

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Are Fiscal Multipliers Estimated with Proxy-SVARs Robust?\*

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Angelini G., Caggiano G., Castelnuovo E., Fanelli L. (2023). Are Fiscal Multipliers Estimated with Proxy-SVARs Robust?\*. OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 85(1), 95-122 [10.1111/obes.12521].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/896722> since: 2022-10-20

*Published:*

DOI: <http://doi.org/10.1111/obes.12521>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

**Angelini, G., Caggiano, G., Castelnovo, E. and Fanelli, L. (2022), Are Fiscal Multipliers Estimated with Proxy-SVARs Robust?\*. Oxf Bull Econ Stat.**

The final published version is available online at:

<https://doi.org/10.1111/obes.12521>

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

# Are Fiscal Multipliers Estimated with Proxy-SVARs Robust?\*

Giovanni Angelini  
University of Bologna

Giovanni Caggiano  
Monash University  
University of Padova

Efrem Castelnuovo  
University of Padova

Luca Fanelli  
University of Bologna

December 2020; This revision: March 2022

## Abstract

We estimate government spending and tax multipliers in the US via a flexible proxy-SVAR model, where identification is achieved by combining fiscal and non-fiscal instruments with additional parametric restrictions. We find that, while the spending multiplier is robustly estimated to be larger than one across different models, the estimate of the tax multiplier is sensitive to the combination of instruments that we use. We unveil that the key factor behind these heterogeneous estimates is the assumption of orthogonality between total factor productivity (our main non-fiscal proxy) and the tax shock. If this assumption is imposed, the tax multiplier is around one. If it is not imposed, the tax multiplier is three times as large. Our results point to the need of accounting for the large uncertainty surrounding the tax multiplier for the design of optimal fiscal policies.

*Keywords:* Fiscal multipliers, fiscal policy, identification, instruments, structural vector autoregressions.

*JEL codes:* C52, E62.

---

\*We thank Anindya Banerjee (Editor) and two anonymous referees for their helpful comments and suggestions. We also thank Ralf Brüggeman, Dario Caldara, Fabio Canova, Pedro Gomis-Porqueras, Giovanni Pellegrino, Bruce Preston, Søren Hove Ravn, Antti Ripatti, Xuan Zhou, and participants at seminars held at the Universities of Bologna, Helsinki, Adelaide, Deakin University. We also thank participants to the 30th EC<sup>2</sup> Conference on "Identification in Macroeconomics" (Oxford University, December 2019) and the IAAE 2002 Annual Conference (King's Business School, June 2022) for useful feedbacks. Special thanks are due to Valerie Ramey and Sarah Zubairy for an email exchange on a fiscal proxy available on Ramey's website and to Giorgio Primiceri for sharing the marginal efficiency of investment shock series estimated in Justiniano, Primiceri, and Tambalotti (2011). Giovanni Angelini thanks the University of Melbourne (where part of this research was developed) for its kind hospitality. Financial support by the Australian Research Council via the Discovery Grants DP160102281 and DP190102802 and the Italian MUR via the PRIN 2017, Grant 2017TA7TYC are gratefully acknowledged. This work is dedicated to Edoardo, Giovanni Angelini's son, who arrived the very same day this paper got accepted for publication. Corresponding author: Efrem Castelnuovo, efrem.castelnuovo@unipd.it .

# 1 Introduction

The COVID-19 pandemic has triggered the most severe recession of the post-war era for all major economies (International Monetary Fund (2020)). With policy rates at their effective lower bound, fiscal policy stimulus packages of unprecedented scale have been implemented in most countries with the aim of sustaining aggregate demand. How effective these interventions are depends on the size of the fiscal multipliers, whose quantification is a difficult task. The reason is that spending and tax revenues are in large part endogenous, and tackling this endogeneity issue to identify the output effects of exogenous variations in fiscal variables - i.e., fiscal shocks - is challenging.

One way to identify causal effects that has recently gained a lot of traction is the ‘proxy-SVAR’ (or ‘SVAR-IV’) approach, which relies on the use of instruments for the identification of the shocks of interest (see Stock and Watson (2012) and Mertens and Ravn (2013) for early contributions, and Stock and Watson (2018) for a review).<sup>1</sup> There is, however, lack of consensus on the size of the estimated multipliers. Using a narrative measure of unanticipated exogenous variations in tax revenues, Romer and Romer (2010) and Mertens and Ravn (2011b, 2012, 2013, 2014) find tax multipliers to peak between 2 and 3. Differently, Caldara and Kamps (2017), who employ Fernald’s (2014) measure of total factor productivity (TFP) to identify exogenous changes of output from fiscal policy rules from which fiscal policy shocks are recovered, estimate the tax multiplier to peak in the range between 0.5 and 0.7. Turning to the spending multiplier, Caldara and Kamps (2017) estimate it to peak between 1 and 1.3, larger than the estimates documented in the papers surveyed by Ramey (2019), which point to a 0.6-1.0 range. In light of policymakers’ need to get reliable and robust indications on the size and relative strength of the spending and tax multipliers, the heterogeneity of the estimates provided by the extant literature is problematic.

**Contributions of this paper.** This paper employs fiscal and non-fiscal policy instruments jointly in a proxy-SVAR model, and provides a framework to reconcile the previous heterogeneous estimates. We make two main contributions to the literature.

First, we show that the size of the tax multiplier is extremely sensitive to the orthogonality (exogeneity) condition between the non fiscal proxy used to identify output shocks (TFP, as in Caldara and Kamps (2017)) and the tax shock. When this assumption is imposed, the peak tax multiplier is estimated to be below 1, as in Caldara and

---

<sup>1</sup>We will use the terms ‘instruments’, ‘proxies’ and ‘external variables’ interchangeably throughout the paper.

Kamps (2017). When this assumption is relaxed as suggested by the procyclicality of tax revenues or by theories of endogenous labor productivity, the tax multiplier is three times as large and is found to be in line with Mertens and Ravn (2014). We then show that the hypothesis that the TFP proxy is uncorrelated with the tax shock is not strongly supported by US data. Considering for example the same time series, time span and predictive regression approach as in Caldara and Kamps (2017), the hypothesis of no correlation between the TFP proxy and the tax shock can be rejected if inference is robustified using heteroskedasticity and autocorrelation consistent (HAC) standard errors. We further show that empirical evidence against the orthogonality between the TFP proxy and tax shock holds across a number of proxy-SVAR specifications, starting from a standard three-variate fiscal VAR and then progressively enlarging the information set. Overall, we interpret our empirical findings as evidence that the orthogonality condition between the TFP proxy and the tax shock commonly imposed in fiscal proxy SVARs is at least questionable. To reiterate, our paper shows that imposing or not this condition has important implications for the size and the uncertainty around the estimated tax multiplier.<sup>2</sup>

Second, we find a government spending multiplier larger than one and show that, unlike the tax multiplier, the estimate is very robust to the use of different instruments and modeling specifications. Our findings show that not only the estimates of the tax and spending multipliers differ in terms of size, uncertainty and heterogeneity across models, but also for the degree of uncertainty surrounding these estimates. We show that while the spending multiplier is estimated with high precision, the uncertainty surrounding the estimated tax multipliers is large.<sup>3</sup>

**Methodology.** We obtain these results by working with a flexible proxy-SVAR that combines the observables and the instruments in an augmented specification, as proposed by Angelini and Fanelli (2019). Our proxy-SVAR specification allows (i) to directly test for the exogeneity of the proxies to the non-instrumented structural shocks

---

<sup>2</sup>Our results are robust to employing the shock to the marginal efficiency of investment (MEI) estimated by Justiniano, Primiceri, and Tambalotti (2011) as an alternative proxy for the output shock. Also in this case, we find a sharp difference in the size of the tax multiplier depending on whether the orthogonality of the MEI proxy to the tax shock is imposed in estimation, or it is relaxed as suggested by the data. Results for the MEI proxy are summarized in the online Appendix G.

<sup>3</sup>See Lewis (2021) for a remarkable exception where the opposite occurs. Lewis (2021) identifies fiscal shocks, and then calculates the fiscal multipliers, relying on the heteroskedasticity captured by the covariance structure of squared VAR residuals. Using a three-equations specification, he finds that the tax multiplier peaks at the value 0.86 after 8 quarters (it is zero on-impact and the first two quarters), while the fiscal (government) spending multiplier peaks at 0.75 after 2 quarters (it is 0.65 on-impact), but it is very imprecisely estimated, considerably more than the tax multiplier.

without resorting to additional external variables other than those used to instrument the target structural shocks, and (ii) to recover fiscal multipliers, under a proper set of conditions, also when the orthogonality of the proxies to the non-instrumented structural shocks does not hold in the data generating process.

In our framework,  $k \leq n$  external instruments can be used to identify all  $n$  latent structural shocks of the system under a set of additional parametric restrictions. The joint (point-)identification of tax and fiscal spending shocks is achieved by complementing the instruments with few additional, possibly non-controversial, parametric restrictions. These few additional restrictions involve not only the on-impact coefficients associated with the target fiscal shocks but also the on-impact coefficients associated with the auxiliary (non-fiscal) shocks. This allows us to trade relatively uncontroversial zero restrictions (e.g., the zero contemporaneous response of fiscal spending to output as in Blanchard and Perotti (2002), or the zero response of tax revenues to fiscal spending shocks as in Caldara and Kamps (2017)) with the exogeneity condition. Doing so, we unveil a negative correlation between the TFP proxy and the tax shock. Though the statistical evidence in favor of a negative correlation between the TFP proxy and the tax shock is not overwhelming, we explore the consequences of the failure of the exogeneity condition. Taking this borderline negative correlation into account, we are able to reconcile the different estimates of the tax multiplier obtained in the literature.

As regards the formal conditions that permit to identify all  $n$  structural (fiscal and non-fiscal) shocks simultaneously using  $k \leq n$  external instruments, we refine some of the results stated in Angelini and Fanelli (2019) who provide the necessary and sufficient rank condition for the identification of proxy-SVARs in the presence of multiple target shocks. In particular, we emphasize the importance of complementing the external instruments with additional restrictions that also involve the role of the non-instrumented structural shocks of the system.<sup>4</sup>

**Instruments.** Our baseline model employs both fiscal and non-fiscal instruments in a standard three-variate fiscal VAR that features quarterly US data on government spending, tax revenues and real GDP. Following Mertens and Ravn (2013), Mertens and Ravn (2014), and Caldara and Kamps (2017), we focus on the period 1950Q1-2006Q4. We consider two fiscal instruments. The first is the unanticipated tax shock proposed

---

<sup>4</sup>See the online Appendix A. Besides the seminal contribution by Mertens and Ravn (2013), a recent related work that addresses the identification of multiple shocks in fiscal proxy-SVARs is Gregory, McNeil, and Smith (2021); see instead McNeil (2022) for the monetary policy framework.

by Mertens and Ravn (2011b). The second is a proxy of (unanticipated) fiscal spending shocks that we construct by ‘purging’ the residuals obtained by regressing fiscal spending over a set of macroeconomic indicators from a measure of news spending shocks proposed by Ramey (2011). The logic behind the construction of this proxy is to remove the component which can actually be anticipated on the basis of narrative records from the one-step ahead fiscal spending forecast error. To our knowledge, ours is the first exercise in which a proxy for unexpected fiscal spending shocks is used to estimate the US fiscal spending multiplier in a proxy-SVAR. In this respect, our contribution complements the results in Ramey (2011) who focuses on the output response to anticipated fiscal spending shocks. The non-fiscal instrument is the factor utilization-adjusted TFP series produced by Fernald (2014), which - following Caldara and Kamps (2017) - we exploit to identify the output shock. While we use the two fiscal proxies to directly identify the fiscal shocks of interest, the latter instrument carry information for the identification of non-fiscal shocks that, via the moments related to the covariance matrix of the fiscal SVAR, can be exploited to recover fiscal elasticities and then spending and tax multipliers from these. Finally, in an extended version of the model, we include also inflation and the nominal interest rate, and one additional non-fiscal proxy, i.e., the oil shocks series proposed by Hamilton (2003) to instrument the inflation shock.<sup>5</sup>

**Relation to the literature.** Our point estimates of the fiscal spending multiplier fall in the 1.6-2.1 range, and are significantly larger than 1 from a statistical viewpoint. These point estimates, which are supported by different sets of instruments and model specifications, are quantitatively in line with the estimates by Caldara and Kamps’ (2017), who work with non-fiscal instruments only, Canova and Pappa (2007), who work with sign restrictions in a panel SVAR framework modeling US and EU data, and Leeper, Traum, and Walker (2017), who work with different micro-founded structural frameworks. Our estimates also support the 1.6 figure used by Christina Romer - at the time Chair of President Obama’s Council of Economic Advisers - to predict the job gains possibly generated by the stimulus package approved by the US Congress in February 2009.<sup>6</sup>

---

<sup>5</sup>As an additional check, we also include the Romer and Romer’s (2004) monetary policy shocks series. Since this series is available from 1969Q1 only, the sample that can be used for estimation shrinks (all remaining variables are available since 1954Q1). Despite in our proxy-SVAR the proxies can cover a sample period shorter than that used to estimate the SVAR, we prefer not to pursue such a route to circumvent possible parameter instabilities. We confine the discussion of this further check in the online Appendix F.

