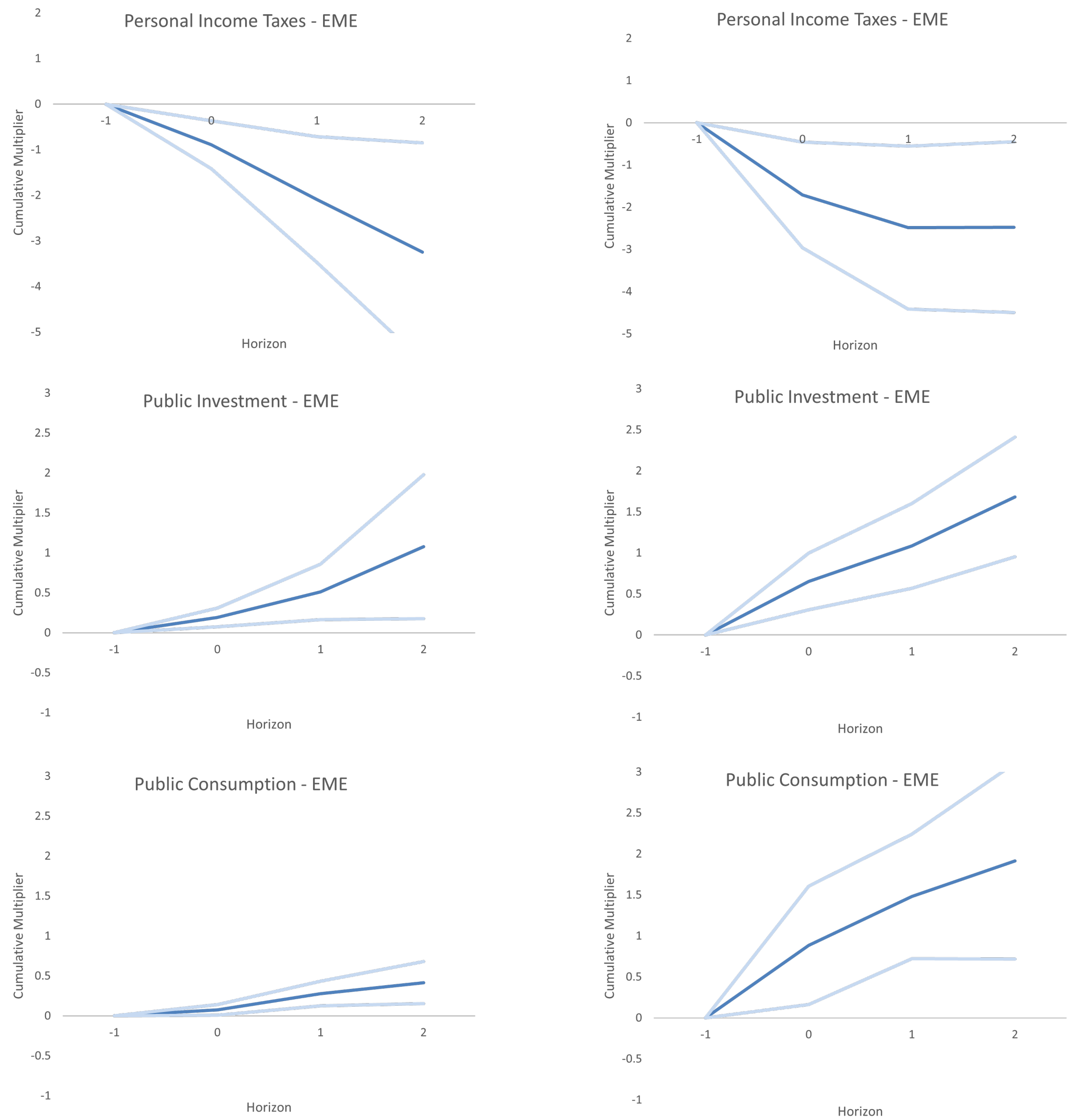
[24]Table 11 in the Appendix reports point estimates and standard errors as a result of directly including the vector of variables used to clean FEs in the two-step IV equations 1 and 2.

the multipliers from the one-stage approach being much lower than the two-stage estimation. This is in line with, for example, lower multiplier estimates obtained by Alichi et al. (2019), Honda, Miyamoto and Taniguchi (2020) or Cacciatore et al. (2021), that estimate multipliers using similar one-stage procedures[25]. This shows that results in the literature are in line to what we find if we were using a similar methodology. However, as we argue that the endogenous fiscal movements following a shock should not be ignored, this indicates that many studies are missing the important dynamic component that our approach addresses.

Ramey and Zubairy (2018) also point out that another advantage of using the two-stage IV regression is that it also stresses the importance of analyzing the relevance of the instruments. Using the one-stage approach usually overlooks this issue. Appendix 6.5 reports the F-Statistics of the first-stage regressions.

[25]A similar one-stage approach is follow by IMF (2018). However, their specification accumulates the shocks over the horizon. Although this tries to control for any future shock that may be influencing future output, it still does not capture the endogenous response of fiscal variables.

Figure 4: One vs Two-stage estimation.

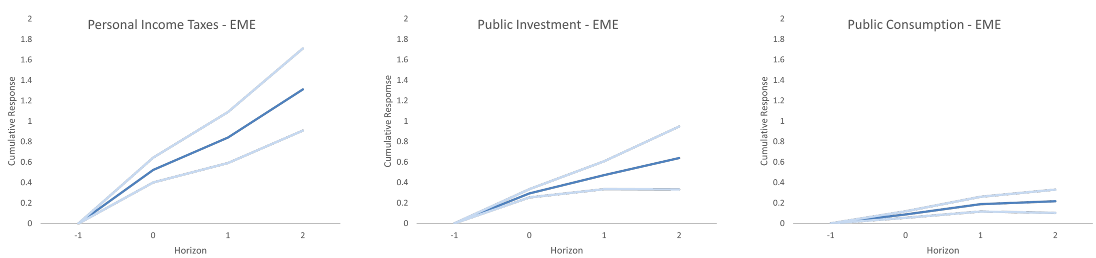


Note: C&P multipliers. One-standard deviation confidence bands using robust standard errors, clustered at the country level. Left column reports the IRFs when using directly the exogenous $shock^f$ on the second-stage regression (regression 4), right column reports the multipliers when using the two-stages IV approach (regression 1).

The reason why the multipliers obtained using the one-stage are different from the multipliers obtained using the two-stage IV approach is that some fiscal shocks create endogenous movements on the fiscal variables across the considered horizon and some others do not. Figure 5 plots the local projections for equation 2, that is the IRF across the considered horizons of the fiscal variable to the fiscal shock. If multipliers are defined as the integral of the output response divided by the integral of the fiscal variable response to a fiscal shock, the one-stage approach of regression 4 would wrongly imply a one-to-one impact of the shock in the fiscal variable across all horizons. This difference is key, especially for government consumption.

As already explained before, this results also highlight why we obtain similar multipliers for government consumption and government investment: the output response to investment shocks is stronger than the response to public consumption shocks, which is a common finding of the literature for EMs. However, our public investment shocks are more persistent - that is, they have a higher impact on public investment also in years following the shock, when compared to public consumption. Hence, when the multiplier definition used is as described in the previous paragraph, multiplier estimates for the two instruments become similar.

