

# US Fiscal Policy Shocks: Proxy-SVAR Overidentification via GMM

Allan W. Gregory, James McNeil, and Gregor W. Smith<sup>†</sup>

April 2022

---

## Abstract

An SVAR in US federal spending, federal revenue, and GDP is a standard setting for the study of the impact of fiscal shocks. An appealing feature of identifying a fiscal shock with an external instrument (proxy variable) is that one can find the effects of that shock without fully identifying the SVAR. But we show that fully or almost fully instrumenting the SVAR allows one to overidentify the model by incorporating the condition that the structural shocks are uncorrelated (via GMM). Over 1948–2019 the overidentifying restrictions are not rejected. The overidentified SVAR yields (a) greater precision in estimating impulse response functions and multipliers and (b) measures of the effects of output shocks even when there is no instrument for them.

---

*JEL classification:* E62, C36.

*Keywords:* structural vector autoregression, fiscal policy, external instruments

<sup>†</sup>Gregory and Smith: Department of Economics, Queen's University, Canada; gregorya@queensu.ca, gregor.smith@queensu.ca. McNeil: Department of Economics, Dalhousie University; mcneilj@dal.ca. We thank the Social Sciences and Humanities Research Council of Canada for support of this research. For very helpful comments we thank co-editor Marco Del Negro and four referees of this journal as well as Luca Fanelli and participants at the IAAE 2021 Annual Conference. We thank colleagues at the Federal Reserve Bank of New York for providing output from the FRBNY DSGE model.

## 1. Introduction

Over the past 20 years, a variety of studies have measured the effects of US fiscal shocks in a structural vector autoregression (SVAR) that comprises federal spending, federal revenue, and GDP (labelled  $\{g_t, \tau_t, y_t\}$  below). Ramey (2011b, 2019) surveys this research. A central concern in this work is to measure the multipliers from shocks to government spending or revenue, so that economists can better predict the effects of these shocks and perhaps also contribute to the design of effective policy in recessions.

Recent studies often identify the SVAR using external instruments, also known as proxy variables. This method was introduced by Stock (2008), and has been used in studies by Stock and Watson (2012), Mertens and Ravn (2013), Gertler and Karadi (2015), Caldara and Kamps (2017), and others. Stock and Watson (2018) show that working with external instruments yields an appealing property: A single instrument for the structural shock to  $g_t$ , for example, can identify the associated impulse response function (IRF).

However, researchers separately identifying features of the SVAR in this way might produce shocks that are correlated, a characteristic that seems incompatible with the definition of a shock. Ramey (2016, p 75) argues that each identified shock “should be uncorrelated with other exogenous shocks; otherwise, we cannot identify the unique causal effects of one exogenous shock relative to another.” Stock and Watson (2018, p 922) write that the “assumption that the structural shocks are mutually uncorrelated accords both with their interpretation as randomly assigned treatments and with their being primitive economic forces.”

We study the fully identified fiscal SVAR by instrumenting each shock and so measure their correlations. We find correlations even when instruments are strong and uncorrelated with other economic shocks and when the sample size spans many decades. We then show how this SVAR-IV procedure can be modified to include the conditions that the shocks are uncorrelated (which we refer to as the covariance restrictions) and show that estimation can proceed by the Generalized Method of Moments (GMM). The covariance restrictions overidentify the SVAR (or can be used to just identify it in the absence of enough external instruments). They thus allow a  $J$ -test, which may be useful as an indicator of a missing variable, or some other misspecification.

Section 2 outlines the method while section 3 then sets it in context, by relating it to the methods of Mertens and Ravn (2013), Angelini and Fanelli (2019), and Angelini *et al* (2020). Section 4 then studies the fiscal SVAR. In this application the  $J$ -test finds that the overidentifying restrictions cannot be rejected. An additional, important advantage is that the estimates from the overidentified SVAR are more efficient than those from the SVAR-IV method. In our application, standard errors on the shock-impact estimates fall by an average of 30%.

Section 5 discusses alternate instruments and, most importantly, shows that the covariance restrictions allow one to identify the effects of an output shock, for example, even without an instrument for that shock. Section 6 provides a Monte Carlo exercise which shows that estimated shocks can be correlated in the just-identified, fully instrumented SVAR, simply because of sampling variability. It also shows that overidentification with the covariance restrictions reduces those sample correlations, provides a  $J$ -test that is properly sized, and adds to efficiency. Section 7 concludes.

## 2. Estimators

Let  $x_t$  be a vector of  $N$  variables, indexed by  $n$ . They follow a  $p$ th-order VAR:

$$x_t = \sum_{i=1}^p B_i x_{t-i} + D w_t + u_t, \quad u_t \sim IID(0, \Sigma), \quad (1)$$

where  $w_t$  is a vector of deterministic terms,  $D$  and  $B_i$  are parameters, and  $\Sigma$  is positive definite. Residuals  $u_t$  are related to structural shocks  $\epsilon_t$  like this:

$$u_t = \Theta \epsilon_t, \quad \epsilon_t \sim IID(0, \Omega), \quad (2)$$

so that  $\Theta$  is an  $N \times N$  matrix. It is assumed to be nonsingular. The diagonal elements of  $\Theta$  are normalized to one for parameter identification. The off-diagonal elements are the parameters of interest because they determine the impact effect of the structural shocks on the variables  $x_t$ .

The variance-covariance matrix of the structural shocks,  $\Omega$ , is diagonal, reflecting the fact that structural shocks are orthogonal. The shock  $\epsilon_{n,t}$  for example, has variance  $\sigma_n^2$ , as

in Stock and Watson (2018). This is equivalent to assigning a unit variance to each shock and adjusting  $\Theta$  accordingly (see Mertens and Ravn, 2013).

Consider a set of  $N$  external instruments  $z_t$  corresponding to the  $N$  elements of  $x_t$ . For an instrument to be valid we require relevance and exclusion restrictions:

$$E(z_{n,t}\epsilon_{n,t}) = \alpha_n \neq 0 \quad (3a)$$

$$E(z_{n,t}\epsilon_{m,t}) = 0, \quad \text{for } m \neq n, \quad (3b)$$

so that a given instrument is correlated with a specific shock and uncorrelated with the other shocks. With these restrictions the instruments are sufficient to identify the off-diagonal elements of  $\Theta$ , and consistent estimation proceeds by IV as shown by Mertens and Ravn (2013) and Stock and Watson (2018). Each instrument  $z_{n,t}$  identifies a column of  $\Theta$  using:

$$E[(\hat{u}_{m,t} - \Theta_{m,n}\hat{u}_{n,t})z_{n,t}] = 0 \quad \text{for } m \neq n, \quad (4)$$

which gives  $N(N-1)$  moments that just identify the off-diagonal elements of  $\Theta$ . We refer to this as the just-identified estimator. Thus with a valid instrument for each shock the SVAR can be identified, and the shocks measured as:

$$\hat{\epsilon}_t = \hat{\Theta}^{-1}\hat{u}_t. \quad (5)$$

As the introduction noted, the shocks  $\hat{\epsilon}_t$ , measured in this way, can be correlated. But the IV estimator can be augmented by restricting the structural shocks to be uncorrelated, which we refer to as the covariance restrictions. Under that assumption the off-diagonal elements of  $\Omega$  are zero, or:

$$E[\epsilon_{m,t}\epsilon_{n,t}] = 0 \quad \text{for } m \neq n. \quad (6)$$

The shocks  $\hat{\epsilon}_t$ , defined in equation (5), depend on  $\Theta$  which is how the sample version of conditions (6) can provide additional information about the model parameters. We estimate the moment conditions (4) and (6) by GMM. We refer to this as the overidentified estimator.