<sup>6</sup>See <https://voxeu.org/article/determining-size-fiscal-multiplier>. Auerbach and Gorodnichenko (2012), Bachmann and Sims (2012), Caggiano, Castelnuovo, Colombo, and Nodari (2015), and Ghas-

Turning to the tax multiplier, depending on the model specification and the instruments we use, we can support point (peak) estimates ranging from 0.7 to 3.6. The ‘low’ tax multiplier is obtained under the assumption that the TFP proxy is orthogonal to the tax shock, and is in line with the estimates reported in Caldara and Kamps (2017). The ‘high’ tax multiplier, obtained when the orthogonality condition is relaxed lines up with Romer and Romer’s (2010) and Mertens and Ravn’s (2011b, 2012, 2013, 2014). Our paper sheds light on the role played by different identification schemes within a class of proxy-SVAR models in delivering substantially different estimates of the tax multiplier.<sup>7</sup>

Our finding that the TFP proxy and the tax shock can be negatively correlated is consistent both with models that highlight the importance of the procyclicality of tax revenues, and with models where labor productivity may endogenously respond to structural shocks. A notable example of the former type of models is Mountford and Uhlig (2009), who characterize ‘business cycle’ shocks as generating a positive conditional correlation between output and tax revenues, and argue that this assumption is consistent with a number of theoretical views. As for the latter, Mertens and Ravn (2011a) show that permanent exogenous changes in income tax rates induce permanent changes in hours worked as well as in labor productivity with relevant implications also in the short run. Building on Mertens and Ravn (2011a), Hussain (2015) shows that exogenous labor tax increases have negative long run effects on TFP and rationalizes this finding through a DSGE model with endogenous TFP and learning-by-doing.

From a methodological viewpoint, we share with Mertens and Ravn (2014) the idea that the proxy-SVAR approach must not necessarily be confined to a ‘partial identification’ logic. Our approach is also close in spirit to Caldara and Kamps (2017). We show that their identification strategy, which exploits non-fiscal proxies to identify fiscal shocks, can be generalized to the case where the exogeneity condition fails but at the same time other restrictions that hold elsewhere in the system can be used for identification. Our approach also shows that the change in the estimated tax multiplier obtained by relaxing the orthogonality condition is due to the tight link between the size of the estimated multipliers and the parameters of the underlying fiscal policy rules.

---

sibe and Zanetti (2019) find this multiplier to be larger in recessions, and Klein and Linnemann (2019) to be particularly large during the Great Recession. For contrasting evidence, see Ramey and Zubairy (2018). Evidence on state-dependent output effects of tax shocks is provided by Sims and Wolff (2018).

<sup>7</sup>A related contribution on the heterogeneity of tax multipliers is Chahrour, Schmitt-Grohe, and Uribe (2012), who use data generated from a theoretical DSGE model in samples of length typically available to macroeconomists to show that small sample uncertainty may account for the observed differences in estimated tax multipliers.



In particular, the automatic stabilizer represented by the elasticity of tax revenues to output is found to be strongly sensitive to the violation/validity of the exogeneity condition, a result which explains why the size of the tax multipliers is also so sensitive to the exogeneity condition.

The remainder of the paper is structured as follows. Section 2 presents the econometric methodology and our identification approach. Section 3 documents our main results. In Section 4 we provide some robustness checks. Section 5 concludes.<sup>8</sup>

## 2 Econometric methodology

In this section, we first describe the identification problem in a standard proxy SVAR. We then present our approach based on a structural VAR that jointly models the observables and the instruments. We have in mind a point-identification setup. Next, we provide a specific example in the context of our baseline three variate fiscal VAR to show the flexibility of our identification approach, which allows to relax, and test for, the orthogonality conditions imposed in a standard proxy SVAR. Finally, we show what is the impact of relaxing the orthogonality conditions on the size of the multipliers.

**Setting up the problem.** Consider the following reduced-form VAR:

$$\Pi(L)Y_t = u_t \quad (1)$$

where  $Y_t$  is a vector of  $n$  observables,  $\Pi(L) \equiv I_n - \Pi_1 L - \Pi_2 L^2 - \dots - \Pi_p L^p$  is a matrix lag polynomial, and  $u_t$  is the vector of innovations with time-invariant covariance matrix  $E(u_t u_t') = \Sigma_u$ .<sup>9</sup>

Let the mapping between the vector of innovations  $u_t$  and that of structural shocks  $\varepsilon_t$  be

$$u_t = B\varepsilon_t, \quad E(\varepsilon_t \varepsilon_t') = I_n. \quad (2)$$

We focus on the identification of a subset of  $k \leq n$  structural shocks  $\varepsilon_{1,t}$ , where  $\varepsilon_t = (\varepsilon'_{1,t}, \varepsilon'_{2,t})'$ ,  $\varepsilon_{1,t}$  collects the  $k$  target shocks of primary interest which in our framework are the fiscal shocks, while  $\varepsilon_{2,t}$  collects the remaining  $n - k$  non-fiscal shocks, henceforth denoted auxiliary shocks. Then, without loss of generality, we can re-write the mapping (2) in the form

---

<sup>8</sup>An online Appendix complements this paper along several dimensions.

<sup>9</sup>Constants and other deterministic terms are here omitted for brevity. The extension of our formal expressions to cases in which constants and deterministic trends are present is straightforward.

$$u_t = B_1 \varepsilon_{1,t} + B_2 \varepsilon_{2,t} \quad (3)$$

where  $B = (B_1, B_2)$ .  $B_1$  contains the instantaneous impact coefficients associated with the shocks in  $\varepsilon_{1,t}$ , and  $B_2$  those associated with the auxiliary shocks in  $\varepsilon_{2,t}$ .<sup>10</sup> It is intended that  $\varepsilon_{1,t} \equiv \varepsilon_t$  and  $B_1 \equiv B$  when  $k = n$ .

Assume that a vector of  $r = k$  instruments,  $v_{z,t}$ , is available for the target shocks,  $\varepsilon_{1,t}$ . For such instruments to be valid, the following two conditions have to hold:

$$E(v_{z,t} \varepsilon'_{1,t}) = \Phi, \quad \text{rank}(\Phi) = k \quad (4)$$

$$E(v_{z,t} \varepsilon'_{2,t}) = 0_{k \times (n-k)}. \quad (5)$$

Condition (4) states that the  $k$  instruments have to be relevant, i.e., significantly correlated with the  $k$  structural shocks of interest  $\varepsilon_{1,t}$ ; condition (5) states that the instruments have to be uncorrelated with the  $n - k$  non-instrumented structural shocks in  $\varepsilon_{2,t}$ .  $\Phi$  is a  $k \times k$  full column rank matrix containing ‘relevance’ parameters, and the rank condition in (4) implies that each column of  $\Phi$  is non-zero and carries important information on the shocks in  $\varepsilon_{1,t}$ .

A key point of this paper is that the orthogonality of the proxies  $v_{z,t}$  to  $\varepsilon_{2,t}$  in (5) can be relaxed under the conditions we discuss below.<sup>11</sup> The setup described by (4)-(5) can be easily extended to the case  $r \geq k$ , meaning that we can allow the number of proxies to be larger than the number of instrumented structural shocks. Notably, in the very special case  $r \geq k = n$ , we can use  $n$  (or more) external variables (if available) to identify the  $n$  structural shocks embedded in the system.

Conditions (4)-(5) are consistent with the linear measurement error model

$$v_{z,t} = \Phi \varepsilon_{1,t} + \omega_t, \quad (6)$$

where  $\omega_t$  is a measurement error orthogonal to the structural shocks  $\varepsilon_t = (\varepsilon'_{1,t}, \varepsilon'_{2,t})'$ . By combining (3) with (6), we obtain the proxy-SVAR moment conditions

$$\Sigma_{u,v_z} = B_1 \Phi', \quad (7)$$

---

<sup>10</sup>We have ordered the target shocks  $\varepsilon_{1,t}$  first for convenience: as it will be clear below, the ordering of the variables is irrelevant in our framework.

<sup>11</sup>As shown later, this provides us with an important degree of flexibility and allows us to unveil the determinants of the heterogeneity in the tax multiplier estimates found in the proxy SVAR literature.

where the covariance matrix  $\Sigma_{u,v_z} = E(u_t v'_{z,t})$  can be estimated from the data using the proxies  $v_{z,t}$ ,  $t = 1, \dots, T$  and the VAR residuals  $\hat{u}_t$ ,  $t = 1, \dots, T$ , under fairly general conditions.

**The proxy-SVAR as an ‘augmented’ SVAR model.** We consider the following ‘augmented’ VAR, which combines the observables and the instruments in the system:

$$\begin{pmatrix} \Pi(L) & 0_{n \times k} \\ \Gamma(L) & \Theta(L) \end{pmatrix} \begin{pmatrix} Y_t \\ Z_t \end{pmatrix} = \begin{pmatrix} u_t \\ v_{z,t} \end{pmatrix} \quad (8)$$

where  $Z_t$  collects the ‘raw’ external variables, and  $v_{z,t} \equiv Z_t - E(Z_t | \mathcal{F}_{t-1})$ , with  $\mathcal{F}_{t-1}$  being the econometrician’s information set at time  $t-1$ .  $\Gamma(L)$  and  $\Theta(L)$  are matrix lag polynomials.<sup>12</sup> The relationship between innovations, instruments and shocks is given by:

$$\begin{pmatrix} u_t \\ v_{z,t} \end{pmatrix} = \begin{pmatrix} B_1 & B_2 & 0_{n \times k} \\ \Phi & 0_{k \times (n-k)} & P_\omega \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \omega_t^o \end{pmatrix} \quad (9)$$

where  $\omega_t^o$  denotes the measurement error normalized to have unit variance, and  $P_\omega$  denotes any symmetric positive definite matrix such that  $\omega_t = P_\omega \omega_t^o$ , and  $\Sigma_\omega = E(\omega_t \omega_t') = P_\omega P_\omega'$ .

The SVAR in (8)-(9) denoted ‘Augmented, Constrained SVAR - AC-SVAR’ in Angelini and Fanelli (2019), can be compacted in the expression

$$\tilde{\Psi}(L)W_t = \eta_t \quad (10)$$

$$\eta_t = \tilde{G}\xi_t \quad (11)$$

where

$$\begin{aligned} W_t &\equiv \begin{pmatrix} Y_t \\ Z_t \end{pmatrix}, \quad \eta_t \equiv \begin{pmatrix} u_t \\ v_{z,t} \end{pmatrix}, \\ E(\eta_t \eta_t') &= \Sigma_\eta = \begin{pmatrix} \Sigma_u & \Sigma'_{u,v_z} \\ \Sigma_{u,v_z} & \Sigma_\omega \end{pmatrix}, \quad E(\xi_t \xi_t') = I_{n+k} \\ \tilde{\Psi}(L) &\equiv \begin{pmatrix} \Pi(L) & 0_{n \times k} \\ \Gamma(L) & \Theta(L) \end{pmatrix}, \quad \tilde{G} \equiv \begin{pmatrix} B_1 & B_2 & 0_{n \times k} \\ \Phi & 0_{k \times (n-k)} & P_\omega \end{pmatrix}, \end{aligned}$$

---

<sup>12</sup>The variables  $Z_t$  are allowed to be persistent (via  $\Theta(L)$ ) and to depend on the lags of  $Y_t$  (via  $\Gamma(L)$ ). Given the large number of coefficients featured by the system of equations (8), in our empirical analysis we impose that  $\Theta(L)$  is diagonal when  $k > 1$ , i.e., the instruments are assumed to be dynamically unrelated to each other. These restrictions are supported by the data, i.e., the (cross-)correlations among the instruments used throughout the analysis are statistically equal to zero. Furthermore, in all estimated models discussed below the lag order  $q$  of  $\Theta(L)$  and  $s$  of  $\Gamma(L)$  is set to four, in line with the VAR lag order  $p$ .

$W_t$  and  $\eta_t$  being  $(n + k)$ -dimensional. With ‘ $\sim$ ’ we mean that by construction the polynomial  $\tilde{\Psi}(L)$  and the matrix  $\tilde{G}$  incorporate zero restrictions.

The necessary and sufficient rank condition for the identification of  $n$  structural shocks from system (10)-(11) can be found in Angelini and Fanelli (2019). Our Proposition 1 in the online Appendix A refines their necessary order condition for identification by emphasizing the role the non-instrumented structural shocks in  $\varepsilon_{2,t}$  play in the analysis.

**Identification strategy: relaxing the orthogonality conditions.** What are the practical advantages of representing the proxy-SVAR in the ‘augmented’ SVAR form (10)-(11)? The main advantage is the flexibility of the specification that allows us to exchange some (possibly controversial) orthogonality restrictions with other (possibly uncontroversial) parameter restrictions. This intuition, formalized in the online Appendix A is illustrated as follows.

