

Gregory, Allan W.; McNeil, James; Smith, Gregor W.

**Working Paper**

## US tax and spending shocks 1950-2019: SVAR overidentification with external instruments

Queen's Economics Department Working Paper, No. 1461

**Provided in Cooperation with:**

Queen's University, Department of Economics (QED)

*Suggested Citation:* Gregory, Allan W.; McNeil, James; Smith, Gregor W. (2021) : US tax and spending shocks 1950-2019: SVAR overidentification with external instruments, Queen's Economics Department Working Paper, No. 1461, Queen's University, Department of Economics, Kingston (Ontario)

This Version is available at:

<https://hdl.handle.net/10419/247203>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



Queen's Economics Department Working Paper No. 1461

# US Tax and Spending Shocks 1950-2019: SVAR Overidentification with External Instruments

Allan W. Gregory  
Queen's University

James McNeil  
Dalhousie University

Gregor W. Smith  
Queen's University

Department of Economics  
Queen's University  
94 University Avenue  
Kingston, Ontario, Canada  
K7L 3N6

7-2021

# US Tax and Spending Shocks 1950–2019: SVAR Overidentification with External Instruments

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

August 2021

## 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 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 restricting the shock covariances to be zero. In this application the overidentifying restrictions are not rejected. Compared to the unrestricted case the restricted SVAR yields (a) greater precision in estimating impulse response functions and multipliers and (b) smaller estimated effects of government spending shocks on output growth.

---

*JEL classification:* E62, C36.

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

<sup>†</sup>Gregory and Smith: Department of Economics, Queen's University, Canada; awg@econ.queensu.ca, smithgw@econ.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 helpful comments we thank Luca Fanelli and participants at the IAAE 2021 Annual Conference. We thank Marco Del Negro and 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 often viewed as incompatible with the definition of a shock. 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. 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 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).

Mertens and Ravn (2013), Angelini and Fanelli (2019), 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. (Arias, Rubio-Ramírez, and Waggoner (2018) and Giacomini, Kitagawa, and Read (2020) study set-identification in this environment using Bayesian methods.) We

consider the case when this assumption is satisfied and the SVAR can be fully identified with sufficient instruments. We have not found other examples of SVARs fully identified by external instruments. A number of studies have only a single instrument, perhaps because finding valid instruments involves painstaking work by researchers like those cited below.

Studying the fully-identified SVAR has advantages that we hope will strike the reader as intuitive. It allows us to test the hypothesis that the structural shocks are mutually uncorrelated. That test arises because the hypothesis of zero correlation also provides overidentifying restrictions on the SVAR and, in fact, is sometimes used to just identify it in the absence of enough external instruments. In the application here the  $J$ -test finds that the overidentifying restrictions cannot be rejected, so that overidentification allows one to isolate orthogonal shocks. An additional, important advantage is that the estimates from the overidentified SVAR may be more efficient than those from the SVAR-IV method. In our application, standard errors fall by an average of 39%, a notable improvement given the wide confidence intervals often associated with SVARs.

Section 2 outlines the measurement of variables and instruments, and describes the two identifications that we study. Section 3 then estimates the just-identified and overidentified SVARs. For shocks to spending and to revenue, and for each identification, we report impulse response functions, dynamic present-value multipliers, and historical decompositions, focusing on the effects on output. Overidentification adds to the precision with which impulse response functions and multipliers are estimated. It greatly reduces the estimated impact on output growth of shocks to government spending growth. It also attenuates the effect of tax shocks, though this change is not statistically significant. Section 4 briefly reports on some alternative specifications of the SVAR and their effects on these findings. Section 5 concludes.

## 2. SVAR Identification and Instruments

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 expressed as changes in logarithms after being deflated by the GDP deflator and expressed in per-capita terms. Figure 1 shows the three quarterly growth rates for 1950–2019.