Figure 5: Dynamic response of fiscal variable to the fiscal shock: Local projections of first stage regressions.



Note: C&P cumulative response of fiscal tools to a shock in the same tool (regression 2). One-standard deviation confidence bands using robust standard errors, clustered at the country level.

A Glance at the Source of Financing

The financing source of a fiscal policy expansion has been recognized as another important factor that may influence the size of multipliers. The neoclassical mechanism suggests that a fiscal expansion may have low multipliers because they decrease the intertemporal wealth of private agents, given the increase in future taxes to finance current increase in expenditure (Kraay, 2012). However, specially at lower horizons, private agents may not internalize into their decisions this decrease in their present value wealth given the higher future taxes, which makes Ricardian equivalence to fail and multipliers to be higher. Ricardian equivalence may not hold for other different reasons as well: Liquidity constrained agents that make their consumption decisions based on their current income are not impacted by changes in their permanent wealth. Alternativelly, myopic agents, which ignore the intertemporal dimension of their decisions, may also make Ricardian equivalence not to hold (Chodorow-Reich, 2019). For this reason, it may or may not matter if the expansion is financed by an increase in contemporaneous revenues or by future revenues (that is, deficit-financed). The literature also highlights the higher multipliers that outside-financed fiscal expansions have relative to deficit-financed shocks. In those cases, it is not that Ricardian equivalence does not hold, but that the local government does not have to raise (present or future) taxes to finance the expansion[26].

Another advantage of estimating fiscal multipliers using the two-stage IV approach is that it allows the researcher to partially address this issue. The local projections for the first stage regression, shown in figure 5, show the dynamic response that a fiscal variable has to a shock in itself. But we can further explore the response of the primary budget balance and analyze the impact that an expenditure shock has on the dynamics of government finances. By doing so, one can understand how, *a posteriori*, an expenditure shock at time t is being financed in the following years. To do so, we compute the local projections for the following regression, for each horizon $h=0,1,\ldots,H$:

$$bb_{i,t+h} = \alpha_{i,h} + \mu_{t,h} + \beta_h shock^{f}_{i,t} + \theta_h X_{i,t-1} + \varepsilon_{i,t+h} \tag{5}$$

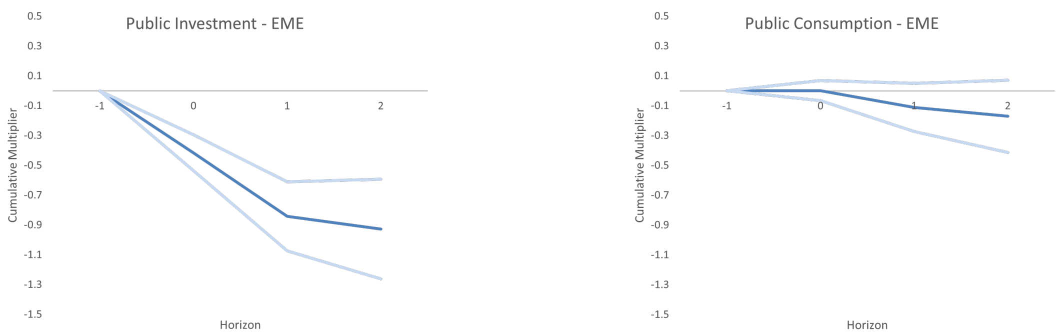
where $bb_{i,t+h}$ is the cumulative change on primary budget balance at a given horizon h.

Public investment shocks have a negative impact on government budget, as figure 6 shows. This result might suggest that the shocks we are considering are mostly deficit-financed. The reduced-form estimates do not allow us to fully disentangle the response of the budget

[26]See, for example, Coelho (2019) for the multipliers of federal transfers in eurozone countries.

balance into what is driven by automatic stabilizers or discretionary measures in response to the expansionary expenditure shock. Still, a deteriorating budget balance indicates that taxes are not increasing enough to compensate for the increase in expenditure. The same cannot be said about our public consumption shocks, as point estimates are close to zero and not statistically significant. Therefore, we cannot exclude the hypothesis that our public consumption shocks are compensated by a decrease in some expenditure component or an increase in revenues (or both).

Figure 6: Cumulative response of government budget balance to public investment and public consumption shocks.



Note: C&P cumulative response of the budget balance to a fiscal shock. One-standard deviation confidence bands using robust standard errors, clustered at the country level.

Forecast errors in different units

Instead of calculating the forecast errors as a share of GDP (or, more precisely, as the difference between the realized fiscal variable as a share of GDP and the forecasted value for the fiscal variables over the forecasted GDP), some papers calculate the forecast errors of the fiscal variable growth rates (Colombo et al., 2022). The drawback of doing so is that the forecast errors are calculated as a percentage of the fiscal variable itself, not as a percentage of GDP[27]. If the procedure used for estimating the multipliers is the one-stage procedure,

[27] For example, IMF (2017) calculates the forecast errors of growth rates of the fiscal variables. Then, they convert these forecast errors into levels using a base year. This new level series is then divided by lagged GDP, so that the variables are expressed as percentage of the same variable in the right and left hand side of equation (1).

in which shocks are included directly on the final regression (as explored throughout this section), then this causes variables on the right and lefts hand side of equation (4) to be in different units and therefore the IRFs cannot be interpreted as multipliers. However, if instead an IV two-stage procedure is used as we propose, as long as the variables in regression 1 are in the same unit, this does not change the final interpretation of the coefficients.

We proceed by showing how using the two-stage IV estimation attenuates this problem. We assess if changing the initial unit in which forecast errors are calculated changes the results. Instead of calculating forecast errors as in equation 3, we calculate them as the difference in realized and projected growth rates:

$$FE^{f}_{i,t} = f^{gr}_{i,t} - f^{gr}_{i,t|t-1}, \tag{6}$$

where $f^{gr}_{i,t}$ is the realized growth rate, and $f^{gr}_{i,t|t-1}$ is the projected growth rate. After this initial difference in their calculation, we follow the same filtering procedure explained in section 3.2.

Table 1 reports the results of using this alternative procedure with the baseline results. We can see some differences in results, especially with estimates for personal income taxes not being now statistically significant even when we consider a one standard deviation confidence interval. However, the point estimates for this fiscal tool and the other two expenditure tools are very close to the estimates obtained using the baseline specification. The main takeaway from this exercise is, therefore, that using the two-stage IV approach can limit the differences in results when identified shocks are not even in the same unit. Using these shocks directly on the the second stage regression, however, would yield IRFs that should be interpreted as elasticities instead of unit multipliers. This is particularly important because the methodology used to calculate forecast errors in different papers is not always the same.

Table 1: Multipliers when using FE of growth rates.