Imposing the covariance restrictions (6) in estimation leads to  $N(N - 1)/2$  additional restrictions and a  $J$ -test with that many degrees of freedom. In our application the number of variables is  $N = 3$  and the VAR is fully instrumented. Thus the GMM estimator yields a  $J$ -test with 3 degrees of freedom.

We have so far discussed the case where there is one instrument per shock (or variable in the VAR) as in equations (3). However, Angelini and Fanelli (2019, Proposition 2, p 963) show that necessary and sufficient conditions for identification of  $\Theta$  hold when there is a shock with a missing instrument (again with one instrument for each of the other shocks). In that case, conditions (4) give  $(N - 1)^2$  restrictions (as each instrument identifies a column of  $\Theta$ ), while the covariance restrictions provide  $N(N - 1)/2$  restrictions as before. The matrix  $\Theta$  contains  $N(N - 1)$  unknown parameters, because the diagonal elements are normalized to one. Thus with a missing instrument there is still overidentification, and a  $J$ -test statistic with degrees of freedom:

$$(N - 1)^2 + \frac{N(N - 1)}{2} - N(N - 1) = \frac{(N - 1)(N - 2)}{2}. \quad (7)$$

The fiscal SVAR with  $N = 3$  thus is overidentified using the covariance restrictions even with only 2 instruments. In that case there are four IV moments (4) and three covariance restrictions (6) to identify six parameters (the off-diagonal elements of  $\Theta$ ), with a  $J$ -test with 1 degree of freedom. In our fiscal application the three instruments appear strong. But sub-section 5.2 reports on an example in which we drop an instrument and still can fully identify the SVAR and apply this  $J$ -test.

### 3. Related Studies

When the SVAR is fully identified, Angelini and Fanelli (2019) propose a maximum likelihood (ML) estimator comparable to our GMM approach. The two estimators differ for the usual reasons that GMM and ML differ. For example, our procedure does not require any distributional assumptions, which may be attractive to practitioners. In our fiscal policy application, the instruments for tax and government spending shocks are highly censored, containing many zero observations. In this case, one may be unsure of how to describe their distribution.

The full-shocks identification approach of Angelini and Fanelli (2019) involves estimating a constrained VAR, augmented to include the instrumental variables. Joint estimation of all the model parameters thus requires that the instruments and VAR variables are available over the same sample period. Because of the challenge of constructing valid instruments, there are many cases where instruments are available only over a shorter sample period. For example, it is now common to identify the effects of monetary policy shocks using Federal Funds futures contracts, which are available beginning only in 1988. Stock and Watson (2018) observe that when the instruments are not available over the full sample, it is still desirable to estimate the parameters from the reduced-form VAR (the  $B$  parameters in our notation) over the full sample period, which provides more efficient estimates of those parameters and hence impulse response functions. Because our procedure follows the usual SVAR-IV approach of separately estimating the parameters from the reduced-form VAR and the impact coefficients, we can readily accommodate different samples.

The GMM procedure is also distinct because the weights placed on the moment conditions are based on their covariability over the estimation period, so that the moments arising from the covariance restrictions are not treated as any more informative than the moments associated with the instrumental variables. By contrast, under the ML procedure the covariance restrictions will hold exactly in the absence of any additional zero restrictions on the  $\Theta$  matrix. If the necessary assumptions hold for the two approaches then both are consistent but they will differ in finite samples.

Mertens and Ravn (2013), and Angelini *et al* (2020) consider cases where more than one instrument is used to identify more than one shock but the assumption that each instrument is correlated with only one of the shocks is not satisfied. In this case additional information is required to identify the model, for example zero restrictions on the  $\Theta$  matrix, as Angelini and Fanelli (2019) outline in detail. We focus on the case where this assumption is satisfied and the SVAR can be fully identified with sufficient instruments.

The just-identified estimator is a natural extension of the typical SVAR-IV approach where a single shock is identified with a single instrument, which must be uncorrelated with all other (unidentified) shocks in the VAR. Because we identify all shocks in the

model, however, this assumption can also be studied jointly. As we show in sub-section 4.3, there is no evidence that any of the identified shocks in our fiscal VAR is correlated with any of the other instruments. Of course, this assumption may not be appropriate in all contexts. Mertens and Ravn (2013), for example, construct instruments for personal and corporate tax changes based on a narrative account of changes to US tax liabilities. They make the reasonable observation that both instruments are likely to be correlated with personal and corporate tax shocks, violating the exclusion restriction (3b), so they replace this assumption with a recursivity assumption.

To our knowledge, very few SVAR studies use a full set of instruments for identification. Angelini *et al* (2020, section 3.4) do use a full set. They study a fiscal SVAR for the US using an auxiliary model of the instruments like that in the ML approach of Angelini and Fanelli (2019). They find a significant correlation between  $\hat{\epsilon}_{\tau,t}$  and  $z_{y,t}$  (where below we do not). They then place restrictions on  $\Theta$  to allow identification and estimation when condition (3b) does not hold. They also use different instruments from ours and a sample ending in 2006 (for comparability with some previous studies), two features which may explain their different findings.

Fengler and Polivka (2021) estimate a multivariate volatility model with  $N - 1$  instruments by ML and show that  $N$  structural shocks are identified. Stock and Watson (2012) use many instruments to estimate many different shocks in a dynamic factor model and document sizeable correlations between many of their identified shocks, which they note may be because many of the instruments are weak and hence may not satisfy either relevance or exclusion restrictions. A number of studies have only a single instrument, perhaps because finding valid instruments involves painstaking work by researchers like those cited in section 4.

Montiel Olea, Stock, and Watson (2020) and Angelini, Cavaliere, and Fanelli (2021) provide tools for inference in proxy-SVARs when instruments are weak. But in the application here the instruments are strong, according to robust, first-stage  $F$ -tests, so we do not draw on their methods.

Several other studies have tested overidentification in SVARs. For example, Bernanke and Mihov (1998) test monetary SVARs overidentified with short-run restrictions. Lanne

and Luoto (2021) test restrictions implied by shocks that are uncorrelated but not necessarily independent, in other words restrictions involving higher-order moments. Guay and Normandin (2018) base identification on third and fourth unconditional moments of reduced-form residuals. Lewis (2020) bases identification on time-varying volatility. These last two studies also apply their methods to the  $\{g_t, \tau_t, y_t\}$  SVAR, so one could combine their methods with external instruments for further overidentification.