We consider a baseline three-variate fiscal proxy-SVAR and the scenario in which we use TFP as instrument for the output shock  $\varepsilon_t^y$ , as in Caldara and Kamps (2017). The counterpart of system (3) becomes:

$$\begin{pmatrix} u_t^{tr} \\ u_t^g \\ u_t^y \end{pmatrix} = \begin{pmatrix} b_{tr,tr} & b_{tr,g} \\ b_{g,tr} & b_{g,g} \\ b_{y,tr} & b_{y,g} \end{pmatrix}_{B_2} \begin{pmatrix} \varepsilon_t^{tr} \\ \varepsilon_t^g \end{pmatrix} + \begin{pmatrix} b_{tr,y} \\ b_{g,y} \\ b_{y,y} \end{pmatrix}_{B_1} \varepsilon_t^y, \quad (12)$$

where  $u_t = (u_t^{tr}, u_t^g, u_t^y)'$  is the vector of the VAR innovations, hence  $u_t^{tr}, u_t^g, u_t^y$  are the disturbances associated with the equation for tax revenues, fiscal spending, and output, respectively;  $\varepsilon_t^{tr}$  and  $\varepsilon_t^g$  denote the tax and spending shocks, while  $\varepsilon_t^y$ , the output shock, is the target shock in this example ( $k = 1$ ). The output shock is directly instrumented by the TFP proxy, denoted  $v_t^{TFP}$ . Hence the counterpart of the linear measurement error model (6) is given by the equation

$$v_t^{TFP} = \phi_1 \varepsilon_t^y + \omega_t^{TFP} \quad (13)$$

where  $\phi_1 = Cov(v_t^{TFP}, \varepsilon_t^y)$  is the relevance parameter which captures the correlation between the TFP proxy and the output shock, while  $\omega_t^{TFP}$  is a measurement error with standard deviation  $\sigma_{\omega,TFP}$ , assumed to be orthogonal to all structural shocks in the system. Under the condition  $\phi_1 \neq 0$ , the TFP proxy would be enough to identify the output shock. Suppose now that the following two additional restrictions hold:  $b_{tr,g} = 0$ , i.e., tax revenues do not instantaneously respond (within the quarter) to the fiscal spending shock; and  $b_{g,y} = 0$ , i.e., fiscal spending does not react contemporaneously

(within the quarter) to changes in economic activity, as in Blanchard and Perotti (2002) and Mertens and Ravn (2014). The TFP proxy in (13) and the conditions  $b_{tr,g} = 0$  and  $b_{g,y} = 0$  imply the following structure for the matrix  $\tilde{G}$  in (9):

$$\tilde{G} = \begin{pmatrix} b_{tr,tr} & 0 & b_{tr,y} & 0 \\ b_{g,tr} & b_{g,g} & 0 & 0 \\ b_{y,tr} & b_{y,g} & b_{y,y} & 0 \\ 0 & 0 & \phi_1 & \sigma_{\omega,TFP} \end{pmatrix}. \quad (14)$$

It can be noticed that in this case the necessary order condition in Proposition 1(a) in Appendix A is satisfied as we have  $r = k = 1 < n = 3$ , and  $\ell_{B_2} = 1 = \frac{1}{2}(n-k)(n-k-1)$  (corresponding to the restriction  $b_{tr,g} = 0$  in  $B_2$ ). It is possible to prove that in this case the necessary order condition for identification is also satisfied and the proxy-SVAR is overidentified (hence is testable) because of  $b_{g,y} = 0$  (in  $B_1$ ).<sup>13</sup>

Consider now the case where the  $v_t^{TFP}$  proxy is not exogenous to the tax shock, so that it potentially brings information on two structural shocks. Then, the linear measurement error model for the TFP proxy becomes

$$v_t^{TFP} = \phi_1 \varepsilon_t^y + \phi_2 \varepsilon_t^{tr} + \omega_t^{TFP}, \quad (15)$$

where  $\phi_1$  has the same interpretation as before and the parameter  $\phi_2 = Cov(v_t^{TFP}, \varepsilon_t^{tr})$  captures the relationship between the tax shock and the TFP proxy. Is the model identified? The matrix  $\tilde{G}$  is now given by:

$$\tilde{G} = \begin{pmatrix} b_{tr,tr} & 0 & b_{tr,y} & 0 \\ b_{g,tr} & b_{g,g} & 0 & 0 \\ b_{y,tr} & b_{y,g} & b_{y,y} & 0 \\ \phi_2 & 0 & \phi_1 & \sigma_{\omega,TFP} \end{pmatrix}, \quad (16)$$

and has structure similar to the matrix  $\tilde{G}$  in (14), the main difference being the presence of the parameter  $\phi_2 \neq 0$  in the first column. Referring explicitly to Proposition 1(b) in the online Appendix A and the notation there used, we have one restriction in  $B_1$  ( $b_{g,y} = 0$ ), i.e.  $\ell_{B_2} = 1$ , one restrictions in  $B_2$  ( $b_{tr,g} = 0$ ), i.e.  $\ell_{B_2} = 1$ , and a novel, possibly non-zero coefficient ( $\phi_2 \neq 0$ ) which leads to a violation of the exogeneity condition relative to the specification in (14), i.e.  $g = 1$ . Thus,  $\ell_{G_1} + \ell_{B_2} = 2 = \frac{1}{2}n(n-1) - k(n-k) + g$  which ensures that the necessary order condition for identification is satisfied. It is possible to prove that with  $\phi_2 \neq 0$  a necessary and sufficient rank

---

<sup>13</sup>Notice that the fiscal elasticities of tax revenues and fiscal spending to output can be recovered from  $\tilde{G}I \equiv \tilde{G}^{-1}$ .

condition is also satisfied, hence the model is (just-)identified. Notably, the specification (16) nests the one in (14).

A key feature of our proxy-SVAR approach is that  $\phi_2$  can be estimated along with its variability, which means that we can quantify the statistical uncertainty that characterizes this crucial parameter. Also, hypotheses of the type  $\phi_2 = \check{\phi}_2$ , where  $\check{\phi}_2$  are pre-specified guess values of  $\phi_2$ , are overidentifying and testable against the data.

**Relaxing the orthogonality condition: the size of the tax multiplier.** Consider the tax policy rule, obtained by inverting the relationship  $\eta_t = \tilde{G}\xi_t$  in (11) and rearranging terms:

$$u_t^{tr} = \psi_g^{tr} u_t^g + \psi_y^{tr} u_t^y + \sigma_{tr} \varepsilon_t^{tr}. \quad (17)$$

Here  $\psi_y^{tr}$  and  $\psi_g^{tr}$  are the elasticities of tax revenues to output and to fiscal spending, respectively, and  $\sigma_{tr}$  is the standard deviation of the tax shock.<sup>14</sup> To simplify the presentation (and aligning with the empirical evidence discussed below), we assume that  $\psi_g^{tr} \approx 0$ , so that:

$$u_t^{tr} = \psi_y^{tr} u_t^y + \sigma_{tr} \varepsilon_t^{tr} \quad (18)$$

and the elasticity  $\psi_y^{tr}$  reads as an automatic stabilizer.

Consider first the case in which the orthogonality condition holds, hence  $v_t^{TFP}$  is given in equation (13). Multiplying both sides of (18) by  $v_t^{TFP}$ , taking expectations, using the orthogonality condition  $E(\varepsilon_t^{tr} v_t^{TFP}) = 0$ , and solving for  $\psi_y^{tr}$  we obtain:

$$\psi_y^{tr} = \frac{E(u_t^{tr} v_t^{TFP})}{E(u_t^y v_t^{TFP})}. \quad (19)$$

Equation (19) shows that the parameter  $\psi_y^{tr}$  equals the ratio of two reduced form covariances, which can be estimated consistently from the data under fairly general conditions.

Assume now that the orthogonality of the TFP proxy to the tax shock does not hold. The proxy  $v_t^{TFP}$  is given in equation (15), hence it is correlated not only with the output shock but also with the tax shock. As before, multiplying (18) by the proxy, taking expectations, and using  $E(\varepsilon_t^{tr} v_t^{TFP}) = \phi_2$  one obtains:

$$\psi_y^{tr} = \frac{E(u_t^{tr} v_t^{TFP}) - \phi_2}{E(u_t^y v_t^{TFP})}. \quad (20)$$

This expression shows that the elasticity  $\psi_y^{tr}$  now depends also on the correlation between the TFP proxy and the tax shock. If, as expected,  $\phi_2 < 0$ , the elasticity computed

---

<sup>14</sup>The parameters  $\psi_g^{tr}$ ,  $\psi_y^{tr}$  and  $\sigma_{tr}$  are highly nonlinear function of the non-zero  $b_{*,*}$ -coefficients that appear in the  $3 \times 3$  left-upper block of the matrix  $\tilde{G}$ .

from (20) will be larger than the elasticity computed erroneously assuming  $\phi_2 = 0$ . Following the argument in Caldara and Kamps (2017), a larger  $\psi_y^{tr}$  implies a larger tax multiplier, unveiling why the tax multiplier is sensitive to the relationship between the TFP proxy and the tax shock.

**Data and instruments.** We use US quarterly data on gross domestic product,  $y_t$ , federal tax revenues,  $tr_t$ , and government spending,  $g_t$ , defined as the sum of government consumption and investment. Following Caldara and Kamps (2017), all series are expressed in logs and real per capita terms, and are detrended by removing a linear trend. All specified VARs are treated as highly persistent, but stationary systems.<sup>15</sup> The sample covers the period 1950Q1-2006Q4, which makes our results directly comparable with those documented in the extant literature (see e.g. Caldara and Kamps (2017)), and avoids the challenge of estimating the fiscal multipliers in presence of the zero lower bound (for contributions on this issue, see Christiano, Eichenbaum, and Rebelo (2011) and Wieland (2018)). In an ‘extended’ model, we also include consumer price inflation  $\pi_t$  and the 3-month (nominal) Treasury bill rate  $i_t$ , so that  $Y_t = (y_t, tr_t, g_t, \pi_t, i_t)'$ .

In the baseline model, we include three proxies in the vector  $Z_t$ , two fiscal and one non-fiscal instrument. The two fiscal instruments are Mertens and Ravn’s (2011b) series of unanticipated tax shock (denoted  $MR$ ) which is a subset of and Romer and Romer’s (2010) shocks identified by studying narrative records on tax policy decisions, and a novel series of unanticipated fiscal spending shocks inspired by Auerbach and Gorodnichenko’s (2012) contribution (denoted  $AG$ ). This latter proxy is the residual of the OLS regression of the log of fiscal spending over a linear trend, the spending news shocks series proposed by Ramey (2011), and three lags of output, fiscal spending, tax revenues (all in logs), and Ramey’s series. Controlling for the contemporaneous (as well as the past) realizations of Ramey’s (2011) anticipated shocks helps us isolate the truly unanticipated component of fiscal spending, which is our object of interest. As stressed by Mertens and Ravn (2014), using instruments that confound unanticipated and news shocks may lead to a failure of the exogeneity assumption, and therefore invalidate the econometric analysis. Turning to the non-fiscal instruments, the proxy employed for the output shock is the total factor productivity series by Fernald (2014), denoted  $TFP$ , which is adjusted for changes in factor utilization. In our ‘extended’ model, which also

---

<sup>15</sup>As regards the way the variables are treated in the empirical analysis, in the online Appendix D we re-visit our proxy-SVAR approach along the following two dimensions: (i) we consider non-detrended variables (in levels) in a VAR endowed with a linear deterministic trend; (ii) we account for possible stochastic trends in the VAR and pursue a cointegration approach. In both cases, the results on the size of fiscal multipliers do not depart too much from the findings we obtain from the baseline specification.

includes inflation and a policy rate, we use one additional non fiscal proxy: the oil shocks series by Hamilton (2003), denoted  $OIL$ , which is a nonlinear function of the changes in the nominal price of crude oil.<sup>16</sup>

The model is estimated via Maximum Likelihood and in all specifications the reduced form VAR includes  $p = 4$  lags and a constant. We quantify the statistical uncertainty surrounding point estimates of the parameters of interest and of fiscal multipliers via the moving block bootstrap proposed by Jentsch and Lunsford (2019a).<sup>17</sup>

**Multipliers.** Let  $P$  be either the level of fiscal spending  $G$  or the level of taxes  $TR$  (not in logs);  $GDP$  be the level of output (not in logs);  $\beta y_h$  be the response of log-output at horizon  $h$  to a (one-standard deviation) fiscal policy shock; and  $\beta p_0$  be the impact of the (one-standard deviation) fiscal policy shock to the corresponding fiscal variable, expressed in logs. Then, the multiplier, defined as the dollar response of output to an effective change in the fiscal variable of 1 dollar, is given by the expression:

$$\mathcal{M}p_h = (\beta y_h / \beta p_0)(GDP/P), \quad (21)$$

where  $GDP/P$  is a policy shock-specific scaling factor converting elasticities to dollars. As in Caldara and Kamps (2017), we set the scaling factors for the two shocks of interest (unexpected change in fiscal spending and tax revenues) to their sample means on the estimation period, i.e.,  $(GDP/G)^{-1} = 0.20$  and  $(GDP/T)^{-1} = 0.18$ , respectively.<sup>18</sup> We consider positive fiscal spending shocks and negative tax shocks to compare multipliers

---

<sup>16</sup>All series but the instrument inspired by the Auerbach and Gorodnichenko (2012) paper are available in the replication package of the Caldara and Kamps (2017) paper, which is available at Dario Caldara's webpage: <https://sites.google.com/view/dariocaldara/publications>. Our AG instrument is available upon request.