	Baseline		FE growth rates	
	Impact	$t+2$	Impact	$t+2$
Personal Income Taxes	-1.71*	-2.47*	-1.32	-2.68
	(1.24)	(2.02)	(1.61)	(4.16)
Public Investment	0.65**	1.68***	0.49***	1.39***
	(0.35)	(0.73)	(0.23)	(0.47)
Public Consumption	0.88*	1.91*	0.91*	1.74***
	(0.72)	(1.19)	(0.58)	(0.88)

Note: C&P multipliers for EME economies. Robust standard errors, clustered at the country level, in parentheses. Stars indicate significance at different confidence intervals: *68%, **90%, ***95%.

4.4 Different approach to accumulate responses

Even if one agrees on the definition of cumulative multipliers we use in this paper, there are still different approaches to achieve it. Variables can be accumulated across each horizon using the C&P approach, as we have been reporting the results. However, Ramey and Zubairy (2018) (R&Z) suggest a slightly different approach. They accumulate variables in the following way[28]: $y_{i,t+h} = \sum_{0}^{h} \frac{Y_{i,t+h}}{Y_{i,t-1}}$ and $f_{i,t+h} = \sum_{0}^{h} \frac{F_{i,t+h}}{Y_{i,t-1}}$. On the other hand, Auerbach and Gorodnichenko (2013) (A&G) accumulate the output response as $y_{i,t+h} = \frac{Y_{i,t+h}-Y_{i,t-1}}{Y_{i,t-1}}$[29]. However, in their paper they use their exogenous fiscal shocks directly in the regression (since they calculate fiscal spillovers from one country to the other, and not exactly multipliers of one country). Since we are using a two-stage IV approach, we also accumulate the fiscal variable in the same way for the A&G case, $f_{i,t+h} = \frac{F_{i,t+h}-F_{i,t-1}}{Y_{i,t-1}}$[30].

Table 2 shows the impact multipliers, as well as one- and two-periods ahead estimates, for the three different accumulating approaches. We report the estimates for the three variables we have being using: personal income tax, public investment and public consumption.

[28] In their paper, the authors normalize the US variables by trend GDP. Since we are working in the context of a large panel of countries, and focusing on EMEs, trend GDP estimates are less reliable and therefore we normalize our variables by GDP at $t-1$.

[29] This is also the most common approach followed in the literature.

[30] Notice that in every considered methodology, variables are normalized by output at $t-1$. As both the cumulative change in output on the right hand side of the equation, and cumulative change on the fiscal variable on the left hand side are normalized by the same variable, IRFs can be directly interpreted as unit multipliers.

Results for public investment and public consumption multipliers using both the C&P approach and the R&Z approach are very similar. However, the R&Z methodology yields higher cumulative multipliers for personal income taxes. For the multipliers of these two types of shocks, we can also see that the A&G approach inflates estimates, specially for the estimates of the multipliers two-periods after the shock occurring.

These results reinforce the importance of comparing estimates in a standardized framework. The three different approaches, because of the two-stage IV estimation, are reporting cumulative dynamic multipliers. Still, minor changes in the way variables are accumulated across the horizons leads to different multiplier estimates.

Table 2: Different approaches to calculate cumulative multipliers.

	Impact	$t+1$	$t+2$
Personal Income Taxes			
C&P	-1.71*	-2.48*	-2.47*
	(1.24)	(1.92)	(2.02)
R&Z	-1.83*	-3.27*	-4.61*
	(1.39)	(2.90)	(4.39)
A&G	-1.71*	-4.31*	-4.18
	(1.24)	(4.04)	(4.92)
Public Investment			
C&P	0.65**	1.09***	1.68***
	(0.35)	(0.52)	(0.73)
R&Z	0.71***	1.26***	1.97***
	(0.33)	(0.49)	(0.72)
A&G	0.65**	1.69***	3.62***
	(0.35)	(0.74)	(1.58)
Public Consumption			
C&P	0.88*	1.48**	1.91*
	(0.72)	(0.76)	(1.19)
R&Z	0.77*	1.35**	1.89*
	(0.65)	(0.71)	(1.19)
A&G	0.88*	1.84***	3.84
	(0.72)	(0.91)	(4.49)

Note: Reported results in the table are for EME economies. Robust standard errors, clustered at the country level, in parentheses. Stars indicate significance at different confidence intervals: *68%, **90%, ***95%. C&P accumulates the variables using the Canova and Pappa (2021) approach, R&Z accumulates the variables using the Ramey and Zubairy (2018) approach, and A&G uses the (slightly modified) Auerbach and Gorodnichenko (2013) approach.

4.5 Alternative data sources

Another source of noisy estimates may come from the use of alternative data sources, which specially for EME and LIC economies can make a considerable difference. In this section, we show how the use of data coming from the system of national accounting (SNA) or coming from the government finance statistics (GFS) may influence the estimates.

Apart from being compiled from different statistical authorities, the main difference between the two systems lies on *when* the transactions are recorded. For most of the EME and LIC economies, the GFS still uses a cash value approach: that is, transactions are recorded at the moment the cash payment related to a certain event is materialized. The SNA, on the other hand, uses an accrual approach, recording the economic transaction at the time the economic event occurs. Since most of the studies for EME and LIC economies rely on annual data, a difference of a couple of months between the economic event and the cash payment may imply that the same transaction is recorded in different years when using the two different systems.

However, the *timing* of the transaction is not the only source of differences between the two approaches. Using cash data also means that some transactions that do not have a cash flow counterpart might not be recorded. For example, donations in kind and debt forgiveness are two examples of transactions which would not be recorded in a cash data framework.

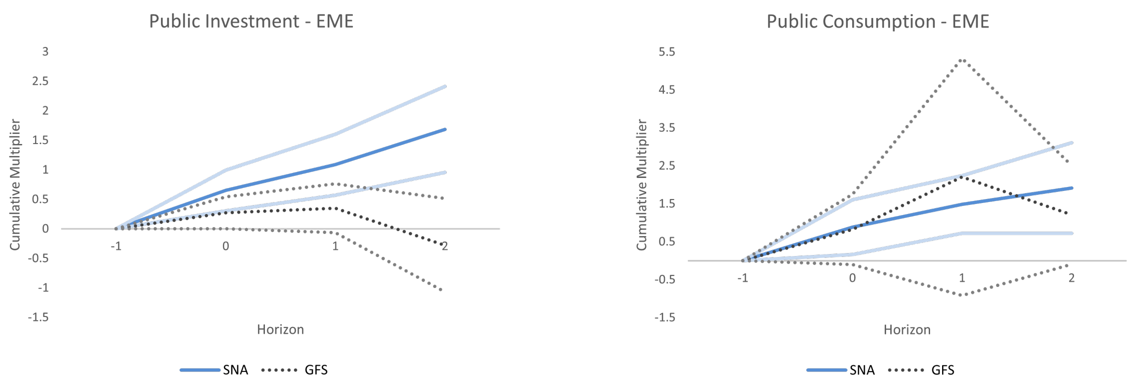
To show how these small details may affect the results, we estimate cumulative multipliers for public investment and public consumption, using both SNA and GFS data[31]. Figure 7 plots the different estimates found when using the different data sources.