## 4. Fiscal SVAR

We next outline the fiscal SVAR, describe the instruments, estimate  $\Theta$ , and then graph multipliers. The focus is on comparing the findings from the just-identified and overidentified cases, so as to document the effect of adding the covariance restrictions.

### 4.1. Specification

Let  $x_t \equiv \{g_t, \tau_t, y_t\}'$  be a vector of quarterly US federal government spending, tax revenue, and GDP in logs of real dollars per capita. Our measurements are designed to follow those in previous studies. Following Mertens and Ravn (2014), government spending,  $g_t$ , is federal government expenditure and gross investment; federal revenue,  $\tau_t$ , is the sum of federal current tax receipts, social insurance contributions, and corporate income taxes; and output,  $y_t$ , is Gross Domestic Product (GDP). All variables are in logarithms after being deflated by the GDP deflator and expressed in per-capita terms. The sample period is 1948–2019.

The vector of deterministic terms,  $w_t$ , includes a constant, a quadratic time trend, and a dummy variable for 1975Q2, and four lags of that dummy variable. Blanchard and Perotti (2002) include this dummy variable so that they can compare the effects of temporary and permanent tax changes. The vector  $x_t$  then follows the VAR:

$$x_t = \sum_{i=1}^4 B_i x_{t-i} + D w_t + u_t, \quad u_t \sim IID(0, \Sigma), \quad (8)$$

where  $\Sigma$  is positive definite. The variable definitions, deterministic terms, and lag length ( $p = 4$  quarters) thus follow Mertens and Ravn (2014) and Lewis (2020) for comparability. This specification of the VAR also fits with other studies of fiscal policy using SVARs in



quarterly US data, including those of Blanchard and Perotti (2002), Perotti (2008), Ramey (2011a), Mertens and Ravn (2013), and Liu and Williams (2019).

We assume that  $\Theta$  is nonsingular so that the structural shocks  $\epsilon_t$  can be recovered from the VAR residuals  $u_t$ . This assumption would be violated under fiscal foresight, where some of the exogenous changes to government spending are anticipated in advance. Ramey (2011a) provides some evidence of this. One way to account for foresight is to include a series of expected government spending in the VAR. Unfortunately, such data from the *Survey of Professional Forecasters* begin only in 1981. To include this data considerably reduces our sample size and as a result both fiscal instruments are too weak to be informative. However, Perotti (2011) compares results from SVARs identified with zero restrictions with and without these expectations and finds very similar results. He also shows that professional forecasters do not predict actual government spending growth with much accuracy. Both of these findings suggest that spending shocks are not anticipated.

## 4.2. Instrument Relevance

In our application to US history, the instrument for government spending shocks,  $z_{g,t}$ , comes from Ramey and Zubairy (2018). This instrument is an updated version of the Ramey (2011a) shocks which are a series of changes in the expected present value of government purchases as a result of military buildups. Following Ramey we deflate the government spending shocks by the the previous period's nominal GDP. The instrument for tax shocks,  $z_{\tau,t}$ , comes from Mertens and Montiel Olea (2018) who construct a narrative series of exogenous tax changes for different tax brackets. We use their instrument which is the average tax change for the lowest 90% of the income distribution. We convert their annual series to quarterly by assigning the tax change to the quarter in which it took effect, in the same manner as Romer and Romer (2010) and Mertens and Ravn (2012).

To find an instrument  $z_{y,t}$  for shocks to GDP growth, we follow the example of Stock and Watson (2012). They suggest and illustrate using the productivity shocks from the Smets-Wouters model as an external instrument. We use the shocks from the FRBNY DSGE model which is the most prominent, ongoing DSGE model for the US. Del Negro *et al* (2017) and Cai *et al* (2019) provide descriptions and applications of the model.

Specifically, we use the posterior mean of the innovation to the permanent productivity shock in that model. We then filter this by regressing it on its own 4 lags and 4 lags of the VAR variables, to ensure it is an innovation in this environment. These three instruments are essentially uncorrelated with each other:  $\rho(z_{g,t}, z_{\tau,t}) = -0.013$ ,  $\rho(z_{g,t}, z_{y,t}) = 0.006$ , and  $\rho(z_{\tau,t}, z_{y,t}) = -0.028$ . And they have no significant autocorrelation.

The instruments are available for a smaller sample than the variables in the VAR, so our GMM estimation uses the shorter sample period. The reduced-form VAR parameters ( $B$  and  $D$ ) are estimated using data from the full 1948–2019 span to add efficiency—as recommended by Stock and Watson (2018, pp 923–933)—while the impact matrix  $\Theta$  is estimated using data from 1960–2012. We also use that sample (1960–2012) to calculate first-stage  $F$ -statistics.

As a check on the relevance of each instrument Andrews and Stock (2018) recommend using the heteroskedasticity-robust  $F$ -statistic, which coincides with the effective  $F$ -statistic of Montiel Olea and Pflueger (2013) when there is a single regressor, as is the case here. When this statistic is above 10, Montiel Olea, Stock, and Watson (2020) suggest that one can use standard methods for inference. The first-stage, robust  $F$ -statistics, equation-by-equation, for  $\{g_t, \tau_t, y_t\}$ , are: 9.94, 24.47, and 10.56. For the spending and revenue equations we thus confirm the first-stage findings of Ramey and Zubairy (2018) and Mertens and Montiel Olea (2014), so that condition (3a) holds. The FRBNY instrument  $z_{y,t}$  appears to be strong enough too.

### 4.3. $\hat{\Theta}$ and Shock Properties

We consider two identifications. First, we just-identify the SVAR using the three instruments  $z_t$  as in equations (4). Second, we also impose the restriction (6) that the shock covariances are zero. Estimates and standard errors are calculated by iterated GMM. (Implementing a parametric bootstrap is challenging because the fiscal instruments have many zero observations.) In the just-identified case, this is equivalent to equation-by-equation two-stage least squares. In the overidentified case, we use an identity weighting matrix to produce first-stage estimates and a Newey-West HAC covariance matrix with 4 lags for later stages. The covariance moments in general are nonlinear in the  $\Theta$  parameters

and we use the BFGS algorithm to solve the nonlinear optimization problem.

The orthogonality of structural shocks adds three overidentifying restrictions. With  $df = 3$  the test statistic is  $J = 0.03$  yielding a  $P$ -value of 0.99. Thus there is no evidence against the overidentifying restrictions. Our next goal is to document the effect of this overidentification on the economic findings. The upper two panels of Table 1 show the estimates  $\hat{\Theta}$  from the two identifications, along with standard errors. (The third panel is discussed in sub-section 5.2 below.)