<sup>17</sup>Bootstrap inference in proxy-SVARs has recently been debated by Mertens and Montiel Olea (2018), Jentsch and Lunsford (2019b), Mertens and Ravn (2019), Jentsch and Lunsford (2019a), and Montiel Olea, Stock, and Watson (2020). Throughout the paper the term 'bootstrap' refers to the moving block bootstrap (MBB). We compute MBB confidence intervals as Hall's percentile intervals - see Brüggemann, Jentsch, and Trenkler (2016) for details.

<sup>18</sup>This definition of the fiscal multipliers enhances the comparability of our results with those documented in the literature. For a discussion on this vs. alternative definitions, see Ramey (2019). It is worth noticing that our definition of fiscal multiplier in (21) corresponds to the 'alternative definition' in Caldara and Kamps (2017) (their Section 5, 'Definition of fiscal multipliers'). In our notation, the baseline definition used in Caldara and Kamps (2017) would correspond to the expression  $\mathcal{M}p_h = (\beta y_h / \sigma_p)(GDP/P)$ , where  $\sigma_p$  is the standard deviation of the policy shock, as implied by the fiscal policy rule for  $P$ . For instance, considering the simple tax policy rule in (18),  $\sigma_p = \sigma_{tr}$ . It is worth noting that the role played by the orthogonality condition in determining the size of the tax multiplier discussed throughout this paper is robust to the different definitions of multiplier.



related to shocks expected to have a positive effect on output.<sup>19</sup>

### 3 Results

In this section, we present our baseline results for three scenarios: first, the case with fiscal instruments for the identification of fiscal shocks; second, the case with the non-fiscal instrument (TFP) to identify output shocks directly, and then indirectly the fiscal shocks; third, the case with both fiscal and non-fiscal instruments to jointly identify fiscal and non-fiscal shocks. A key result is that different assumptions on the correlation between TFP shocks and tax revenues shocks lead to dramatically different estimates of the tax multiplier. Instead, the estimates related to the output effects of fiscal spending shocks are relatively robust across scenarios. We then discuss the link between different estimates of the output-tax elasticity and the corresponding tax multiplier.

#### 3.1 Fiscal instruments only approach

##### **Fiscal spending shock: Auerbach and Gorodnichenko’s (2012) instrument.**

We begin our analysis by instrumenting the fiscal spending shock with our novel AG proxy, which is meant to identify unexpected changes in fiscal spending. In this case,  $Y_t = (y_t, tr_t, g_t)'$ ,  $Z_t = (AG_t)$ , and  $\varepsilon_{1,t} \equiv \varepsilon_t^g$ , and we estimate the ‘augmented’ model for  $W_t = (Y_t', Z_t')' = (y_t, tr_t, g_t, AG_t)$ . While the proxy  $AG_t$  identifies the fiscal spending shock  $\varepsilon_t^g$ , we achieve just identification of all shocks (i.e. also the tax shock and the output shock in  $\varepsilon_{2,t} \equiv (\varepsilon_t^{tr}, \varepsilon_t^y)$ ) by imposing that fiscal spending does not instantaneously respond to output shocks.<sup>20</sup> The robust first-stage F-statistic for this instrument is

---

<sup>19</sup>We also report in Appendix F the cumulative spending multiplier, defined as

$$\mathcal{M}_{p_h}^c = \frac{\sum_h \beta y_h}{\sum_h \beta p_h} \frac{GDP}{P}$$

which accounts for the persistence of the response of government spending to its own shock. We do not report cumulative tax multipliers. Given the strong feedback from GDP to tax revenues, the concept of a cumulative tax multiplier is not well defined, and its calculation is problematic (see Mertens and Ravn (2013) and Ramey (2019)).

<sup>20</sup>Formally, this is the constraint  $b_{g,y} = 0$  discussed in Section 3 (albeit for a different proxy-SVAR). Blanchard and Perotti (2002) impose a zero contemporaneous response of fiscal spending to all shocks affecting output. The two restrictions are equivalent if output is not affected by fiscal shocks at time  $t$ . If it is, our restriction is less stringent than Blanchard and Perotti’s (2002). The difference in these restrictions is due to the fact that they work with an “AB-model” (Lütkepohl (2005)) which accounts also for the contemporaneous relationships among the variables. Differently, we work with a “B-model”, which focuses directly on the mapping going from the structural shocks to the VAR innovations.

2019.58. For brevity, the maximum likelihood estimates of the implied matrix  $\tilde{G}$ , along with 68%-bootstrap confidence intervals, are confined in Appendix C.

Figure 1 (left panel) plots the fiscal spending multiplier obtained from this specification. The on-impact multiplier ( $\mathcal{M}_{g_0}$  in our notation) is about 1.1, it increases to about 1.6 after 2 quarters, it stays at that level for about one year, then it gradually declines. The 68%-bootstrap confidence interval associated with the peak multiplier, reported in Table 1, ranges from 1.1 to 2. While the just identified model cannot be offered formal statistical support by the overidentification restriction test, we notice that the estimated relevance parameter, which connects the  $AG_t$  instrument to the fiscal shock  $\varepsilon_t^g$ , is  $\hat{\phi}_{AG} = 0.0129$ , is strongly significant, and implies a correlation of 96% with the identified fiscal shock.<sup>21</sup>

Table 1 collects our estimate of the output-spending elasticity, given by  $\psi_y^g = -(\tilde{GI}_{3,1}/\tilde{GI}_{3,3})$ , where  $\tilde{GI} \equiv \tilde{G}^{-1}$ , and  $\tilde{GI}_{i,j}$  is the element located in the  $i$ -th row and  $j$ -th column of the  $\tilde{GI}$  matrix. We get a point estimate of  $\hat{\psi}_y^g = -0.0029$ , and the associated 68%-bootstrap confidence interval is  $(-0.027, 0.025)$ . This finding supports Blanchard and Perotti's (2002) choice of calibrating such elasticity to zero. Caldara and Kamps' (2017) analytical derivations show that a zero elasticity implies an on-impact multiplier equal to 1, which is in line with what we find.

**Tax shock: MR instrument.** We now turn to the identification of the tax revenues shock. The instrument we use is the series of unanticipated tax shock produced by Mertens and Ravn (2011b),  $MR_t$ . Since  $Y_t = (y_t, tr_t, g_t)'$ ,  $Z_t = (MR_t)$  and  $\varepsilon_{1,t} \equiv \varepsilon_t^{tr}$ , we estimate the augmented model for  $W_t = (Y_t', Z_t')' = (y_t, tr_t, g_t, MR_t)'$ . In this case  $MR_t$  identifies directly the tax shock  $\varepsilon_t^{tr}$  but, consistently with the previous case, we achieve identification of all shocks (i.e. also the fiscal spending shock and the output shock in  $\varepsilon_{2,t} \equiv (\varepsilon_t^g, \varepsilon_t^y)$ ) by imposing the restriction that fiscal spending does not instantaneously respond to output shocks. The robust first-stage F-statistic for this instrument is 1.55. The correlation between the residual associated with the tax revenue equation of the VAR,  $\hat{u}_t^{tr}$ , and the  $MR_t$  instrument is equal to 12%.<sup>22</sup> The point estimate of the rele-

<sup>21</sup>An alternative framework to estimate fiscal multipliers using external proxies is local projections. We find that our estimates of the on-impact fiscal spending multiplier, reported in Figure 1 (left panel), are in line with those obtained by Ramey and Zubairy (2018) through a linear local projections approach à la Stock and Watson (2018). Despite employing different definitions of variables, a different instrument for the fiscal spending shock, and a longer span of data, Ramey and Zubairy (2018) estimate the on-impact fiscal spending multiplier to be around 1.3, with 95% confidence bands given by [0.6, 1.9], which clearly include our estimated on-impact spending multiplier.

<sup>22</sup>If the correlation is computed by considering only the non-zero elements of  $MR_t$  (and the corresponding elements in  $\hat{u}_t^{tr}$ ), the correlation increases to 35%.

vance parameter for the  $MR_t$  instrument is  $\hat{\phi}_{MR} = 0.043$ , which implies a correlation of 27% with the identified tax shock.<sup>23</sup>

Figure 1 (right panel) plots the implied tax multiplier. The multiplier is 2.1 on impact ( $\mathcal{M}tr_0$ ), and reaches a peak value of 3.1 after 3 quarters. The size of the multiplier is in line with the estimates by Mertens and Ravn (2014) and part of the literature cited therein. The 68%-bootstrap confidence interval for the peak tax multiplier ranges from 1.4 to 4.8. We then recover the output-tax elasticity as  $\psi_y^{tr} = -(\widetilde{GI}_{2,1}/\widetilde{GI}_{2,2})$ . Conditional on the estimated model, the point estimate is  $\hat{\psi}_y^{tr} = 3.36$ , close to that reported in Mertens and Ravn (2014) and Mertens and Ravn (2011a), 3.13 and 3.7 respectively. The confidence interval for  $\psi_y^{tr}$  is [2.25, 4.45]. Although it reflects sizeable uncertainty about the value of this elasticity, the lower bound is still higher than the value 2.08 used by Blanchard and Perotti (2002), who rely on an application of the OECD methodology documented in Giorno, Richardson, Roseveare, and van den Noord (1995), and is considerably higher than the value 1.7 produced by Follette and Lutz (2010) for the US economy. We postpone the discussion on the plausibility of an output-tax elasticity around the value 3 to Section 3.3.

### 3.2 TFP only approach

We use Fernald’s (2014) measure of TFP,  $TFP_t$ , as an instrument for output shocks. While such shocks are not of direct interest for the computation of the fiscal multipliers, as shown by Caldara and Kamps (2017), the information related to their impulse vector can be fruitfully combined with that of the covariance matrix of our VAR to achieve full identification and recover the output effects of fiscal spending and tax shocks.<sup>24</sup> Thus, we have  $Y_t = (y_t, tr_t, g_t)'$ ,  $Z_t = (TFP_t)$  and  $\varepsilon_{1,t} \equiv \varepsilon_t^y$ , and we estimate the ‘augmented’ model for  $W_t = (Y_t', Z_t')' = (y_t, tr_t, g_t, TFP_t)'$ . The robust first-stage F-statistic for this instrument is 61.61. To identify the three structural shocks of the system, we impose the two restrictions  $b_{tr,g} = 0$  and  $b_{g,y} = 0$  in the matrix  $\tilde{G}$  in (14). While  $b_{g,y} = 0$  (fiscal spending does not instantaneously respond to output shocks) is consistent with the proxy-SVARs estimated in the fiscal instruments only approach, the restriction  $b_{tr,g} = 0$  (tax revenues do not instantaneously respond to fiscal spending shocks) is necessary for the identification of the model. The proxy-SVAR is overidentified, and

<sup>23</sup>These results would motivate the use of the weak-instrument robust approach for proxy-SVAR developed by Montiel-Olea, Stock and Watson (2020). For comparative purposes we stick to Mertens and Ravn’s (2014) approach and do not pursue the test inversions route.

<sup>24</sup>For an early study on the connection between policy rules and policy shocks with an application to the identification of monetary policy shocks, see Leeper, Sims, and Zha (1996).

the overidentification restrictions test returns a p-value of 0.41, which leads to not rejecting the model specification. The point estimate of the relevance parameter is  $\hat{\phi}_1 = 1.86$ , which implies a correlation of 57% with the identified output shock.

As shown by Figure 1, the point estimates of the fiscal spending multipliers identified with the TFP proxy turns out to be in line with the ones computed with the AG instrument. The impact multiplier ( $\mathcal{M}_{g0}$ ) is equal to 1.1, while the peak - which occurs after two quarters - is equal to 1.9, with associated 68%-bootstrap confidence interval which ranges from 1.3 to 2.4. The point estimate of the elasticity of fiscal spending to output  $\psi_y^g$  is negative, and zero is not included in the confidence interval (even though the upper bound is very close to zero). Overall, these results are close to those reported in Caldara and Kamps (2017).

Turning to the tax multiplier, the estimate is substantially lower relative to that obtained with the MR instrument. On impact, the multiplier is estimated to be 0.4, and the peak value - 0.76 - realizes 5 quarters after the shock. The 68%-bootstrap confidence interval for the peak tax multiplier ranges from a value slightly less than zero to 0.93. Figure 1 shows that the drop of the tax multiplier relative to the MR case is substantial for at least 25 quarters after the shock. What is the driver of this drastic change in the tax multiplier when moving from the MR case to the use of the TFP proxy for the output shock? Table 1 collects the estimated value of the tax policy coefficient  $\psi_y^{tr}$  in this scenario, which is 2.1, with associated confidence interval [1.8, 2.5]. The estimated elasticity, as well as the associated confidence interval, is significantly lower than the estimate obtained when using the MR instrument only. The fact that lower values of the tax elasticity,  $\psi_y^{tr}$ , are associated with lower values of the multiplier,  $\mathcal{M}_{tr}$ , is consistent with the simulations proposed in Mertens and Ravn (2014), and with the analytical derivations documented in Caldara and Kamps (2017).

### 3.3 TFP only approach: Relaxing the TFP-tax shock orthogonality condition

In all the previous proxy-SVARs, the multipliers have been estimated assuming the orthogonality of the TFP instrument to both fiscal shocks. While the exogeneity assumption for the spending shock is based on the well-known delays characterizing fiscal spending decisions and implementations (Blanchard and Perotti (2002)), the hypothesis that TFP is uncorrelated with the tax shock needs further discussion.