What we show is that the IRFs for the different approaches differ significantly. The impact on output found when using GFS data from a public investment shock is lower compared to the impact found when using SNA data. Multipliers estimated are not statistically significant across any considered horizon when using GFS data, which sharply contrasts whith the results

[31] For SNA data, we use the following variables from the WEO database: $NFIG$ for public investment, NCG for public consumption. For GFS data, for public investment we use $GGAANT$, which stands for general government net acquisition of nonfinancial assets, and for public consumption we sum $GGECE$ and $GGEGS$, the compensation of employees and purchase of goods and services, respectively.

from SNA data. The same conclusion can be taken for public consumption. When using GFS data, considerably higher standard errors make estimates not significantly different than zero.

Figure 7: SNA vs GFM data.



Note: C&P multipliers. One-standard deviation confidence bands using robust standard errors, clustered at the country level. Shocks to public investment and public consumption are expressed as the cumulative response of GDP to a unit increase in the respective cumulative fiscal variable, at a given horizon.

5 Conclusion and Policy Implications

Fiscal multipliers in a country depend on specific economic and institutional characteristics and state contingencies. This hardens comparisons across studies, budgetary items, or countries. Moreover, methodological differences, such as the definition of the fiscal multiplier itself, can limit comparisons even further. This issue is key for policy makers, who are often confronted with limited empirical evidence for the country they work for when designing measures. Although, as Ramey (2019) shows, standardizing the methodologies to calculate multipliers can help narrowing the range of estimates reported, a uniform methodological framework is a prerequisite for comparison among studies and sound policy advise.

We conduct our exercises using a sample of 177 countries divided into three income groups (AEs, EMEs and LICs) using the same analytical framework. Although our baseline results are in line with previous findings in the literature, we show how some methodological details can largely affect the size, persistence, and precision of fiscal multipliers estimates. This analysis also allow us to emphasize good practice. Focusing on EMEs, we show that, even under the same identification strategy which relies on forecast errors of fiscal variables, slight changes in the procedure to extract fiscal innovations from forecast errors can largely bias final estimates. For example, using non-filtered forecast errors as instrumental variables leads to tax multipliers which are almost twice as big as the estimates we get when using our proposed filtering approach, for any horizon. We also explore the implications of relying on different reporting approaches (i.e. the way fiscal multipliers are defined). In general, we find that the approach followed in Auerbach and Gorodnichenko (2013) can inflate estimates in comparison with Ramey and Zubairy (2018) or Canova and Pappa (2021), especially in the medium term. Finally, we also show that using cash-based statistics can largely bias downwards the multiplier effect of government investment, an issue usually unnoticed by the literature.

In addition, our results stress the importance of using a two-stage IV procedure for the estimation of fiscal multipliers and argue against estimating them as simply the response of output to an exogenous fiscal shock. This is important mainly for two reasons. First, doing so can bias the results downwards, specially for public consumption. In addition, the dynamic endogenous movements that a fiscal shock triggers can be also informative for policy makers, as disregarding this would give a distorted picture of the impact that the policy action has not only in the economy but also in terms of the budget.

The methodological framework proposed is easily extendable to study the impact of other non-methodological differences such as economic and institutional factors. We leave that for the future.

References

Abiad, Abdul, Davide Furceri, and Petia Topalova. 2016. "The macroeconomic effects of public investment: Evidence from advanced economies." *Journal of Macroeconomics*, 50: 224–240.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91(5): 1369–1401.

Alichi, Ali, Mr Ippei Shibata, Kadir Tanyeri, et al. 2019. "Fiscal Policy Multipliers in Small States." IMF Working Paper.

An, Zidong, Joao Tovar Jalles, Prakash Loungani, and Ricardo M Sousa. 2018. "Do IMF fiscal forecasts add value?" *Journal of Forecasting*, 37(6): 650–665.

Auerbach, Alan J, and Yuriy Gorodnichenko. 2012. "Measuring the output responses to fiscal policy." *American Economic Journal: Economic Policy*, 4(2): 1–27.

Auerbach, Alan J, and Yuriy Gorodnichenko. 2013. "Output spillovers from fiscal policy." *American Economic Review*, 103(3): 141–46.

Avellán, Leopoldo, Arturo José Galindo Andrade, and John León-Díaz. 2020. "The role of institutional quality on the effects of fiscal stimulus." IDB Working Paper Series IDB-WP-01113, Washington, DC:Inter-American Development Bank (IDB).

Barnichon, Regis, Davide Debortoli, and Christian Matthes. 2022. "Understanding the size of the government spending multiplier: Its in the sign." *The Review of Economic Studies*, 89(1): 87–117.

Batini, Nicoletta, Luc Eyraud, and Anke Weber. 2014. *A simple method to compute fiscal multipliers.* International Monetary Fund.

Blanchard, Olivier, and Roberto Perotti. 2002. "An empirical characterization of the dynamic effects of changes in government spending and taxes on output." *the 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.

Cacciatore, Matteo, Romain Duval, Davide Furceri, and Aleksandra Zdzienicka. 2021. "Fiscal multipliers and job-protection regulation." *European Economic Review*, 132: 103616.

Caldara, Dario, and Christophe Kamps. 2017. "The analytics of SVARs: a unified framework to measure fiscal multipliers." *The Review of Economic Studies*, 84(3): 1015–1040.

Canova, Fabio, and Evi Pappa. 2021. "What are the likely macroeconomic effects of the EU Recovery plan?" *CEPR Discussion Paper*.

Carrière-Swallow, Yan, Antonio C David, and Daniel Leigh. 2021. "Macroeconomic Effects of Fiscal Consolidation in Emerging Economies: New Narrative Evidence from Latin America and the Caribbean." *Journal of Money, Credit and Banking*, 53(6): 1313–1335.

Chodorow-Reich, Gabriel. 2019. "Geographic cross-sectional fiscal spending multipliers: What have we learned?" *American Economic Journal: Economic Policy*, 11(2): 1–34.

Coelho, Maria. 2019. "Fiscal stimulus in a monetary union: Evidence from eurozone regions." *IMF Economic Review*, 67(3): 573–617.

Colombo, Emilio, Davide Furceri, Pietro Pizzuto, and Patrizio Tirelli. 2022. "Fiscal Multipliers and Informality." IMF Working Paper.

Ederveen, Sjef, and Richard Nahuis. 2006. "Fertile Soil for Structural Funds? A Panel Data Analysis of the Conditional Effectiveness of European Cohesion Policy." *Kyklos*, 59(1): 17+. 17.

Furceri, Davide, and Bin Grace Li. 2017. *The macroeconomic (and distributional) effects of public investment in developing economies.* International Monetary Fund.

Hernández de Cos, Pablo, and Enrique Moral-Benito. 2016. "Fiscal multipliers in turbulent times: the case of Spain." *Empirical Economics*, 50(4): 1589–1625.

Honda, Mr Jiro, Hiroaki Miyamoto, and Mina Taniguchi. 2020. *Exploring the output effect of fiscal policy shocks in low income countries.* International Monetary Fund.

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

IMF. 2015. "Making Public Investment More Efficient." *IMF Policy Papers*, June 2015.