**Table 1: Estimates  $\hat{\Theta}$**

IV			
	$g$	$\tau$	$y$
$g$	1.000	0.248 (0.159)	-0.433 (0.656)
$\tau$	-0.134 (0.382)	1.000	2.172 (1.006)
$y$	0.206 (0.119)	-0.052 (0.070)	1.000

Overidentified GMM			
	$g$	$\tau$	$y$
$g$	1.000	0.253 (0.144)	-0.500 (0.484)
$\tau$	-0.148 (0.205)	1.000	2.175 (0.547)
$y$	0.191 (0.078)	-0.054 (0.057)	1.000

Overidentified GMM Missing $z_{y,t}$			
	$g$	$\tau$	$y$
$g$	1.000	0.251 (0.148)	-0.596 (0.787)
$\tau$	-0.123 (0.283)	1.000	2.207 (0.637)
$y$	0.205 (0.115)	-0.051 (0.066)	1.000

Notes: Table 1 reports estimates  $\hat{\Theta}$  from  $u_t = \Theta \epsilon_t$ . In the first panel these are just-identified with instruments  $z_t$ . In the second panel they are overidentified with covariance restrictions. In the third panel  $z_{y,t}$  is dropped with the covariance restrictions again imposed.

The most notable effect of overidentification is the increase in precision. Moving from the IV estimation (in the first panel) to the overidentified estimation with covariance restrictions (in the second panel) the standard errors fall by an average of 30% with little change in the point estimates themselves. In the third row of each panel the coefficients  $\hat{\Theta}_{y,g}$  and  $\hat{\Theta}_{y,\tau}$  measure the impact effect on output of shocks to government spending and tax revenues respectively. These values begin the associated IRFs and multipliers. In Table 1, overidentification reduces the standard error on  $\hat{\Theta}_{y,g}$  by 35% and the standard error on  $\hat{\Theta}_{y,\tau}$  by 19%. Thus the efficiency gain for the IRF and multiplier for  $g$ -shocks will be greater than for  $\tau$ -shocks.

While our main focus is on these efficiency gains, one also can compare the estimates to those in other studies. For example, the estimates of  $\hat{\Theta}_{\tau,y}$  (with standard errors in brackets) are 2.172 (1.006) in the just-identified case and 2.175 (0.547) in the overidentified case. These point estimates fall within the range spanned by the estimates of Blanchard and Perotti (2002), who calibrate a value of 2.08, and Mertens and Ravn (2014), who estimate a value of 3.13. Caldara and Kamps (2017) and Lewis (2020) discuss the effects of this coefficient.

We next calculate the time series of the three structural shocks from  $\hat{\Theta}$  and the reduced-form shocks as in equation (5). (They are graphed in the online appendix.) As we noted earlier, there may be correlations between the shocks from the traditional identification that uses one external instrument to identify each shock. Here those values are  $\rho(\hat{\epsilon}_{g,t}, \hat{\epsilon}_{y,t}) = -0.13$ ,  $\rho(\hat{\epsilon}_{g,t}, \hat{\epsilon}_{\tau,t}) = .16$ , and  $\rho(\hat{\epsilon}_{\tau,t}, \hat{\epsilon}_{y,t}) = 0.02$ . Of course we can uncover these correlations only because we identify more than one shock: They would go undetected if one were studying a single instrument. In the GMM case these correlations are smaller but not exactly zero, for two reasons. First, there is overidentification so that the covariance restrictions need not hold exactly. Second, these correlations are from the entire sample: They are closer to zero during the 1960–2012 period for which all the instruments are available.

Table 2 shows the shock-instrument correlations from both our identifications. Note that the off-diagonal elements (such as  $\rho(\hat{\epsilon}_{g,t}, z_{\tau,t})$  or  $\rho(\hat{\epsilon}_{g,t}, z_{y,t})$ ) are close to zero, a result which is not a formal test but accords with the exclusion restrictions (3b). Angelini and

Fanelli (2019, Table 1) report similar statistics as a diagnostic in their SVAR. Lewis (2020) uses the fact that one of those correlations is not zero in a fiscal SVAR to question the validity of an instrument. But no such question is raised by the diagnostic in Table 2.

**Table 2: Shock-Instrument Correlations**

IV			
	$\hat{\epsilon}_{g,t}$	$\hat{\epsilon}_{\tau,t}$	$\hat{\epsilon}_{y,t}$
$z_{g,t}$	0.202	0.002	-0.001
$z_{\tau,t}$	0.005	0.299	0.001
$z_{y,t}$	0.001	-0.001	0.289

Overidentified GMM			
	$\hat{\epsilon}_{g,t}$	$\hat{\epsilon}_{\tau,t}$	$\hat{\epsilon}_{y,t}$
$z_{g,t}$	0.202	-0.001	0.006
$z_{\tau,t}$	0.005	0.299	0.003
$z_{y,t}$	0.007	0.002	0.289

Notes: Table 2 reports correlation coefficients between estimated structural shocks and instruments first from just-identified IV and then with covariance restrictions also imposed.

#### 4.4. Multipliers

Label the impulse response at horizon  $h$  as  $\psi_h$ . Then  $\psi_0 = \Theta$ , and the rest of the impulse response functions (IRFs) are, in recursive form:

$$\psi_h = \sum_{i=1}^{\min h,p} B_i \psi_{h-i}, \quad (9)$$

where the  $B_i$  are from equation (8) with  $B_i = 0$  for  $h \geq p$ , and with lag length  $p = 4$  in this application. The online appendix contains the 3x3 matrix of all IRFs.

To focus on the effects of fiscal-policy shocks, and for comparability with previous studies, we report the dynamic multipliers. Label  $\psi_{y,g,h}$  as the response of output to a shock to  $g$  after  $h$  quarters, and  $\psi_{y,\tau,h}$  as the corresponding measure for tax shocks. These multipliers measure the ratio of the  $y$ -response IRF to the impact response of the fiscal variable (here normalized to one), divided by the fiscal/GDP ratio over the sample period.

So, for government spending for example, define  $\overline{G/Y}$  as the average ratio of the level of government spending to output. Then the dynamic multiplier is:

$$\frac{\psi_{y,g,h}}{\overline{G/Y}}, \quad (10)$$

with a similar expression for the tax multiplier.

The upper panel of Figure 1 shows these multipliers for government spending (on the left) and taxes (on the right). The just-identified case is shown in black while the overidentified case is shown in red. Dashed lines show 68% asymptotic confidence intervals which are calculated using the delta method, the same method used by Mertens and Montiel Olea (2018) and Mertens and Ravn (2019). (Brüggemann, Jentsch, and Trenkler (2016) and Jentsch and Lunsford (2019) show that the wild bootstrap is not valid for structural IRFs.)

Figure 1 shows that the point estimates for the dynamic  $g$ -multipliers are slightly smaller in the overidentified case while for the  $\tau$ -multipliers they are very similar across the two estimators. The main effect of the overidentification is in precision, where the (red) confidence intervals for the  $g$ -multipliers in particular are notably narrower.

Our main aim is to compare these just-identified and overidentified cases, but the findings also can be compared to those in other studies. The largest values of the  $\tau$ -multipliers are at -1.2, which falls in between the values found by Blanchard and Perotti (2002) and Mertens and Ravn (2014). As noted in sub-section 4.3, this partly reflects the fact that our estimates  $\hat{\Theta}_{\tau,y}$  also lie between the values found in those studies. Our results are little changed if we end the sample in 2006, for comparison with Mertens and Ravn (2004) and Lewis (2020). The  $g$ -multipliers peak at values above 2, higher than values reviewed by Ramey (2019, table 1) or reported by Lewis (2020, Figure 4) using several methods. But the 68% confidence intervals generally include the point estimates found in other studies.