**TFP-tax shock orthogonality condition: Empirical and theoretical evi-**

**dence.** Our model specification allows to analyze empirically the link between the TFP proxy and the tax shock identified with our proxy-SVAR. The analysis is carried out ‘internally’ to the proxy-SVAR, i.e. without resorting to other external proxies in addition to those used for the instrumented shocks. This marks a difference with Caldara and Kamps (2017), who instead need an external observable proxy for the tax shock in their predictive regression approach.

We consider the model (15)-(16), where the TFP jointly serves as an instrument for the output shock and the tax shock, respectively. Given the matrix  $\tilde{G}$  in (16), we can estimate and make inference not only on the parameter  $\phi_1$ , i.e. the relevance parameter for output shocks, but also on  $\phi_2$ , which connects the TFP instrument to the tax shock. We find that the connection of the TFP instrument to the output shock is strongly supported by the data, and the connection of the TFP instrument to the tax shock, though less strong, appears non-negligible. More precisely, in line with the evidence in Caldara and Kamps (2017), the TFP turns out to be an instrument significantly correlated with the output shock: the point estimate of the relevance coefficient,  $\hat{\phi}_1$ , is equal to 1.63, with a 68%-bootstrap confidence interval of  $[1.42, 2.01]$ , while the implied correlation with the output shock is 49.7%. More important for our analysis, the TFP proxy turns out to be linked with the tax shock as well: the estimated relevance coefficient is  $\hat{\phi}_2 = -0.89$ , with the 68%-bootstrap confidence interval given by  $[-1.51, -0.64]$ , and the resulting correlation between TFP and the identified tax shock is -27%.<sup>25</sup>

To shed further light on the orthogonality condition, we revisit Caldara and Kamps’s (2017) inference on the exogeneity of TFP to the tax shock based on a predictive

---

<sup>25</sup>The following remark is in order. As is known, classical confidence intervals are, at least implicitly, defined by inverting a test. Our 68%-bootstrap confidence interval for  $\phi_2$  can then be regarded as the set of parameter values for which a bootstrap Wald-type test does not reject the null hypothesis  $\phi_2 = \check{\phi}_2$  at the 32% nominal significance level. Thus, exploiting the duality between hypothesis testing and confidence intervals, we can conclude that all values that lie outside the interval  $[-1.51, -0.64]$ , including  $\phi_2 = \check{\phi}_2 = 0$ , are rejected at the 32% nominal significance level. However, it is important to notice that, when the parameter  $\phi_2$  is pre-fixed to the value  $\phi_2 = \check{\phi}_2$ , the proxy-SVAR becomes overidentified. We could then use likelihood-ratio (LR) tests for restrictions of this type, which might not be consistent with the inference based on the bootstrap confidence interval. For instance, when we fix  $\phi_2 = \check{\phi}_2 = -1.51$  (which corresponds to the lower bound of the confidence interval for  $\phi_2$ ), the p-value of the LR overidentification restrictions test is 0.25; for  $\phi_2 = \check{\phi}_2 = -0.64$  (which corresponds to the upper bound of the confidence interval for  $\phi_2$ ) the p-value of the LR overidentification restrictions test is 0.75; for  $\phi_2 = \check{\phi}_2 = -0.25$  the p-value is 0.51 (0.54 if a bootstrap version of the test is considered) and for  $\phi_2 = \check{\phi}_2 = 0$  the p-value is 0.41 (0.44 if a bootstrap version of the test is considered). According to the p-values implied by the LR overidentification tests, the evidence in favour and against the orthogonality hypothesis is not conclusive.

regression approach. Differently from our approach, Caldara and Kamps (2017) use an observable external proxy for the tax shock in order to run their test of exogeneity. They regress the TFP proxy against the narrative the Mertens and Ravn’s (2011) narrative measure of tax shock and the Ramey’s (2011) narrative measure of expected exogenous changes in military spending. We replicate their analysis and confirm that the  $t$  and  $F$ -tests support evidence in favour of the assumption of exogeneity of the TFP proxy. However, the same predictive regression approach returns  $t$  and  $F$ -tests that, once the inference is robustified computing HAC-type standard errors to account for residuals autocorrelation, lead to (marginally) reject the ‘no correlation’ hypothesis. Detailed results are reported in Appendix B. Overall, we find that the empirical evidence is not convincingly supportive of the condition  $\phi_2 = Cov(v_t^{TFP}, \varepsilon_t^{tr}) = 0$ .

The connection between factor productivity and exogenous tax shocks has been examined by Mertens and Ravn (2011a). They show that, in the context of a stochastic growth model, permanent changes in income tax rates induce permanent changes in labor productivity. This finding violates the standard long-run identification strategy for technology shocks, based on the assumption that neutral technology shocks are the only source of long run changes in labor productivity. Mertens and Ravn (2011a) also show empirically using a VECM that tax shocks affect productivity significantly both in the short and in the long run. They also highlight several channels through which permanent income tax changes can have permanent effects on labor productivity: in models with endogenous changes in labor productivity, such as models with educational choices or human capital accumulation, increases in labor income taxes lower the return on skills, and decrease labor productivity. In life cycle models, changes in labor income tax rates can affect the retirement decisions of older workers, and this may negatively affect labor productivity if skills are accumulated over the life cycle. An alternative mechanism works through the government budget constraint. For a given level of public spending, a change in labor income taxes will lead to a change with opposite sign in capital income taxes, which affects the long run level of labor productivity. Taken together, their findings point to rejection of the standard identifying assumption for productivity shocks based on the long run orthogonality between tax changes and factor productivity. Hussain (2015) provides further VAR-based evidence that exogenous permanent increases in taxes have strong, permanent, and negative effects on TFP. He then rationalizes this finding with a DSGE model with endogenous TFP and human capital accumulation. In his model, learning-by-doing takes the form of an externality: tax increases reduce human capital accumulation and labor productivity, because the

TFP of all firms depends on the aggregate level of human capital.

We now turn to the key question of whether relaxing the exogeneity condition makes an important difference as regards the size of the estimated tax multiplier.

**TFP-tax shock orthogonality condition: Implications for the multipliers.**

What are the implications for the multipliers? We first look at the peak fiscal spending multiplier, which is estimated to be around 2 with confidence interval ranging from 1.4 to 2.6. This figure is slightly larger than, but not statistically different from, those found when imposing the TFP-tax shock orthogonality condition. Quite differently, the impact on the tax multiplier is dramatic, with the peak value jumping from 0.7 to 3.6. This latter figure is statistically in line with the tax multiplier around 3 estimated with the  $MR$  instrument. Admittedly, allowing for the non-zero correlation does not come without costs. The confidence interval for the peak tax multiplier ranges from 0.2 to 5.9, hence it tends to be larger relative to the confidence interval obtained with the  $MR_t$  instrument alone. We will discuss this issue in more depth when we present the results obtained when using multiple instruments.

What is the driver of the substantial difference between the small tax multiplier found when imposing the TFP-tax shock orthogonality and the one around 3 obtained by relaxing such restriction? Mertens and Ravn (2014) and Caldara and Kamps (2017) document the mapping between the output-tax elasticity and the tax multiplier. In particular, Caldara and Kamps (2017) derive an analytical expression for the tax multiplier and show that, if  $\psi_y^{tr}$  belongs to the  $(-1, 4)$  range, there is a positive correlation between the elasticity and the multiplier. Table 1 documents the substantial change in such elasticity when the TFP-tax shocks orthogonality is relaxed, with  $\hat{\psi}_y^{tr}$  moving from 2.1 (orthogonality imposed) to 3.8 (orthogonality relaxed). This latter number is pretty close to the 3.7 estimate provided by Mertens and Ravn (2011a), who employ long run restrictions to identify movements in output due to a technology shock to tackle the tax-output endogeneity bias. Moreover, the associated confidence interval (2.3, 4.9) implies that estimates around 3 that are often found in the literature are statistically equivalent to ours.

**Output-tax elasticity equal to 3: How sensible?** As stated above, Blanchard and Perotti (2002) rely on an output-tax elasticity equal to 2.08, which is the one estimated by the OECD (Giorno, Richardson, Roseveare, and van den Noord (1995)). Such elasticity is slightly larger than that estimated by Follette and Lutz (2010) on yearly data (1.7). Instead, our results point to output-tax elasticities equal to 3 or larger.

Are such large elasticities sensible? Mertens and Ravn (2014) critically review the construction of output-tax elasticity by the OECD, which is a weighted average of the output elasticities for different tax revenue components (personal income taxes, social security contributions, indirect taxes and corporate income taxes). Each component-specific elasticity is a product of two elasticities, i.e., the tax base-tax revenues one and the output-tax base one. Mertens and Ravn (2014) point out that, while both elasticities are (somewhat necessarily) computed by relying on many questionable assumptions, the second one in particular is typically estimated via OLS regressions that do not tackle the obvious endogeneity issue affecting the output-tax relationship. Importantly, Mertens and Ravn (2014) show that such endogeneity issue is likely to induce a negative bias in the estimated output-tax elasticity. As pointed out above, Mertens and Ravn (2011a) tackle this bias by estimating the response of the US federal tax revenues to a technology shock identified with long run restrictions, and find a value for the elasticity equal to 3.7. Caldara and Kamps (2017) derive the output-tax elasticity implied by the sign restriction approach pursued by Mountford and Uhlig (2009), and find a value equal to 3. Overall, a value of the output-tax elasticity equal to 3 or larger does not seem at odds with the US data.

### 3.4 Multiple instruments approach

As stressed in the Introduction and in Section 3, the methodology we work with allows us to jointly employ multiple instruments. In this Section, we combine all instruments jointly to re-estimate both multipliers. This is a novelty in the proxy-SVAR literature, since in our case the number of employed external instruments,  $k$ , is the same as the number of variables,  $n$ , of the original SVAR, i.e., all structural shocks of the system are instrumented.<sup>26</sup> We estimate the ‘augmented’ model for  $W_t = (Y_t', Z_t')' = (y_t, tr_t, g_t, MR_t, AG_t, TFP_t)'$ . As before, we analyze two cases, one in which we impose the the TFP-tax shock orthogonality condition, and the other one in which we do not.

#### **Fiscal shocks: AG & MR & TFP instruments - orthogonality condition.**

Figure 2 shows the fiscal spending and tax multipliers when we assume that the orthogonality condition holds. The fiscal spending multiplier peaks at a value equal to

---

<sup>26</sup>Gregory, McNeil and Smith (2021) also consider a model in which the number of external instruments is the same as the number of variables, but in a more restrictive setup where the matrix  $\Phi$  in (6) is diagonal, a necessary condition to allow the parameters in the matrix  $B_1 \equiv B$  to be unrestricted (see Proposition 1(d) in Appendix A).



1.8, which is relatively similar to those found in the one-instrument scenarios. Again, this multiplier is precisely estimated as the associated confidence interval ranges from 1.3 to 2.2. The peak realization of the tax multiplier is 1, with associated confidence interval ranging from 0.4 to 1.3. While being larger than the one estimated with the TFP instrument only under the assumption of TFP-tax shock orthogonality (0.76), this value is three times smaller than the one obtained with the TFP instrument only when the orthogonality condition is relaxed. From a statistical standpoint, this model - which is overidentified - is supported by the data, the p-value of the overidentification restriction test being 0.72.

**Fiscal shocks: AG & MR & TFP instruments - non orthogonality.** We next turn to the case where the TFP-tax shock orthogonality condition is not imposed. Figure 2 documents the spectacularly different implications for the two multipliers. The impact of relaxing the orthogonality condition on the estimated fiscal spending multiplier is virtually zero, i.e., the multiplier is exactly the same as the one estimated when imposing such condition. Differently, the tax multiplier records a peak value of 2.8 vs. the value of 1 estimated when imposing the orthogonality condition; the associated confidence interval ranges from 0.4 to 4.3. Figure 2 shows that the estimated tax multiplier under non-orthogonality is clearly not contained in the confidence interval surrounding the point estimates of the tax multiplier conditional on the assumption of orthogonality. As before, the driver of this dramatic increase of the value of the tax multiplier under non-orthogonality is the impact of the orthogonality/non-orthogonality assumption on the estimated output-tax elasticity, which moves from 2.3 (orthogonality imposed) to 3.3 (orthogonality not imposed). Turning to the output-fiscal spending elasticity, our model allows us to estimate it jointly with the rest of the system. Our point estimate, which is zero, lends once again support to the Blanchard and Perotti (2002) zero restriction typically used in this literature. Finally, this model estimated with multiple instruments and the relaxation of the TFP-tax shock orthogonality condition is overidentified and supported by the data with a p-value of 0.89.

Also in this case, a note on the uncertainty surrounding our tax multiplier estimates conditional on the non-orthogonality case is in order. While the empirical analysis points to a non-zero correlation between the TFP proxy and the tax shock, we observe that the uncertainty surrounding the point estimates of the tax multiplier is much larger in this case than when orthogonality is imposed. In this respect, the estimated tax multiplier appears less robust than the estimated fiscal spending multiplier.