IMF. 2017. "WEO October 2017: Seeking Sustainable Growth. Chapter 4: Cross-border impacts of fiscal policy: Still Relevant?" *World Economic Outlook*, October 2017.

IMF. 2018. "Regional Economic Outlook April 2018: Seizing the momentum. Chapter 4: Fiscal Multipliers: how will consolidation affect Latin America and the Caribbean?" *Regional Economic Outlook*, Western Hemisphere, April 2018.

Izquierdo, Alejandro, Ruy E Lama, Juan Pablo Medina, Jorge P Puig, Daniel Riera-Crichton, Carlos A Vegh, and Guillermo Vuletin. 2019. "Is the public investment multiplier higher in developing countries? An empirical investigation." National Bureau of Economic Research.

Jordà, Òscar. 2005. "Estimation and inference of impulse responses by local projections." *American economic review*, 95(1): 161–182.

Kraay, Aart. 2012. "How large is the government spending multiplier? Evidence from World Bank lending." *The Quarterly Journal of Economics*, 127(2): 829–887.

Miyamoto, Wataru, Thuy Lan Nguyen, and Dmitriy Sergeyev. 2018. "Government spending multipliers under the zero lower bound: Evidence from Japan." *American Economic Journal: Macroeconomics*, 10(3): 247–77.

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

Nakamura, Emi, and Jon Steinsson. 2014. "Fiscal stimulus in a monetary union: Evidence from US regions." *American Economic Review*, 104(3): 753–92.

Raga, Sherillyn. 2022. "Fiscal Multipliers: a review of fiscal stimulus options and impact on developing countries." *Supporting Economic Transformation.*

Ramey, Valerie A. 2016. "Macroeconomic shocks and their propagation." *Handbook of macroeconomics*, 2: 71–162.

Ramey, Valerie A. 2019. "Ten years after the financial crisis: What have we learned from the renaissance in fiscal research?" *Journal of Economic Perspectives*, 33(2): 89–114.

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.

Restrepo, Jorge. 2020. *How big are fiscal multipliers in latin america?* International Monetary Fund.

Rodríguez-Pose, Andrés, and Enrique Garcilazo. 2015. "Quality of Government and the Returns of Investment: Examining the Impact of Cohesion Expenditure in European Regions." *Regional Studies*, 49(8): 1274–1290.

Romer, Christina D, and David H Romer. 2010. "The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks." *American Economic Review*, 100(3): 763–801.

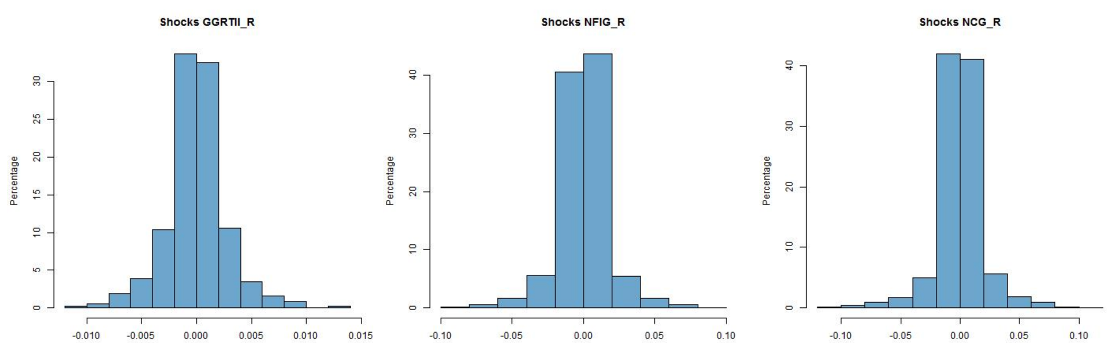
6 Appendix

6.1 Descriptive Data

Table 3 shows the classification of each country by income group. On table 4, we show some of the descriptive statistics for fiscal variables and our calculated fiscal shocks, fro the three different income groups. Additionally, on table 5 we report how does the sample size evolve on our compiled data set.

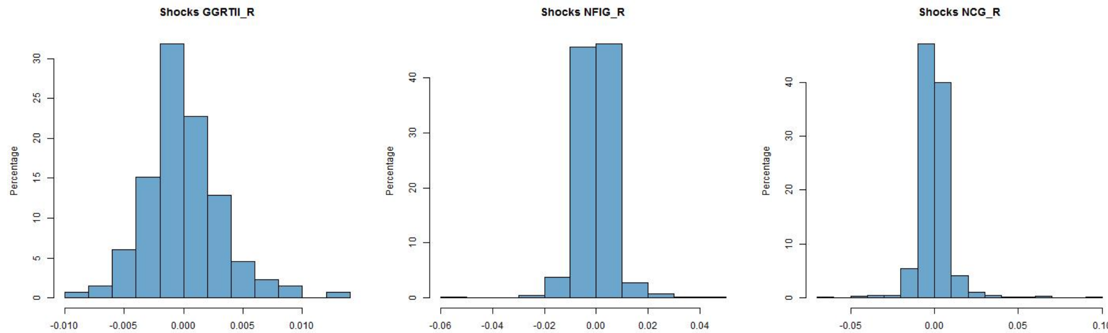
Figure 8 plots the histogram of the shocks for all sample, figure 9 plots the fiscal shocks for AEs, figure 10 plots the hocks for EMEs, and figure 11 the shocks for LICs. Shocks are in percentage of GDP.

Figure 8: Histogram of fiscal shocks for all sample, including AEs, EMEs and LICs.



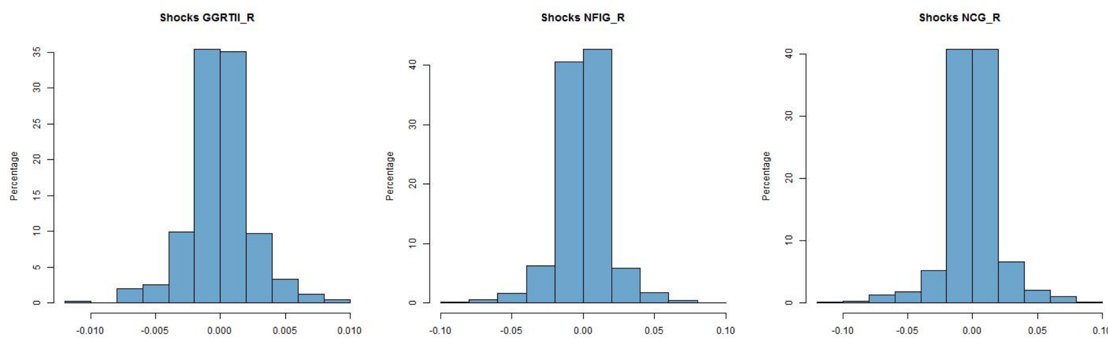
Note: GGRTII stands for personal income taxes, NFIG for public investment and NCG for public consumption. All variables are in real terms.

Figure 9: Histogram of fiscal shocks for AEs.