We next calculate the cumulative spending and tax multipliers, following the formulas of Mountford and Uhlig (2009), as discussed by Ramey (2019). In the case of government spending, this multiplier measures the ratio of the present value of the output response over time to the present value of the government spending shock over time. Define  $\bar{i}$  as

the average 3-month treasury bill rate over the sample. Then the cumulative multiplier is given by:

$$m_{y,g,h} = \frac{1}{G/Y} \frac{\sum_{j=0}^h (1 + \bar{i})^{-j} \psi_{y,g,h}}{\sum_{j=0}^h (1 + \bar{i})^{-j} \psi_{g,g,h}}. \quad (11)$$

The tax multiplier  $m_{y,\tau,h}$  is computed similarly.

The lower panel of Figure 1 shows the cumulative multipliers. Again the just-identified case is in black and the over-identified case is in red. Dashed lines show 68% confidence intervals. The overidentification leads to point estimates of the  $g$ -multipliers that are lower and more precisely estimated than in the just-identified case. For example at the 10-quarter horizon the just-identified multiplier is 1.49 with a standard error of 1.05 while the overidentified one is 1.37 with a standard error of 0.74. Thus the standard error again drops by 30%. As just noted for the dynamic multipliers, these values are somewhat higher than the cumulative multipliers reported by Ramey (2019, Table 1) though the confidence intervals include typical values like 1.

The cumulative  $\tau$ -multipliers are in the lower right panel of Figure 1. The efficiency gains from overidentification are small in this dimension, as for the dynamic multiplier above. For both identifications the cumulative  $\tau$ -multipliers tend to increase in scale over time, as Ramey (2019) noted is true of most tax multipliers. These values lie between those found using the methods of Blanchard and Perotti (2002) and Mertens and Ravn (2015) as reported by Lewis (2020, Table 2). Note that these multipliers include feedback from induced changes in output to tax revenues, which leads to large measures as explained by Ramey (2019, p 98 and Table 2).

We conclude this application with two observations. First, the historical decompositions implied by the two identifications are very similar, because the shocks and IRFs are themselves similar. Second, in the just-identified case the forecast error variance decompositions (FEVDs) may be misleading because the measured shocks may be correlated. For example, the FEVD of output at horizon zero is calculated as:

$$\frac{\hat{\Theta}_{y,n}^2 \hat{\sigma}_n^2}{\hat{\Sigma}_{y,y}} \quad (12)$$

for each of the  $n = g, \tau, y$  structural shocks. Because the forecast errors  $\hat{u}_t$  are just combinations of these structural shocks, together the shocks should explain all the variation

in the forecast errors of output and the sum of equation (12) over  $n$  should equal one. But this will not automatically hold in the just-identified case. However, the departure from one is not large in this application: The value of the sum is 1.11.

## 5. Alternate or Missing Instruments

This section reports on two variations on the fiscal SVAR. We first discuss alternate instruments. We then report on the effects of dropping an instrument (specifically  $z_{y,t}$ ).

### 5.1 Alternate Instruments

One might wonder about alternate instruments or whether further overidentification can be provided by additional instruments. While valid instruments typically are scarce in macroeconomic applications, several alternate instruments are available for this VAR. In the role of  $z_{\tau,t}$  we consider the instrument for unanticipated tax changes developed by Mertens and Ravn (2013) and extended by Liu and Williams (2019). We also consider the Mertens and Montiel Olea (2008) instrument that is the average tax change across all brackets. In both cases these are weaker than the baseline tax instrument of sub-section 4.2 and so might require different tools for inference.

In the role of  $z_{y,t}$  we consider the quarterly series of utilization-adjusted, total factor productivity constructed by Fernald (2014). This has a first-stage  $F$ -statistic similar to that of our baseline  $z_{y,t}$ . And we consider the posterior mean of the innovation to the temporary productivity shock in the FRBNY DSGE model (as opposed to the innovation in the permanent shock, which is used in our baseline case). This has a much larger first-stage  $F$ -statistic. However, in both these cases we find a larger correlation  $\rho(\hat{\epsilon}_{g,t}, z_{y,t})$  than shown for the baseline case in Table 2, which casts doubt on whether the exclusion restriction (3b) holds. And in both cases the  $J$ -test has a much smaller  $P$ -value (roughly 0.10) than in the baseline case, which may be detecting exactly this ineligibility of these candidate instruments.

### 5.2 A Missing Instrument

So far the main effect of the covariance restrictions has been to add precision to  $\hat{\Theta}$  and hence to estimates of IRFs and multipliers. But these restrictions also may allow



identification of each element of  $\Theta$  when an instrument is missing. For example, with 3 variables and 2 instruments, section 2 noted that the SVAR is overidentified, with one overidentifying restriction.

To illustrate this scenario we omit  $z_{y,t}$  and estimate the model using  $z_{g,t}$  and  $z_{\tau,t}$  along with the covariance restrictions. The instrument sample is again 1960–2012 to allow direct comparison with section 4. The  $J$ -test statistic is 0.002 with a  $P$ -value of 0.97, so there is no evidence against the model from this test. The third panel of Table 1 shows the estimates  $\hat{\Theta}$  and their standard errors.

There are three things to note. First, the key finding is that the final column of  $\Theta$  is identified even though  $z_{y,t}$  is absent. Valid instruments may be scarce in macroeconomic applications. In this example the covariance restrictions allow the researcher to study the causal effect of an output shock  $\epsilon_{y,t}$  without a known, valid instrument. Second, the point estimates  $\hat{\Theta}$  are quite similar to those from the fully-instrumented IV and GMM estimators. For that reason we do not graph IRFs and multipliers from this third estimator. Third, by comparison with the IV estimator we have dropped  $z_{y,t}$  but added the covariance restrictions. How this swap affects efficiency is an empirical question specific to the application. In this case the standard errors on average are 10% smaller in the third panel than in the first panel, where the SVAR is fully instrumented but the covariance restrictions are not imposed. Each standard error is smaller with the exception of that on  $\hat{\Theta}_{g,y}$ .

## 6. Monte Carlo

To study the properties of the estimators, we next consider a simulation environment in which the structural shocks are uncorrelated, the instruments are strong, and the sample size is realistically large. Here several of the differences between the just-identified and overidentified results that are found in the historical application can arise only due to sampling variability, for both estimators are consistent.