The higher uncertainty surrounding the estimated tax multiplier relative to the

fiscal spending multiplier is not surprising. Chahrour, Schmitt-Grohe, and Uribe (2012) generate artificial data from a real business cycle model featuring a number of exogenous shocks and real rigidities that have been shown to be important for fitting the US postwar business cycle. They report that for samples of size similar to the length of the postwar period, small sample issues can be substantial. In particular, given two distinct identification schemes that correctly recover the tax shock (Blanchard and Perotti (2002) vs. Romer and Romer (2010)), one cannot reject, at standard significance levels, the hypothesis that the observed differences in the estimated tax multipliers are due to small sample uncertainty.<sup>27</sup> Their evidence points out that even in theory-driven models, the variability associated with the tax multiplier is inherently large.

## 4 Alternative models

**Monetary policy.** Our baseline model is a fiscal policy-only model. Research on the fiscal-monetary policy mix shows that the output effects of fiscal shocks are importantly affected by the systematic monetary policy in place (see Leeper (1991) for an early theoretical contribution, Leeper and Leith (2016) for a more recent review, and Rossi and Zubairy (2011) for an empirical analysis). To control for the role of monetary policy, we work with an enriched model featuring also CPI inflation ( $\pi_t$ ) and the 3-month Treasury bill rate ( $r_t$ ). Hence, our vector of modeled variables becomes  $Y_t = (y_t, tr_t, g_t, \pi_t, r_t)'$ . We estimate this model by augmenting the set of instruments employed so far (AG, MR, and TFP) with one additional shock series: the measure of oil shocks (OIL) proposed by Hamilton (2003) as an instrument for the inflation shock, as done by Caldara and Kamps (2017). Hence, we estimate one additional model, where  $W_t = (Y_t', Z_t')' = (y_t, tr_t, g_t, \pi_t, r_t, MR_t, AG_t, TFP_t, OIL_t)'$ . As in the previous sections, we study two different scenarios, one in which the TFP-tax shock orthogonality condition is imposed, and one in which it is relaxed.<sup>28</sup>

<sup>27</sup>To illustrate, given the true (peak) tax multiplier of 1.78, Table 2 in Chahrour, Schmitt-Grohe, and Uribe (2012) shows that in samples of  $T = 250$  quarters, and conditional on the correct identification of tax innovations, the 68% confidence interval is (0.65, 2.57) when the tax shock is identified by the Blanchard and Perotti's (2002) method, and is (-0.09, 3.55) when regressions a la Romer and Romer (2010) are used. Interestingly, they show that in these samples the tax multiplier estimated with one method can be larger or smaller than the tax multiplier estimated with the other method with almost equal probability.

<sup>28</sup>As already mentioned in the Introduction, in an additional check we add the monetary policy shocks series proposed by Romer and Romer (2004). We estimate two more proxy-SVAR models, one with  $W_t = (Y_t', Z_t')' = (y_t, tr_t, g_t, \pi_t, r_t, MR_t, AG_t, TFP_t, RR_t)'$ , and one with  $W_t = (Y_t', Z_t')' = (y_t, tr_t, g_t, \pi_t, r_t, MR_t, AG_t, TFP_t, OIL_t, RR_t)'$ . We confine to the Appendix F the results of these

Figure 3 shows the estimated multipliers in these two scenarios for each of the two cases. As before, the estimated fiscal spending multiplier is insensitive to the treatment of the orthogonality condition, and peaks at a value equal to 1.8 regardless of the assumption on the TFP-tax shock orthogonality. The uncertainty around these estimates is also relatively low, with confidence intervals ranging from 1.5 to 2.3. Quite differently, the peak of the tax multiplier varies substantially: it is equal to 1.1 when the orthogonality condition is imposed, while it jumps to 3.1 when the orthogonality condition is not imposed. As in the baseline case, the uncertainty surrounding the estimated tax multipliers is again larger compared with that of the spending multiplier. It is important to stress that this latter model, which is again overidentified, is supported by the overidentification restrictions test which delivers a p-value of 0.99. For this proxy-SVAR, the estimated coefficient for the relevance of TFP proxy as an instrument for the output shock is  $\hat{\phi}_1 = 1.69$  and implies a correlation with the output shock of 57.4%, while the estimated coefficient for the relevance of TFP as an instrument for the tax shock is  $\hat{\phi}_2 = -0.64$  (with associated confidence interval equal to  $(-0.98, -0.46)$ ) and implies a correlations with the tax shock of -21.7%. Overall, these empirical results tend to confirm those documented in the previous sections with a more parsimonious VAR.

**Fiscal foresight.** Anticipation effects are likely to be of great relevance for the identification and transmission of fiscal policy shocks. This phenomenon, often referred to as ‘fiscal foresight’, makes SVAR analysis complicated. Standard SVARs, which rely on current and past shocks to interpret the dynamics of the modeled variables, can be ‘non-fundamental’, in that they do not embed the information related to ‘news shocks’, i.e., future shocks anticipated by rational agents. Leeper, Walker, and Yang (2013) work with different fiscal models and show that the anticipation of tax policy shocks severely affects VAR exercises aiming at identifying fiscal shocks. Ramey (2011) shows that government spending shocks estimated with standard fiscal SVARs are predictable, i.e., they are non-fundamental. Forni and Gambetti (2014) propose a test for ‘sufficient information’ to detect non-fundamentality that is based on checking the predictability of the VAR shocks of interest with information external to the VAR. We implement their test by regressing the identified fiscal shocks against lagged realizations of the factors extracted from the large set of macroeconomic and financial variables put together

---

checks since the monetary policy shock series is available from 1969Q1. As a consequence, the sample used to estimate the proxy-SVAR when we include this proxy is shorter (1969Q1-2006Q4), relative to the baseline model (1954Q1-2006Q4).

by McCracken and Ng (2016).<sup>29</sup> We use two sets of regressors: i) the first estimated factor, which explains about 55% of the variance of the data; ii) the first four factors, which explain almost 90%. Table 2 collects the p-values of the F-tests for information sufficiency we run over all our models. For each shock or combination of shocks, we consider two scenarios: a) an univariate scenario in which each fiscal shock is regressed over a constant and the estimated factors (first two rows of each shock/combination of shocks); b) a multivariate one in which the vector of fiscal shocks is regressed over constants and the estimated factors (last row of each shock/combination of shocks). All models pass the information sufficiency test.<sup>30</sup>

**Consumption.** A controversial issue concerns the response of private consumption to a fiscal spending shock.<sup>31</sup> As pointed out by Galí, López-Salido, and Vallés (2007), the standard RBC model predicts a negative response of consumption because households (whose decisions are based on an intertemporal budget constraint) associate the increase in public spending with a decrease in the discounted value of their after-tax income. Hence, given this negative wealth effect, households cut consumption, and the model tends to predict a multiplier lower than one. In contrast, the standard IS-LM model predicts an increase in consumption by non-Ricardian (Keynesian) households due the demand-driven increase in output, leading to a multiplier larger than one. The share of non-Ricardian households in OECD countries is sizeable, as it ranges from 20% to 35% as documented by Kaplan, Violante, and Weidner (2014). To account for non-Ricardian households, New-Keynesian models of the business cycle have then assumed the presence of rule-of-thumb consumers, whose consumption is a function of their current (as opposed to intertemporal) disposable income (Debortoli and Galí' (2018), Bilbiie (2020)). As shown by Galí, López-Salido, and Vallés (2007), adding rule-of-thumb consumers helps reconcile the theoretical predictions of micro-founded DSGE frameworks with the extant empirical evidence pointing to spending multipliers larger than one.

---

<sup>29</sup>To maximize the number of observations to compute the factors, we work with monthly data. We convert monthly factors in quarterly ones by taking the last realization of the factors in each quarter. Given that the factors are estimated with a sample starting in 1959, our regressions regard the sample 1959-2006.

<sup>30</sup>Canova and Sahneh (2018) note that Granger-causality tests might over-reject fundamentalness because of aggregation issues affecting the variables modeled with the VAR. The Forni and Gambetti (2014) tests we conducted over the different specifications of our VARs never reject fundamentalness. Hence, our VARs are not subject to the Canova-Sahneh critique.

<sup>31</sup>We thank an anonymous referee for suggesting to look into the response of consumption to a fiscal spending shock identified with our econometric strategy.

Our econometric framework enables us to document the response of private consumption to a public spending shock. We do so by adding private consumption to our SVAR and identifying such a shock via non-fiscal instruments (in our case, TFP). Given the robustness of our findings on the spending multiplier, we focus on the orthogonality case to keep the number of remaining zero restrictions at a minimum. Identification is achieved by imposing the following zero restrictions: (i) a spending shock has no contemporaneous effect on public revenues; (ii) shocks to real GDP and consumption do not affect public spending contemporaneously. Imposing these restrictions allows us to treat consumption symmetrically with respect to output, i.e., consumption and output are allowed to respond to all (fiscal and non-fiscal) shocks contemporaneously.

Figure 4 reports the response of consumption to a public spending shock and the output multiplier associated with this version of our proxy-SVAR. Two findings are worth highlighting. First, the response of consumption is positive, hump-shaped, persistent, and significant. If used to perform model selection, this impulse response points to the relevance of modeling Keynesian consumers and/or alternative mechanisms to generate a private consumption boost after an unexpected increase in public spending. Second, the dynamic evolution of the output multiplier is similar to the one previously documented with models without consumption. We take this evidence as supportive of the robustness of our findings concerning the size of the spending multiplier, which according to our proxy-SVAR analysis is larger than one.

## 5 Conclusions

This paper estimates US government spending and tax multipliers using a flexible proxy-SVAR approach which has two main advantages relative to other fiscal proxy SVARs: (i) multiple external instruments can be used, under certain conditions, to recover all (fiscal and non-fiscal) structural shocks of the system; (ii) when the orthogonality between the external instruments and the non-instrumented structural shocks does not hold, the impulse response functions of interest can be estimated consistently by exploiting a set of alternative (point-)restrictions that involve the impact that the non-instrumented structural shocks exert on the variables.

We estimate the peak fiscal spending multiplier to be about 1.6-2.1, no matter what the model specification and the set of fiscal and non fiscal instruments are. Differently, we find the peak tax multiplier to be 3.1 when a tax instrument only is employed, while its estimate drops to 0.7 when TFP is used as an instrument to estimate the

effects of output shocks, and the tax multiplier is then recovered via the moments associated to the covariance matrix of the VAR residuals. We show that these different estimates, which replicate those obtained by key contributions in the literature, are due to the imposition of the TFP-tax shock orthogonality condition when TFP is used as an instrument for the output shock. When we relax such assumption (imposing non-binding restrictions elsewhere in the proxy-SVAR), we find a peak tax multiplier that ranges from 2.8 to 3.6 across a set of proxy-SVARs, with 3.1 being our preferred estimate. Crucially, we show that the data do not speak loudly on the validity/failure of the TFP-tax shock orthogonality condition and that the possible connection between these two variables can be motivated by theoretical arguments. In line with what observed in large part of the literature, our tax multipliers tend to be surrounded by larger statistical uncertainty relative to what we document for the fiscal spending multiplier. These findings are robust to the joint use of fiscal and non-fiscal instruments, and to enlarging the system to account for the role of monetary policy.

From a modeling standpoint, our estimates confirm the positive relationship between changes in the output-tax elasticity and variations in the tax multiplier previously detected via counterfactual simulations by Mertens and Ravn (2014) and analytically worked out by Caldara and Kamps (2017). Policy-wise, our paper unveils a trade-off fiscal policymakers might have to face when designing their fiscal plans. On the one hand, our point estimates deliver a tax multiplier larger than the spending multiplier. On the other hand, the former is surrounded by a larger statistical uncertainty. Hence, policymakers with an aversion towards parameter uncertainty may want to assign a larger weight to the fiscal spending lever than to taxes.

A recent strand of the literature exploits the heteroskedasticity found in the data and/or the non-normality of VAR innovations to identify fiscal multipliers; see Lewis (2021) and references therein. Combining the suggested proxy-SVAR approach with the information provided by shifts in the volatility of the data induced by changes in macroeconomic regimes may help to strengthen the identification of fiscal multipliers and possibly to reduce the statistical uncertainty surrounding their estimates. This topic will be the natural continuation of our research agenda.