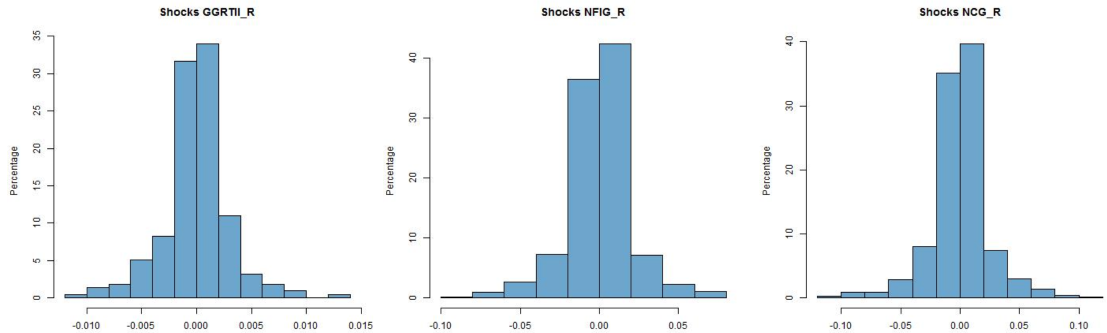
Note: GGRTII stands for personal income taxes, NFIG for public investment and NCG for public consumption. All variables are in real terms.

Figure 10: Histogram of fiscal shocks for EMEs



Note: GGRTII stands for personal income taxes, NFIG for public investment and NCG for public consumption. All variables are in real terms.

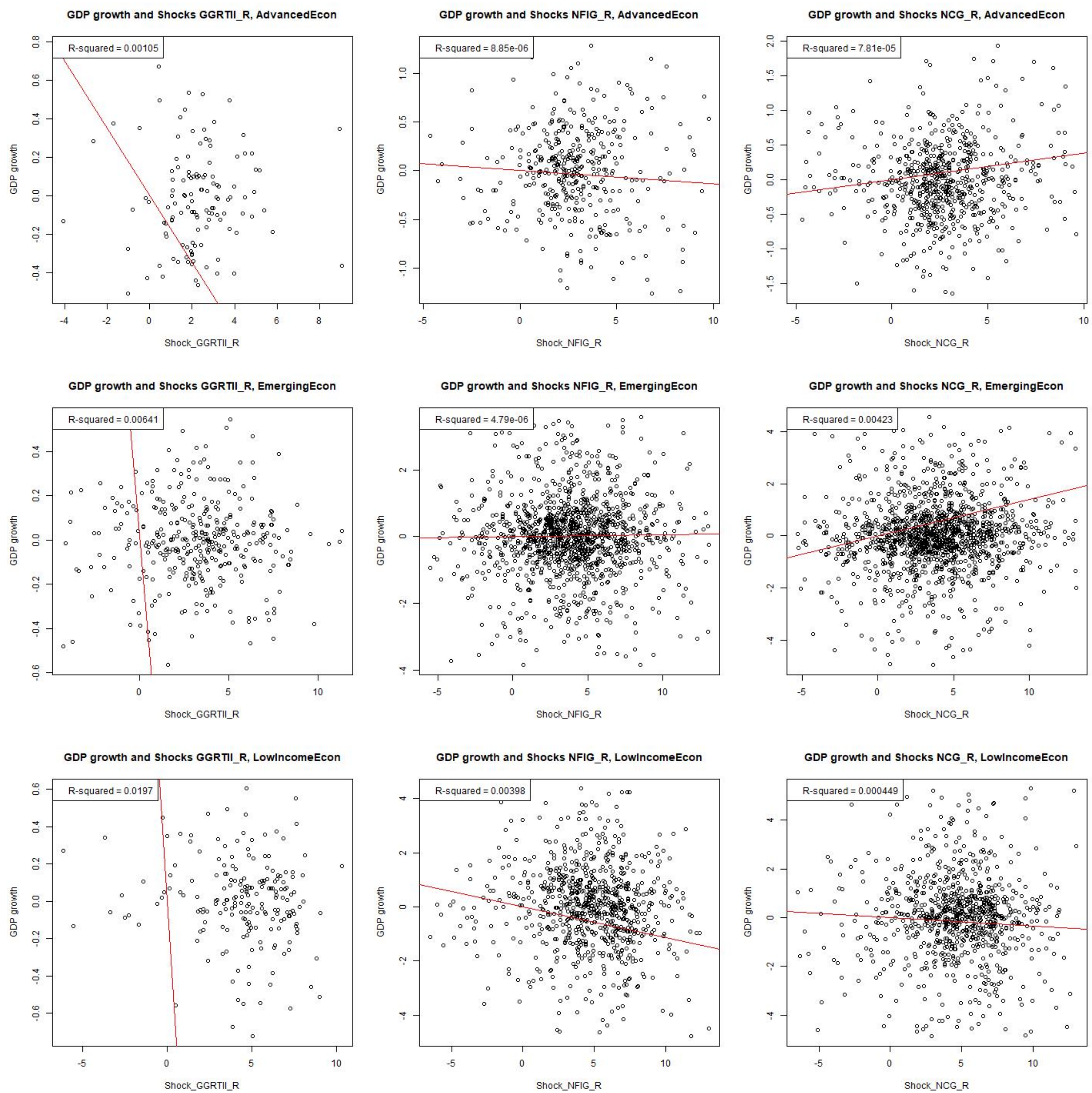
Figure 11: Histogram of fiscal shocks for LICs.



Note: GGRTII stands for personal income taxes, NFIG for public investment and NCG for public consumption. All variables are in real terms.

On panel 12 we plot the relationship between GDP growth and our estimated exogenous fiscal forecast errors. As it is clear, there seems to be no clear contemporaneous correlation between the shocks and GDP growth, in any of the three cases (AEs, EMEs and LICs). R-squared for each scatter plot, presented in the top left of each graph, is approximately zero.

Figure 12: GDP growth and Fiscal Shocks.



Note: Variables in percentage points. Fiscal shocks are the calculated forecast errors, cleaned from endogenous movements in macro variables and other forecast errors. Variables trimmed at 2.5% in each tail of the distribution for these scatter plots. First line for advanced economies, second for emerging economies, third for low income economies.

Table 3: Income classification of countries.