To illustrate the finite-sample properties, we conduct a simple Monte Carlo experiment with 3 variables ( $N = 3$ ) and roughly the same number of observations ( $T = 275$ ) as in the historical data. We abstract from the estimation of the VAR matrices  $B_i$  by setting

$x_t = u_t$  so that the reduced-form residuals are observed. The economic shocks are all *NID* with unit variance and no covariances. The off-diagonal elements of  $\Theta$  are all 0.2. Like Montiel Olea, Stock, and Watson (2020) we generate the instrumental variables using a linear measurement-error model:

$$z_{n,t} = 0.2\epsilon_{n,t} + \sigma_z v_{n,t}, \quad (13)$$

where  $v_{n,t}$  are independent standard normal variables. This setup implies the exclusion restrictions, because each instrument depends on only one shock. We set  $\sigma_z = 0.316$ , which gives an average, first-stage  $F$ -statistic of roughly 100. This choice means that there will not be replications with weak instruments, respecting the guideline of Montiel Olea, Stock, and Watson (2020) that allows standard inference. It also ensures we do not present evidence that is rigged against the just-identified case, for it has three strong instruments. The instruments thus satisfy assumptions (3a) and (3b). We simulate 5000 Monte Carlo replications.

Figure 2 shows the results. The first panel shows the simulated density of the  $J$ -test statistic (in red) along with the  $\chi^2(3)$  density (in grey). The two coincide closely, so that the  $J$ -test is not subject to much size distortion in this environment.

The second panel shows the density of  $\hat{\Theta}_{12}$ , which is completely representative because the VAR is symmetric. The just-identified case is in black and the overidentified case is in red. The overidentified estimator is much more efficient. Recall that the IRF (9) begins with  $\psi_0 = \Theta$  so, in a simulated VAR with  $B_i \neq 0$ , the confidence intervals for IRFs also would be narrower in the overidentified case, as they are in Figure 1.

The third panel shows the density of a representative, sample correlation between two estimated shocks — $\rho(\hat{\epsilon}_{1,t}, \hat{\epsilon}_{2,t})$ —again with the overidentified case in red and the just-identified case in black. In the overidentified case the simulated shock correlations are highly concentrated around zero, the population value. In the just-identified case their density is much more dispersed. Thus there is a much higher probability of finding a non-zero shock correlation, like those found in the historical application.

## 7. Conclusion

This paper studies SVARs identified with external instruments (proxy variables), with one instrument per shock. IV estimation can lead to measured shocks that are correlated, so we propose adding zero-covariance restrictions and estimating the combined moment conditions by GMM. This approach has several appealing features: (a) It can easily accommodate different sample spans for instruments and SVAR variables; (b) It provides overidentification and so yields a  $J$ -test of the model; (c) It also allows overidentification when an instrument is missing.

Our application is to the  $\{g_t, \tau_t, y_t\}$  SVAR using a full set of instruments (carefully constructed by several researchers). There the measured shocks in the just-identified IV case have correlations that range from -0.13 to 0.16. When we add the covariance restrictions the  $J$ -test statistic is small so that the restrictions are readily accepted. Given these two findings, one might expect that adding the restrictions is innocuous. But we find the standard errors on  $\hat{\Theta}$  fall by an average of 30% in the overidentified case. Thus confidence intervals for multipliers are narrower. We also illustrate the identification of the SVAR when an instrument is missing, using the example of the instrument for output shocks.

Monte Carlo simulations demonstrate the possibility of finding sample shock correlations when the restrictions are not imposed, even with strong instruments and in realistic sample sizes. They also illustrate the efficiency gains when the restrictions are added in GMM estimation.

## References

- Andrews, I. and Stock, J.H. (2018) Weak instruments and what to do about them. Short course lectures, NBER Summer Institute. <https://www.nber.org/conferences/si-2018-methods-lectures-weak-instruments-and-what-do-about-them>
- Angelini, G. and Fanelli, L. (2019) Exogenous uncertainty and the identification of structural vector autoregressions with external instruments. *Journal of Applied Econometrics* 34, 951–971.
- Angelini, G., Caggiano, G., Castelnuovo, E., and Fanelli, L. (2020) Are fiscal multipliers estimated with proxy-SVARs robust? Research Discussion Papers 13/2020, Bank of Finland.
- Angelini, G., Cavaliere, G., and Fanelli, L. (2021) An identification strategy for proxy-SVARs with weak proxies. manuscript, University of Bologna.
- Bernanke, B. and I. Mihov (1998) Measuring monetary policy. *Quarterly Journal of Economics* 113, 869–902.
- Blanchard, O. and Perotti, R. (2002) An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics* 117, 1329–1368.
- Brüggemann, R., Jentsch, C., and Trenkler, C. (2016) Inference in VARs with conditional heteroskedasticity of unknown form. *Journal of Econometrics* 191, 69–85.
- Cai, M., M. Del Negro, M.P. Giannoni, A. Gupta, P. Li, and E. Moszkowski (2019) DSGE forecasts of the lost recovery. *International Journal of Forecasting* 35, 1770–1789.
- Caldara, D. and Kamps, C. (2017). The analytics of SVARs: A unified framework to measure fiscal multipliers. *Review of Economic Studies* 73, 195–218.
- Del Negro, M., M.P. Giannoni, D. Giannone, and A. Tambalotti (2017) Safety, liquidity, and the natural rate of interest. *Brookings Papers on Economic Activity* Spring, 235–294
- Fengler, M.R. and J. Polivka (2021) Identifying structural shocks to volatility through a proxy-MGARCH model. Universität St. Gallen, Discussion paper no. 2021-3.
- Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco, Working Paper No. 2012-19.
- Gertler, M. and Karadi, P. (2015) Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics* 7, 44–76.
- Guay, A. and Normandin, M. (2018) Identification of structural vector autoregressions through higher unconditional moments. mimeo, HEC Montréal.
- Jentsch, C. and Lunsford, K.G. (2019) The dynamic effects of personal and corporate income tax changes in the United States: Comment. *American Economic Review* 109, 2655–2678.

- Lanne, M. and Luoto, J. (2021) GMM estimation of non-Gaussian structural vector autoregression. *Journal of Business and Economic Statistics* 39, 69–81.
- Lewis, D.J. (2020) Identifying shocks via time-varying volatility. mimeo, Federal Reserve Bank of New York.
- Liu, C. and Williams, N. (2019) State-level implications of federal tax policies. *Journal of Monetary Economics* 105(C), 74–90.
- Mertens, K. and Ravn, M.O. (2012) Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. *American Economic Journal: Economic Policy* 4, 145–181.
- Mertens, K. and Ravn, M.O. (2013) The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review* 103, 1212–1247.
- Mertens, K. and Ravn, M.O. (2014) A reconciliation of SVAR and narrative estimates of tax multipliers. *Journal of Monetary Economics* 68(S), 1–19.
- Mertens, K. and Ravn, M.O. (2019) The dynamic effects of personal and corporate income tax changes in the United States: Reply. *American Economic Review* 109, 2679–2691.
- Mertens, K. and Montiel Olea, J.L. (2018). Marginal tax rates and income: New time series evidence. *Quarterly Journal of Economics* 133, 1803–1884.
- Montiel Olea, J.L. and Pflueger, C. (2013) A robust test for weak instruments. *Journal of Business and Economic Statistics* 31, 358–369.
- Montiel Olea, J.L., J.H. Stock, and M.W. Watson (2020) Inference in structural vector autoregressions identified with an external instrument. *Journal of Econometrics* forthcoming
- Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics* 24, 960–992.
- Perotti, R. (2008) In search of the transmission mechanism of fiscal policy. *NBER Macroeconomics Annual 2007*, 22, 169–226.
- Perotti, R. (2011) Expectations and fiscal policy: An empirical investigation. IGIER working paper 429.
- Ramey, V. A. (2011a) Identifying government spending shocks: It’s all in the timing. *Quarterly Journal of Economics* 126, 1–50.
- Ramey, V. A. (2011b) Can government purchases stimulate the economy? *Journal of Economic Literature* 49, 673–685.
- Ramey, V.A. (2016) Macroeconomic shocks and their propagation. chapter 2 in *Handbook of Macroeconomics*. eds. J.B. Taylor and H. Uhlig. Amsterdam: Elsevier.
- Ramey, V. A. (2019) Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives* 33 (Spring), 89–114.