## References

- ANGELINI, G., AND L. FANELLI (2019): “Exogenous uncertainty and the identification of Structural Vector Autoregressions with external instruments,” *Journal of Applied Econometrics*, 34(6), 951–971.
- AUERBACH, A., AND Y. GORODNICHENKO (2012): “Measuring the Output Responses to Fiscal Policy,” *American Economic Journal: Economic Policy*, 4(2), 1–27.
- BACHMANN, R., AND E. SIMS (2012): “Confidence and the transmission of government spending shocks,” *Journal of Monetary Economics*, 59(3), 235–249.
- BILBIIE, F. O. (2020): “The New Keynesian Cross,” *Journal of Monetary Economics*, 114, 90–108.
- BLANCHARD, O., AND R. 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.
- BRÜGGEMANN, R., C. JENTSCH, AND C. TRENKLER (2016): “Inference in VARs with conditional volatility of unknown form,” *Journal of Econometrics*, 191, 69–85.
- CAGGIANO, G., E. CASTELNUOVO, V. COLOMBO, AND G. NODARI (2015): “Estimating Fiscal Multipliers: News From a Nonlinear World,” *Economic Journal*, 125(584), 746–776.
- CALDARA, D., AND C. KAMPS (2017): “The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers,” *Review of Economic Studies*, 84, 1015–1040.
- CANOVA, F., AND E. PAPPA (2007): “Price differentials in monetary unions: The role of fiscal shocks,” *Economic Journal*, 117, 713–737.
- CANOVA, F., AND M. H. SAHNEH (2018): “Are Small-Scale SVARs Useful for Business Cycle Analysis? Revisiting Nonfundamentalness,” *Journal of the European Economic Association*, 6(4), 1069–1093.
- CHAHROUR, R., S. SCHMITT-GROHE, AND M. URIBE (2012): “A Model-Based Evaluation of the Debate on the Size of the Tax Multiplier,” *American Economic Journal: Economic Policy*, 4(2), 28–45.
- CHRISTIANO, L. J., M. EICHENBAUM, AND S. REBELO (2011): “When is the Government Spending Multiplier Large?,” *Journal of Political Economy*, 119(1), 78–121.
- DEBORTOLI, D., AND J. GALI’ (2018): “Monetary Policy with Heterogeneous Agents: Insights from TANK models,” available at <https://crei.cat/people/davide-debortoli-2/>.
- FERNALD, J. (2014): “A Quarterly, Utilization-Adjusted Series on Total Factor Productivity,” Federal Reserve Bank of San Francisco Working Paper No. 2012-19.
- FOLLETTE, G., AND B. F. LUTZ (2010): “Fiscal Policy in the United States: Automatic Stabilizers, Discretionary Fiscal Policy Actions, and the Economy,” Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 2010-43.

- FORNI, M., AND L. GAMBETTI (2014): “Sufficient information in structural VARs,” *Journal of Monetary Economics*, 66, 124–136.
- GALÍ, J., J. LÓPEZ-SALIDO, AND J. VALLÉS (2007): “Understanding the Effects of Government Spending on Consumption,” *Journal of the European Economic Association*, 5, 227–270.
- GHASSIBE, M., AND F. ZANETTI (2019): “State Dependence of Fiscal Multipliers: The Source of Fluctuations Matters,” University of Oxford, mimeo.
- GIORNO, C., P. RICHARDSON, D. ROSEVEARE, AND P. VAN DEN NOORD (1995): “Estimating Potential Output, Output Gaps, and Structural Budget Balances,” OECD Economics Department Working Paper No. 152.
- GREGORY, A. W., J. MCNEIL, AND G. W. SMITH (2021): “US Tax and Spending Shocks 1950-2019: SVAR Overidentification with External Instruments,” Working Paper No. 1461, Economics Department, Queen’s University.
- HAMILTON, J. D. (2003): “What Is an Oil Shock?,” *Journal of Econometrics*, 113, 363–398.
- HUSSAIN, S. M. (2015): “The contractionary effects of tax shocks on productivity: An empirical and theoretical analysis,” *Journal of Macroeconomics*, 43, 93–107.
- INTERNATIONAL MONETARY FUND (2020): “World Economic Outlook: The Great Lockdown,” April.
- JENTSCH, C., AND C. K. LUNSFORD (2019a): “Asymptotically Valid Bootstrap Inference for Proxy SVARs,” *Journal of Business and Economic Statistics*, forthcoming.
- JENTSCH, C., AND K. G. LUNSFORD (2019b): “The dynamic effects of personal and corporate income tax changes in the United States: Comment,” *American Economic Review*, forthcoming.
- JUSTINIANO, A., G. E. PRIMICERI, AND A. TAMBALOTTI (2011): “Investment shocks and the relative price of investment,” *Review of Economic Dynamics*, 14(1), 101–121.
- KAPLAN, G., G. L. VIOLANTE, AND J. WEIDNER (2014): “The Wealthy Hand-to-Mouth,” *Brookings Papers on Economic Activity*, Spring, 77–138.
- KLEIN, M., AND L. LINNEMANN (2019): “Macroeconomic effects of government spending: the Great Recession was (really) different,” *Journal of Money, Credit and Banking*, 51, 1237–1264.
- LEEPER, E. (1991): “Equilibria Under ‘Active’ and ‘Passive’ Monetary and Fiscal Policies,” *Journal of Monetary Economics*, 27, 129–147.
- LEEPER, E. M., AND C. LEITH (2016): “Understanding Inflation as a Joint Monetary-Fiscal Phenomenon,” in: J.B. Taylor and H. Uhlig (Eds.): *Handbook of Macroeconomics*, Elsevier Science, forthcoming.
- LEEPER, E. M., C. SIMS, AND T. ZHA (1996): “What Does Monetary Policy Do?,” *Brookings Papers on Economic Activity*, 27, 1–78.



- LEEPER, E. M., N. TRAUM, AND T. B. WALKER (2017): “Clearing Up the Fiscal Multiplier Morass,” *American Economic Review*, 107(8), 2409–2454.
- LEEPER, E. M., T. B. WALKER, AND S.-C. S. YANG (2013): “Fiscal Foresight and Information Flows,” *Econometrica*, 81(3), 1115–1145.
- LEWIS, D. (2021): “Identifying Shocks via Time-Varying Volatility,” *Review of Economic Studies*, forthcoming.
- LÜTKEPOHL, H. (2005): “New Introduction to Multiple Time Series Analysis,” Springer-Verlag, Berlin and Heidelberg.
- MCCRACKEN, M. W., AND S. NG (2016): “FRED-MD: A Monthly Database for Macroeconomic Research,” *Journal of Business and Economic Statistics*, 34(4), 574–589.
- MCNEIL, J. (2022): “Over-identifying structural vector autoregressions with scarce external instruments,” mimeo.
- MERTENS, K., AND J. L. MONTIEL OLEA (2018): “Marginal Tax Rates and Income: New Time Series Evidence,” *Quarterly Journal of Economics*, 133(4), 1803–1884.
- MERTENS, K., AND M. RAVN (2019): “The dynamic effects of personal and corporate income tax changes in the United States: Reply to Jentsch and Lunsford,” *American Economic Review*, forthcoming.
- MERTENS, K., AND M. O. RAVN (2011a): “Technology-hours redux: Tax changes and the measurement of technology,” NBER International Seminar of Macroeconomics, 41–76.
- (2011b): “Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks,” *Review of Economic Dynamics*, 14(1), 27–54.
- (2012): “Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks,” *American Economic Journal: Economic Policy*, 4(2), 214–237.
- (2013): “The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States,” *American Economic Review*, 103(4), 1212–1247.
- (2014): “A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers,” *Journal of Monetary Economics*, 68, S1–S19.
- MONTIEL OLEA, J., J. STOCK, AND M. WATSON (2020): “Inference in SVARs identified with an external instrument,” *Journal of Econometrics*, forthcoming.
- MOUNTFORD, A., AND H. UHLIG (2009): “What Are the Effects of Fiscal Policy Shocks?,” *Journal of Applied Econometrics*, 24, 960–992.
- RAMEY, V. A. (2011): “Identifying Government Spending Shocks: It’s All in the Timing,” *Quarterly Journal of Economics*, 126(1), 1–50.
- (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, V. A., AND S. ZUBAIRY (2018): “Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data,” *Journal of Political Economy*, 126(2), 850–901.
- ROMER, C., AND D. ROMER (2004): “A New Measure of Monetary Policy Shocks: Derivation and Implications,” *American Economic Review*, 94(4), 1055–1084.
- ROMER, C. D., AND D. 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.
- ROSSI, B., AND S. ZUBAIRY (2011): “What Is the Importance of Monetary and Fiscal Shocks in Explaining U.S. Macroeconomic Fluctuations?,” *Journal of Money, Credit and Banking*, 43(6), 1247–1270.
- SIMS, E., AND J. WOLFF (2018): “The State-Dependent Effects of Tax Shocks,” *European Economic Review*, 107, 57–85.
- STOCK, J., AND M. W. WATSON (2018): “Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments,” *Economic Journal*, 128, 917–948.
- STOCK, J. H., AND M. W. WATSON (2012): “Disentangling the Channels of the 2007-2009 Recession,” *Brookings Papers on Economic Activity*, Spring, 81–135.
- WIELAND, J. F. (2018): “State-dependence of the Zero Lower Bound Government Spending Multiplier,” available at <https://sites.google.com/site/johannesfwieland/>.

| <b>Instruments</b>              | $\psi_y^g$                   | $\psi_y^{tr}$             | $\mathcal{M}g$            | $\mathcal{M}tr$            |
|---------------------------------|------------------------------|---------------------------|---------------------------|----------------------------|
| AG only                         | −0.0029<br>(−0.0275;0.0245)  | —                         | 1.6531<br>(1.1518;2.0546) | —                          |
| MR only                         | —                            | 3.3615<br>(2.2459;4.4506) | —                         | 3.0863<br>(1.4182;4.8065)  |
| TFP only - orth.                | −0.1434<br>(−0.2446;−0.0441) | 2.1142<br>(1.8285;2.4671) | 1.9134<br>(1.2678;2.3752) | 0.7583<br>(−0.0015;0.9313) |
| TFP only - non orth.            | −0.3430<br>(−0.4398;−0.1192) | 3.8566<br>(2.3135;4.9939) | 2.1842<br>(1.3902;2.5508) | 3.5831<br>(0.2393;5.8781)  |
| AG & MR & TFP - orth.           | −0.0053<br>(−0.0295;0.0214)  | 2.3115<br>(2.0936;2.6435) | 1.7885<br>(1.3421;2.2185) | 1.0409<br>(0.3851;1.2642)  |
| AG & MR & TFP - non orth.       | −0.0052<br>(−0.0293;0.0220)  | 3.3487<br>(2.4437;4.3210) | 1.7826<br>(1.2885;2.1725) | 2.8299<br>(0.3795;4.2754)  |
| AG & MR & TFP & OIL - orth.     | −0.0175<br>(−0.0521;−0.0029) | 2.6225<br>(1.7481;3.2631) | 1.8062<br>(1.5118;2.3833) | 1.0586<br>(0.2510;1.4164)  |
| AG & MR & TFP & OIL - non orth. | −0.0174<br>(−0.0497;0.0014)  | 3.6022<br>(2.5275;4.8102) | 1.7982<br>(1.4892;2.3558) | 3.1246<br>(0.6505;4.9533)  |

Table 1: **Estimated elasticities and multipliers: Linearly detrended data.**  
 Bootstrapped (16th,84th) percentiles below point estimates based on 1,000 repetitions  
 and the Moving Block-Bootstrap method. Multipliers: Peak values.

| Instruments                     | Shocks                     | $F_t = (F_{1,t})$ | $F_t = (F_{1,t}, F_{2,t}, F_{3,t}, F_{4,t})$ |
|---------------------------------|----------------------------|-------------------|--|
| AG only                         | $\hat{\varepsilon}_t^{tr}$ | 0.3612            | 0.2484                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.8156            | 0.6457                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.6922            | 0.4343                                       |
| MR only                         | $\hat{\varepsilon}_t^{tr}$ | 0.1414            | 0.1326                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.8028            | 0.7641                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.3719            | 0.3248                                       |
| TFP only - orth.                | $\hat{\varepsilon}_t^{tr}$ | 0.3600            | 0.2697                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.8942            | 0.5046                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.6843            | 0.4088                                       |
| TFP only - non orth.            | $\hat{\varepsilon}_t^{tr}$ | 0.0250            | 0.1128                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.6242            | 0.3856                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.1298            | 0.1983                                       |
| AG & MR & TFP - orth.           | $\hat{\varepsilon}_t^{tr}$ | 0.4615            | 0.2487                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.8104            | 0.6452                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.7619            | 0.4432                                       |
| AG & MR & TFP - non orth.       | $\hat{\varepsilon}_t^{tr}$ | 0.1293            | 0.1598                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.8107            | 0.6451                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.3808            | 0.3287                                       |
| AG & MR & TFP & OIL - orth.     | $\hat{\varepsilon}_t^{tr}$ | 0.9990            | 0.5354                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.3827            | 0.3207                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.6597            | 0.4623                                       |
| AG & MR & TFP & OIL - non orth. | $\hat{\varepsilon}_t^{tr}$ | 0.1289            | 0.3608                                       |
|                                 | $\hat{\varepsilon}_t^g$    | 0.3208            | 0.3826                                       |
|                                 | $\hat{\varepsilon}_t^y$    | 0.2190            | 0.3959                                       |

Table 2: **Informational sufficiency: Forni and Gambetti (2014) test.** P-values of F-tests reported in the Table. Per each shock or combination of shocks, we consider two scenarios: a) each fiscal shock regressed over a constant and the estimated factors (first two rows of each shock/combination of shocks); b) the vector of fiscal shocks regressed over constants and the estimated factors (last row of each shock/combination of shocks). Two lags of the factors included in all cases.