Country	Group	Country	Group	Country	Group
Afghanistan	LIC	Germany	AE	North Macedonia	EME
Albania	EME	Ghana	LIC	Norway	AE
Algeria	EME	Greece	AE	Oman	EME
Angola	EME	Grenada	EME	Pakistan	EME
Anguilla	EME	Guatemala	EME	Palau	EME
Antigua and Barbuda	EME	Guinea-Bissau	LIC	Panama	EME
Argentina	EME	Guinea	LIC	Papua New Guinea	LIC
Armenia	EME	Guyana	EME	Paraguay	EME
Aruba	EME	Haiti	LIC	Peru	EME
Australia	AE	Honduras	LIC	Philippines	EME
Austria	AE	Hong Kong SAR	AE	Poland	EME
Azerbaijan	EME	Hungary	EME	Portugal	AE
Bahamas	EME	Iceland	AE	Puerto Rico	AE
Bahrain	EME	India	EME	Qatar	EME
Bangladesh	LIC	Indonesia	EME	Romania	EME
Barbados	EME	Iran	EME	Russia	EME
Belarus	EME	Iraq	EME	Rwanda	LIC
Belgium	AE	Ireland	AE	Saint Lucia	EME
Belize	EME	Israel	AE	Samoa	EME
Benin	LIC	Italy	AE	San Marino	AE
Bhutan	LIC	Jamaica	EME	Saudi Arabia	EME
Bolivia	EME	Japan	AE	Senegal	LIC
Bosnia and Herzegovina	EME	Jordan	EME	Serbia	EME
Botswana	EME	Kazakhstan	EME	Seychelles	EME
Brazil	EME	Kenya	LIC	Sierra Leone	LIC
Brunei Darussalam	EME	Kiribati	LIC	Singapore	AE
Bulgaria	EME	Korea	AE	Slovak Republic	AE
Burkina Faso	LIC	Kosovo	EME	Slovenia	AE
Burundi	LIC	Kuwait	EME	Solomon Islands	LIC
Cabo Verde	EME	Kyrgyz Republic	LIC	Somalia	LIC
Cambodia	LIC	Lao P.D.R.	LIC	South Africa	EME
Cameroon	LIC	Latvia	AE	South Sudan	LIC
Canada	AE	Lebanon	EME	Spain	AE
Central African Republic	LIC	Lesotho	LIC	Sri Lanka	EME
Chad	LIC	Liberia	LIC	St. Kitts and Nevis	EME
Chile	EME	Libya	EME	St. Vincent and the Grenadines	EME
China	EME	Liechtenstein	AE	Sudan	LIC
Colombia	EME	Lithuania	AE	Suriname	EME
Comoros	LIC	Luxembourg	AE	Sweden	AE
Congo	LIC	Macao SAR	AE	Switzerland	AE
Cook Islands	EME	Madagascar	LIC	Syria	EME
Costa Rica	EME	Malawi	LIC	São Tomé and Príncipe	LIC
Croatia	EME	Malaysia	EME	Tajikistan	LIC
Cuba	EME	Maldives	EME	Tanzania	LIC
Cyprus	AE	Mali	LIC	Thailand	EME
Czech Republic	AE	Malta	AE	Timor-Leste	LIC
Côte d'Ivoire	LIC	Marshall Islands	EME	Togo	LIC
Democratic Republic of the Congo	LIC	Mauritania	LIC	Tokelau	EME
Denmark	AE	Mauritius	EME	Tonga	EME
Djibouti	LIC	Mexico	EME	Trinidad and Tobago	EME
Dominica	EME	Micronesia	EME	Tunisia	EME
Dominican Republic	EME	Moldova	LIC	Turkey	EME
Ecuador	EME	Mongolia	EME	Turkmenistan	EME
Egypt	EME	Montenegro, Rep. of	EME	Tuvalu	EME
El Salvador	EME	Montserrat	EME	Uganda	LIC
Equatorial Guinea	EME	Morocco	EME	Ukraine	EME
Eritrea	LIC	Mozambique	LIC	United Arab Emirates	EME
Estonia	AE	Myanmar	LIC	United Kingdom	AE
Eswatini	EME	Namibia	EME	United States	AE
Ethiopia	LIC	Nauru	EME	Uruguay	EME
Fiji	EME	Nepal	LIC	Uzbekistan	LIC
Finland	AE	Netherlands	AE	Vanuatu	EME
France	AE	New Zealand	AE	Venezuela	EME
Gabon	EME	Nicaragua	LIC	Vietnam	LIC
Gambia, The	LIC	Niger	LIC	West Bank and Gaza	LIC
Georgia	EME	Nigeria	LIC	Yemen	LIC

Note: The panel of countries used in the local projections is as unbalanced panel. Whenever available, we use the maximum number of countries to which we have data to, in a given year.

Table 4: Descriptive Statistics.

	AEs	EMEs	LICs
GDP growth			
Mean	2.75%	3.97%	4.35%
Std. Deviation	3.64%	7.43%	5.27%
PIT			
Mean	0.20%	0.11%	0.11%
Std. Deviation	0.47%	0.42%	0.39%
Public Investment			
Mean	0.09%	0.26%	0.31%
Std. Deviation	0.5%	2.00%	2.92%
Public Consumption			
Mean	0.47%	0.57%	0.61%
Std. Deviation	0.67%	2.14%	3.38%
Shock PIT			
Mean	0.00%	0.00%	0.00%
Std. Deviation	0.32%	0.24%	0.31%
Shock Ig			
Mean	0.00%	0.00%	0.00%
Std. Deviation	0.67%	1.69%	2.01%
Shock G			
Mean	0.00%	0.00%	0.00%
Std. Deviation	0.98%	1.98%	2.37%

Note: Sample between 1995-2019 (for PIT shocks, between 2010-2019). PIT, public investment and public consumption are presented as growth in percentage of GDP. Shock variables are the cleaned forecast errors of the respective fiscal variable, calculated as explained in the text.

Table 5: Sample size: number of countries per year.

	Shock PIT	Shock Ig	Shock G
Year			
1995	-	-	-
1996	-	-	-
1997	-	95	104
1998	-	97	102
1999	-	103	111
2000	-	117	121
2001	-	-	-
2002	-	-	-
2003	-	110	144
2004	-	118	151
2005	-	118	153
2006	-	117	155
2007	-	120	156
2008	-	121	161
2009	-	134	162
2010	50	138	159
2011	60	142	162
2012	70	149	168
2013	72	148	165
2014	76	146	169
2015	75	149	167
2016	82	152	171
2017	86	156	176
2018	85	154	177
2019	87	153	175

Note: Shocks are already the cleaned forecast errors, as a percentage of previous year GDP.

6.2 Summary of Main Estimations

Table 6: Summary of Results.

	Impact	$t+1$	$t+2$
Personal Income Taxes			
All	-0.45	-0.38	-1.35
	(1.09)	(1.59)	(2.39)
AE	-1.70*	-1.27	-2.11
	(1.32)	(1.41)	(2.37)
EME	-1.71*	-2.49*	-2.47*
	(1.24)	(1.93)	(2.02)
LIC	1.81	4.55*	2.48
	(1.90)	(3.99)	(7.59)
Public Investment			
All	0.34*	0.69*	1.14**
	(0.28)	(0.45)	(0.66)
AE	0.71*	1.28**	0.28
	(0.55)	(0.74)	(0.81)
EME	0.65**	1.09***	1.68***
	(0.35)	(0.52)	(0.73)
LIC	-0.44	-0.28	0.00
	(0.49)	(0.53)	(0.68)
Public Consumption			
All	0.96*	1.46*	1.66
	(0.74)	(0.95)	(1.82)
AE	-0.41	-1.57	-4.32
	(1.28)	(2.02)	(4.38)
EME	0.88*	1.48**	1.91*
	(0.72)	(0.76)	(1.19)
LIC	0.39	-0.36	4.18
	(2.08)	(6.18)	(11.42)

Note: C&P multipliers. Robust standard errors, clustered at the country level, in parentheses. Stars indicate significance at different confidence intervals: *68%, **90%, ***95%.