- Ramey, V. A. and Zubairy, S. (2018). Government spending multipliers in good times and in bad: Evidence from US historical data. *Journal of Political Economy* 126, 850–901.
- Stock, J.H. (2008) What’s new in econometrics: time series, lecture 7. Short course lectures, NBER Summer Institute. <http://www.nber.org/minicourse-2008.html>.
- Stock, J.H. and Watson, M.W. (2012) Disentangling the channels of the 2007–2009 recession. *Brookings Papers on Economic Activity*, 1(1), 81–135.
- Stock, J. H. and Watson, M. W. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *Economic Journal* 128, 917–948.

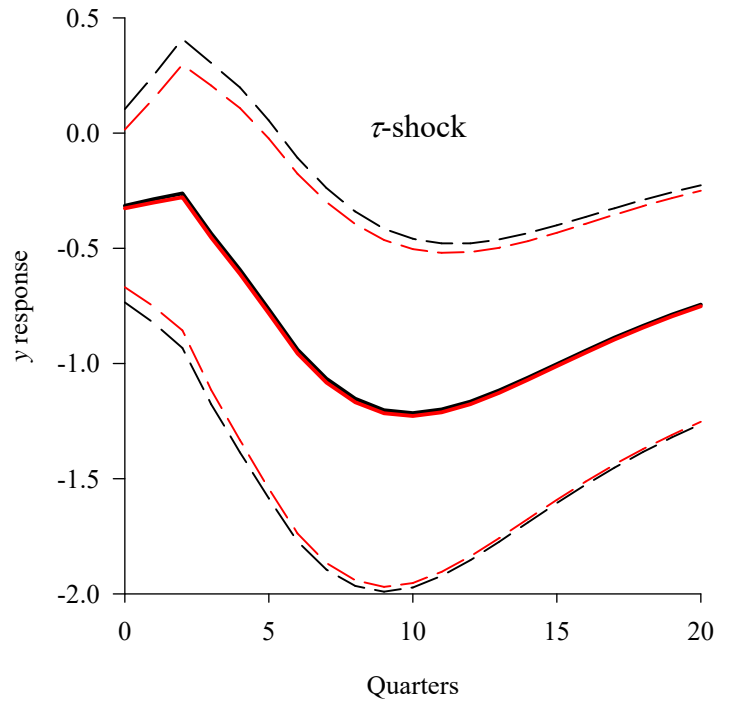
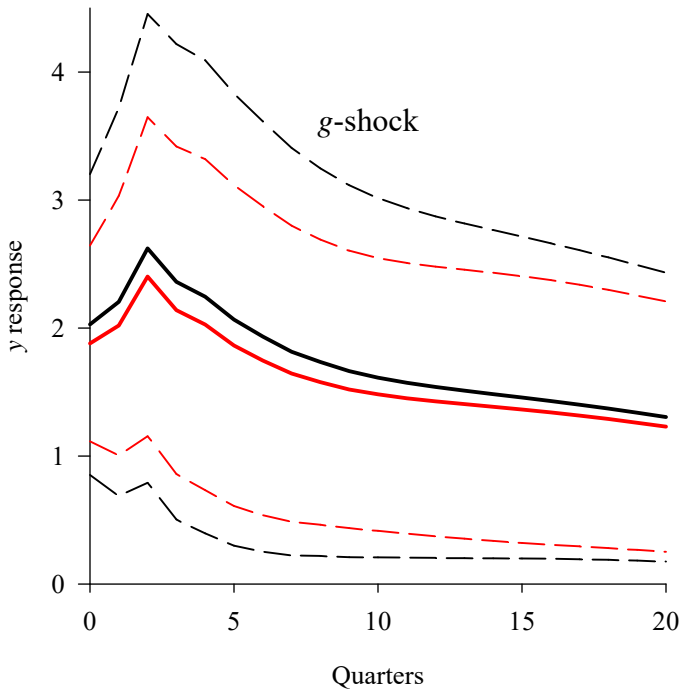
## Data Sources

Data are quarterly for 1948:I–2019:IV for US federal government spending, tax revenue, and GDP in real dollars per capita. Government spending,  $g_t$ , is federal government expenditure and gross investment; federal revenue  $\tau_t$ , is the sum of federal current tax receipts, social insurance contributions, and corporate income taxes; output,  $y_t$ , is Gross Domestic Product (GDP) all from the BEA. All variables are expressed as changes in logarithms after being deflated by the GDP deflator and expressed in per-capita terms.