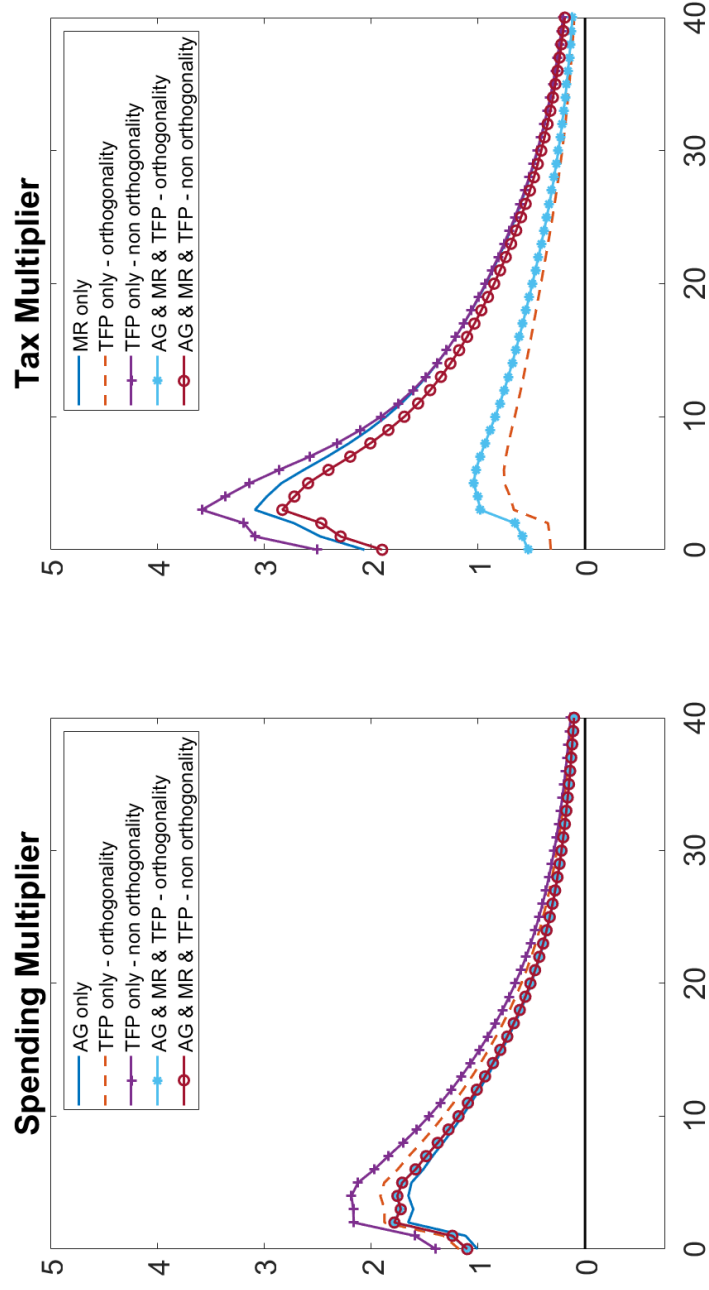


Figure 1: **Fiscal multipliers: Role of instruments.** Model with fiscal spending, taxes, output. AG: Auerbach and Gorodnichenko's (2012) fiscal spending instrument. MR: Mertens and Ravn's (2012) tax spending instrument. TFP: Fernald's (2014) TFP instrument.

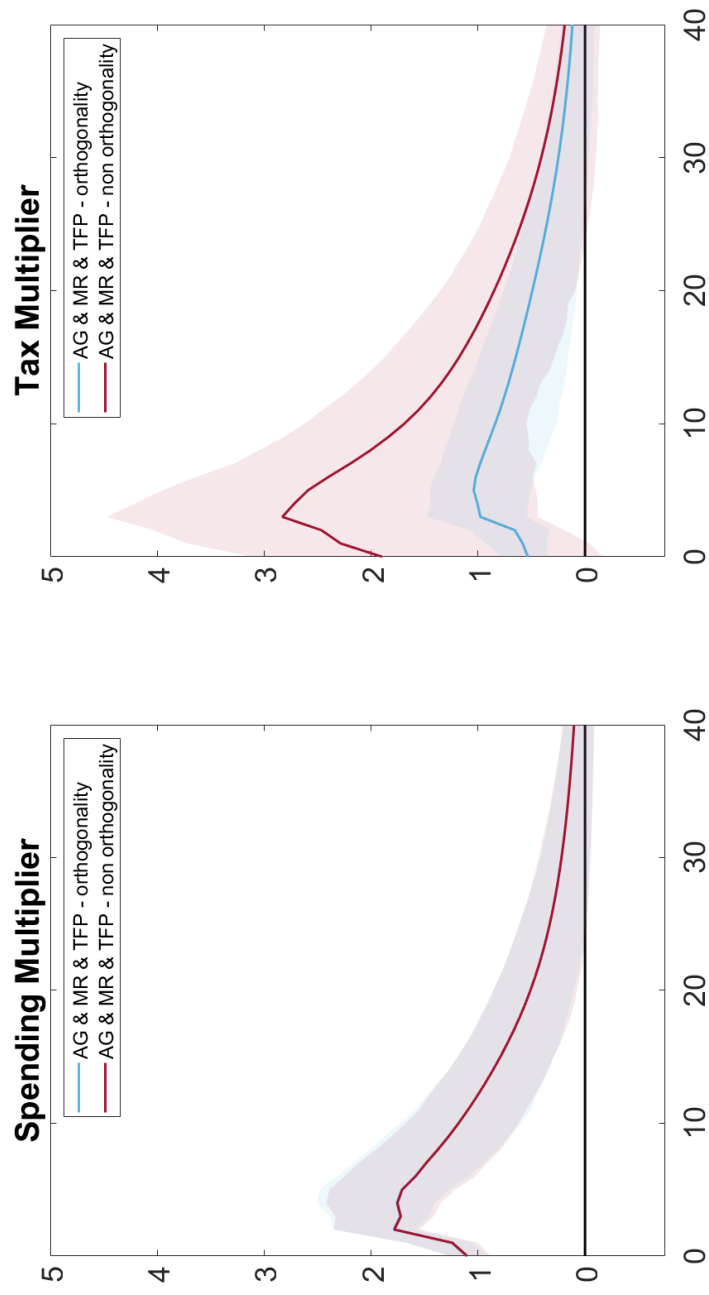


Figure 2: **Fiscal multipliers with alternative sets of instruments: Statistical difference.** Shaded areas: 68%-MBB confidence bands, see Jentsch and Lunsford (2019b). AG: Auerbach and Gorodnichenko's (2012) fiscal spending instrument. MR: Mertens and Ravn's (2012) tax spending instrument. TFP: Fernald's (2014) TFP instrument.

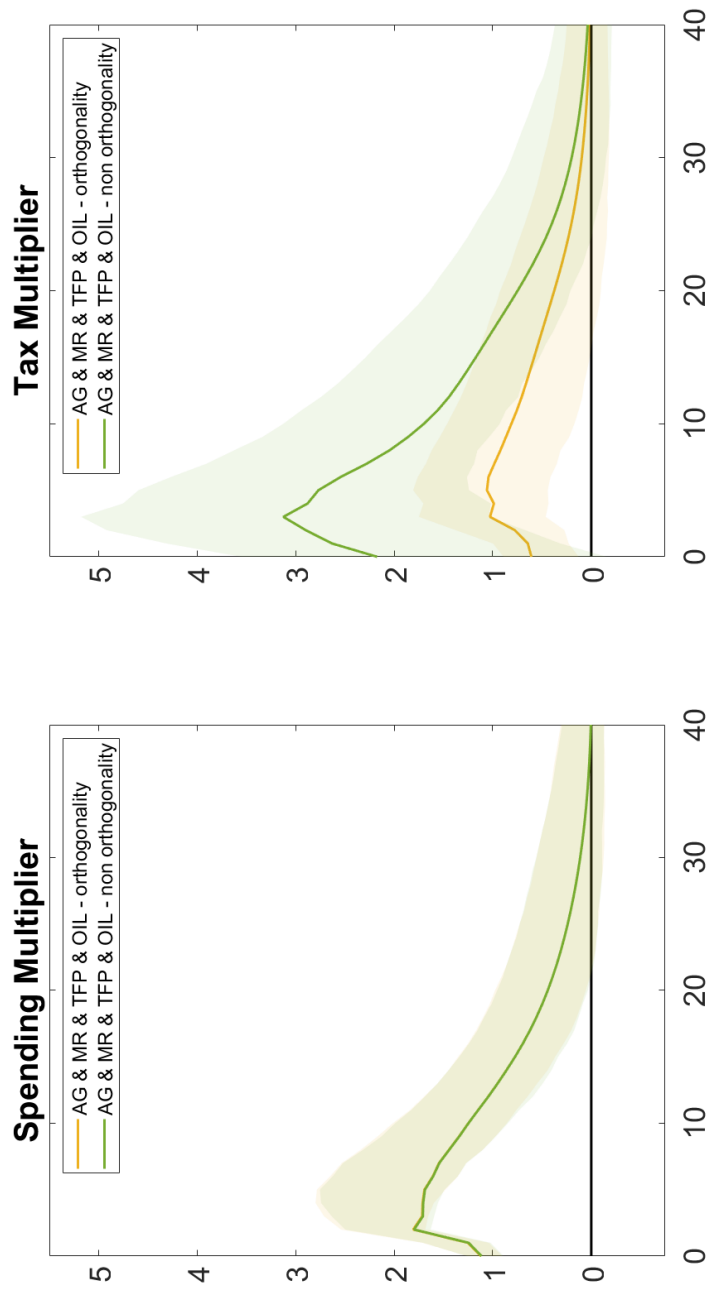


Figure 3: **Fiscal multipliers: Model with Monetary Policy.** Model with fiscal spending, taxes, output, inflation, 3-month Treasury bill rate. Shaded areas: 68%-MBB confidence bands, see Jentsch and Lunsford (2019b). AG: Auerbach and Gorodnichenko's (2012) fiscal spending instrument. MR: Mertens and Ravn's (2012) tax spending instrument. TFP: Fernald's (2014) TFP instrument. OIL: Hamilton's (2003) instrument.

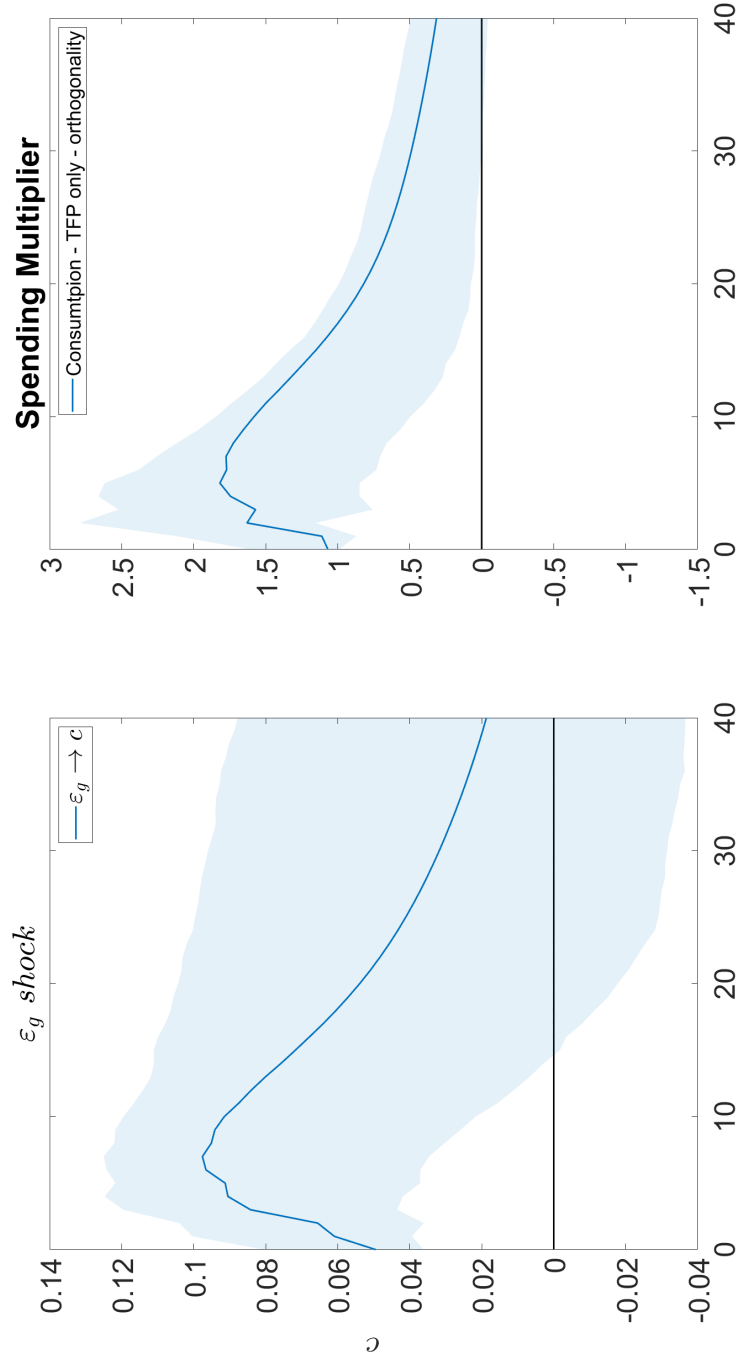


Figure 4: **Fiscal multipliers: Model with Monetary Policy.** Model with fiscal spending, taxes, output, inflation, 3-month Treasury bill rate. Shaded areas: 68%-MBB confidence bands, see Jentsch and Lunsford (2019b). AG: Auerbach and Gorodnichenko's (2012) fiscal spending instrument. MR: Mertens and Ravn's (2012) tax spending instrument. TFP: Fernald's (2014) TFP instrument. OIL: Hamilton's (2003) instrument.