6.3 Different Specifications

Table 7 shows the multipliers when we include Covid years in the exercises. Some differences in estimates are worth mentioning: perhaps because the personal income taxes exercise uses less years than the exercises that calculate expenditure side multipliers, including these atypical years has a higher impact on tax multipliers and bias the results. Additionally, the public consumption impact multiplier also increases when we include this years: the strong fiscal response in some countries, higher fiscal forecast errors together with abnormal movements in output, as well as fast recoveries, may have influence the estimate.

Table 7: Local projections including Covid years in the estimation.

	Impact	$t+1$	$t+2$
Personal Income Taxes	1.64	2.72	-7.69
	(3.74)	(4.11)	(8.23)
Public Investment	0.66**	0.89**	1.23**
	(0.36)	(0.51)	(0.74)
Public Consumption	1.53***	1.82***	1.19
	(0.60)	(0.64)	(1.34)

Note: C&P multipliers for EME economies. Robust standard errors, clustered at the country level, in parentheses. Stars indicate significance at different confidence intervals: *68%, **90%, ***95%.

On table 8 we show the change in results for including only country fixed effects or time fixed effects separately. And table 9 shows results for different lag structures.

Table 8: Different fixed-effect specifications.

	Baseline		Country		Time	
	Impact	$t+2$	Impact	$t+2$	Impact	$t+2$
Personal Income Taxes	-1.71*	-2.47*	-1.98*	-3.45*	-1.84*	-2.94
	(1.24)	(2.02)	(1.29)	(2.5)	(1.39)	(4.29)
Public Investment	0.65**	1.68***	0.58*	1.48**	0.52**	1.11*
	(0.35)	(0.73)	(0.40)	(0.89)	(0.3)	(0.74)
Public Consumption	0.88*	1.91*	0.96*	2.15**	0.85*	1.83*
	(0.72)	(1.19)	(0.79)	(1.26)	(0.74)	(1.58)

Note: C&P multipliers for EME economies. Robust standard errors, clustered at the country level, in parentheses. Stars indicate significance at different confidence intervals: *68%, **90%, ***95%. *Baseline* uses both country and time fixed effects. *Country* specification uses only country fixed effects, *Time* specification uses only time fixed effects.

Table 10 reports the standard errors when clustered at the year-level, to account for the possibility of cross-sectional dependence, and compares with the standard errors clustered at the country-level. We see that standard errors remain practically unchanged.

Table 9: Different lag specifications.

	Baseline		No dynamic panel		No lags		Two lags	
	Impact	$t+2$	Impact	$t+2$	Impact	$t+2$	Impact	$t+2$
PIT	-1.71*	-2.47*	-1.61*	-2.26*	-1.62*	-2.18*	-1.59*	-2.49*
	(1.24)	(2.02)	(1.26)	(1.89)	(1.27)	(1.86)	(1.22)	(2.08)
Ig	0.65**	1.68***	0.68**	1.74***	0.84***	2.25***	0.60**	1.45***
	(0.35)	(0.73)	(0.39)	(0.75)	(0.39)	(0.71)	(0.34)	(0.59)
G	0.88*	1.91*	0.95*	2.20**	0.98*	2.32**	0.97*	2.05**
	(0.72)	(1.19)	(0.82)	(1.23)	(0.85)	(1.39)	(0.73)	(1.21)

Note: C&P multipliers for EME economies. Robust standard errors, clustered at the country level, in parentheses. Stars indicate significance at different confidence intervals: *68%, **90%, ***95%. *Baseline* specification includes one lag of $y_{i,t}$ and one lag of $f_{i,t}$. *No dynamic panel* specification excludes the lagged $y_{i,t}$ component, so that the panel model is no longer a dynamic one, *No lags* specification uses no lags of these variables, *Two lags* specification uses two lags of these two variables. All specifications include one lag $shock^{f}_{i,t}$ as a control.

Table 10: Different clustering levels of Standard Errors.

	PIT		Public Investment		Public Consumption	
	Impact	$t+2$	Impact	$t+2$	Impact	$t+2$
Estimate	-1.71	-2.47	0.65	1.68	0.88	1.91
Clustered Country-level	(1.24)	(2.02)	(0.35)	(0.73)	(0.72)	(1.19)
Clustered Year-level	(1.32)	(2.21)	(0.36)	(0.69)	(0.78)	(1.67)

Note: C&P multipliers for EME economies. Robust standard errors, clustered at the country-level and at the year-level, in parentheses.

6.4 Uncertainty surrounding Forecast Errors

Focusing on EMEs, in table 11 we report the point estimates and standard errors if the vector of variables used to clean the FEs, as explain in section 3.2, is directly included in equations 1 and 2 as controls. In this case unfiltered forecast errors are used as fiscal shocks.

This methodological change has, as expected, two main implications. First, it shows the robustness of baseline point estimates reported in table 2 from sub-section 4.4. Second, it increases standard errors, as the uncertainty surrounding the cleaning procedure of FEs is now directly incorporated. In any case, the final impact in terms of inference can only be assessed when incorporating the main drivers of multiplier heterogeneity.

Table 11: Estimated vs not estimated fiscal shock.

	Filtered FE		Unfiltered FE w/ controls	
	Impact	$t+2$	Impact	$t+2$
Personal Income Taxes	-1.71*	-2.47*	-1.10	-0.55
	(1.24)	(2.02)	(1.32)	(2.27)
Public Investment	0.65**	1.68***	0.49*	1.66**
	(0.35)	(0.73)	(0.37)	(0.87)
Public Consumption	0.88*	1.91*	0.20	1.08
	(0.72)	(1.19)	(0.79)	(1.54)

Note: C&P multipliers for EME economies. Robust standard errors, clustered at the country level, in parentheses. Stars indicate significance at different confidence intervals: *68%, **90%, ***95%. Filtered FE are the fiscal shocks used throughout the paper (see section 3.2). Unfiltered FE with controls represents the regression in which unfiltered forecast errors are used as IVs, but we add as controls in the regression the same variables used to filter the FE in the baseline exercises: the FE of GDP growth, FE of exchange rate, and FE of inflation, as well as lagged GDP growth, lagged government expenditure and revenues growth as percentage of GDP, lagged growth in the exchange rate and lagged inflation rate.

6.5 Instrument Relevance

Table 12 shows the F-statistic of the first-stage regressions of our baseline exercises, for each of the three different ways of accumulating responses.

Table 12: F-statistic of first stage regression for instrument relevance.

	Impact	$t+1$	$t+2$
Personal Income Taxes			
C&P	5.42	5.12	4.36
R&Z	160.45	82.11	51.53
A&G	5.42	3.68	1.52
Public Investment			
C&P	20.88	10.12	26.33
R&Z	105.25	35.00	20.61
A&G	20.88	31.32	44.43
Public Consumption			
C&P	5.79	5.27	4.99
R&Z	61.72	42.93	37.46
A&G	5.79	5.09	4.25

Note: First-stage F-statistics across different horizons and ways of accumulating multipliers. EME economies sample.

PUBLICATIONS

Getting into the Nitty-Gritty of Fiscal Multipliers: Small Details, Big Impacts

Working Paper No. WP/2023/029