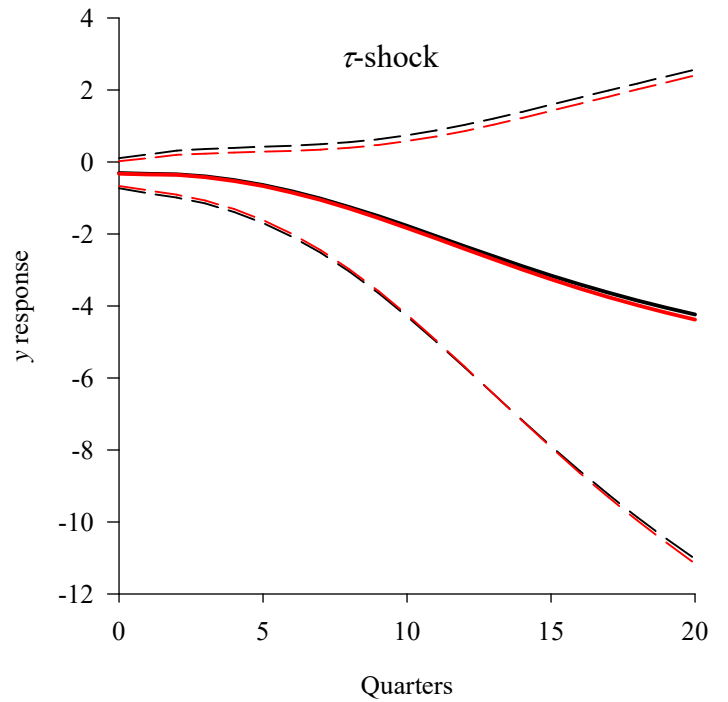
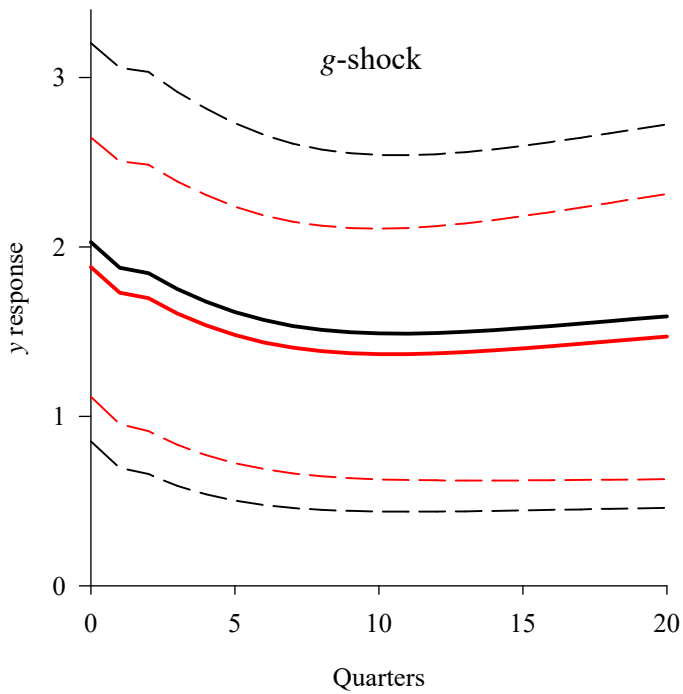
The instrument  $z_{g,t}$  comes from Ramey and Zubairy (2018). The instrument is an updated version of the Ramey (2011a) shocks which are a series of changes in the expected present value of government purchases as a result of military buildups. Following Ramey we deflate the government spending shocks by the the previous period’s nominal GDP.  $z_{\tau,t}$  is the main instrument from Mertens and Montiel Olea (2018) which is the average tax change for the lowest 90% of the income distribution. We convert their annual series to quarterly by assigning the tax change to the quarter it took effect.  $z_{y,t}$  is the the posterior mean of the innovation to the permanent productivity shock in the FRBNY DSGE model. We filter this by regressing it on its own 4 lags and 4 lags of the VAR variables. Documentation is in the references given in the text or at <https://github.com/FRBNY-DSGE/DSGE.jl>

**Figure 1: Multipliers**

**Dynamic**

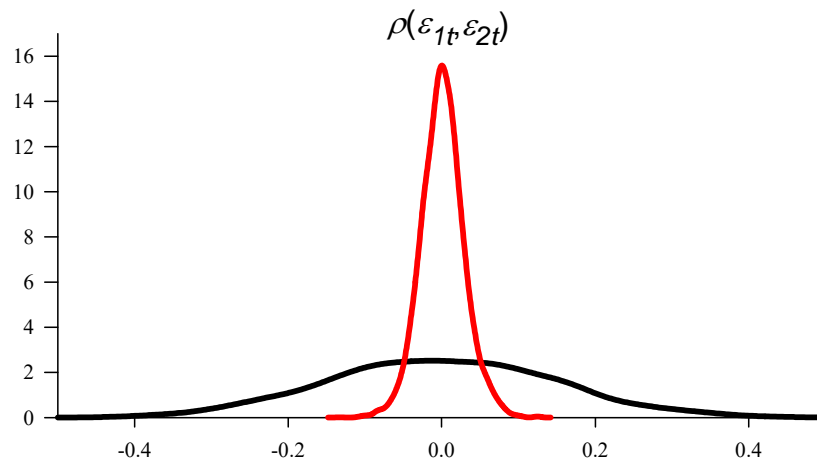
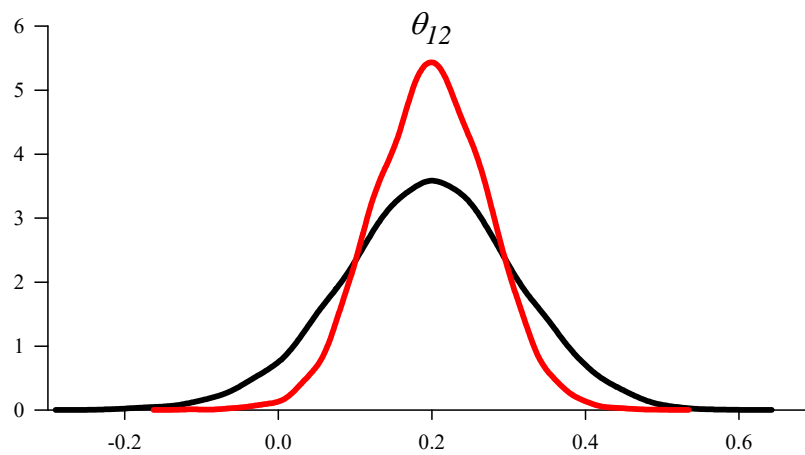
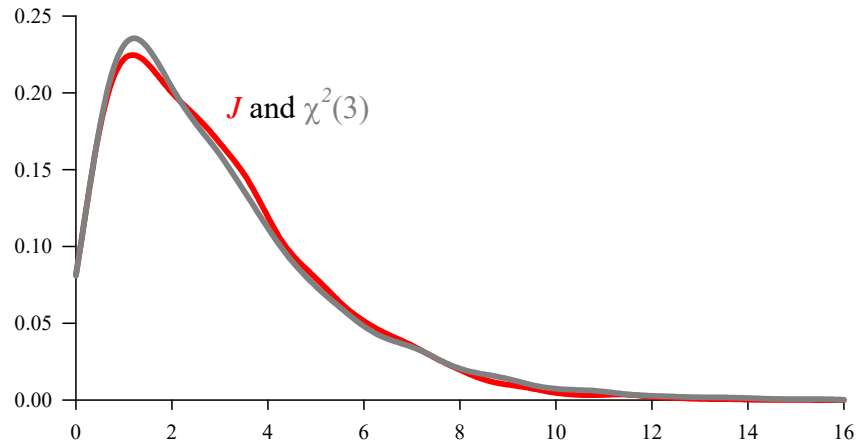


**Cumulative**



Notes: The upper (lower) panels show the dynamic (cumulative) multipliers for  $g$ -shocks (on the left) and  $\tau$ -shocks (on the right). Dashed lines show 68% confidence intervals. The just-identified case is in black and the overidentified case is in red.

**Figure 2: Monte Carlo Densities**



Notes: The upper panel shows the density of the  $J$ -test statistic (in red) along with that of the  $\chi^2(3)$  distribution (in grey) based on  $r=5000$  replications. The central panel shows the density function of  $\theta_{12}$ . The lower panel shows the density functions for the correlation between the two estimated structural shocks. The just-identified case is shown in black. The overidentified case is shown in red. The bandwidth is  $1.06 s r^{-0.2}$  where  $s$  is the standard deviation.