Let  $w_t$  be a vector of deterministic terms: a constant, 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  $\Delta x_t$  then follows the VAR:

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

where  $\Sigma$  is positive definite. Residuals  $u_t$  are related to structural shocks

$\epsilon_t \equiv \{\epsilon_{g,t}, \epsilon_{\tau,t}, \epsilon_{y,t}\}'$  like this:

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

so that  $\Theta$  is a  $3 \times 3$  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 positive definite. We shall later test whether it is diagonal, implying orthogonal shocks. The shock  $\epsilon_{g,t}$  for example, has variance  $\sigma_g^2$ , 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 three external instruments  $z_t \equiv \{z_{g,t}, z_{\tau,t}, z_{y,t}\}'$  corresponding to the three elements of  $x_t$ . For an instrument to be valid, say for identifying the first column of  $\Theta$ , we require relevance and exclusion restrictions:

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

$$E(z_{g,t} \epsilon_{\tau,t}) = E(z_{g,t} \epsilon_{y,t}) = 0, \quad (3b)$$

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 2SLS as shown by Mertens and Ravn (2013) and Stock and Watson (2018). Thus with a valid instrument for each shock the SVAR can be identified, and the shocks measured as

$$\hat{\epsilon}_t = \hat{\Theta}^{-1} \hat{u}_t. \quad (4)$$

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 focus only on their main instrument which is the average tax change across brackets. 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). The tax instrument is available until 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. To span the 1960–2019 period, we use the shocks from the FRBNY DSGE model which is the most prominent, ongoing DSGE model for the US. Specifically, we use the posterior mean of the innovation to the temporary productivity shock in that model. Del Negro *et al* (2017) and Cai *et al* (2019) provide descriptions and applications of the model.

We assume that the SVAR is invertible 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.

Figure 2 graphs the three instruments. The instruments are available for a smaller sample than the variables in the VAR, so our GMM procedure is estimated using the shorter sample period. The reduced-form VAR parameters ( $B$  and  $D$ ) are estimated using data from the full 1950–2019 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.

Under the identification with external instruments, the structural shocks can be correlated. That is not precluded by the relevance (3a) and exogeneity (3b) of the instruments. For example, suppose that the two policy shocks have a common component  $\epsilon_{c,t}$ . We could respecify the policy shocks as:

$$\begin{aligned}\epsilon_{g,t} &= \phi_g \epsilon_{c,t} + e_{g,t}, \\ \epsilon_{\tau,t} &= \phi_\tau \epsilon_{c,t} + e_{\tau,t},\end{aligned}\tag{5}$$

where the common component  $\epsilon_{c,t}$  is orthogonal to the idiosyncratic components  $\{e_{g,t}, e_{\tau,t}\}$ . Here  $\Omega$  is not diagonal. Rigobon (2003) and Rigobon and Sack (2003) model common shocks in a similar manner and propose that they could be interpreted as unobservable macroeconomic news. Suppose that conditions (3a) and (3b) hold for the idiosyncratic components, and that the instruments are uncorrelated with the common component. In this case (3a) and (3b) also hold for the composite shocks  $\{\epsilon_{g,t}, \epsilon_{\tau,t}\}$  even though they are correlated. Of course these instruments will not provide much information if most of the variation in the composite shocks comes from the common component.

A small number of additional studies view shock correlation as natural, especially when the correlation is among policy variables as in the example (5) and so may reflect policy coordination. International macroeconomic models typically allow for monetary policy shocks to be correlated across countries. Alesina, Favero, and Giavazzi (2019a,



2019b) discuss correlated shocks to spending and revenue, which might be termed deficit or austerity shocks, measured using detailed narrative evidence.

The  $g$ - $\tau$ - $y$  VAR operates at quarterly frequency. New information thus arrives at a stately pace, because of the cost of constructing these measures. In practice spending and revenue may both react to financial-market indicators or political developments within the quarter. One might proceed by including these variables in the VAR, but they may not be observed or known to the econometrician. Weber (2010) considers this possibility and describes identifying an SVAR using a parametric model of conditional covariances that does not require them to be zero.

However, most researchers view an absence of correlation in structural shocks as intrinsic to the definition of a shock in an SVAR. As Bernanke (1986) and Cúrdia and Reis (2010) note, detecting such a correlation therefore may be useful as an indicator of a missing variable, so that the model is misspecified. 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.”

The fully-instrumented environment allows a test of the hypothesis that the shocks are uncorrelated—so that  $\Omega$  is a diagonal matrix—which we refer to as the covariance restrictions. Under that hypothesis the off-diagonal elements of  $\Omega$  are zero. Imposing that condition in estimation leads to three additional restrictions (or in general  $N \times (N - 1)/2$  for a VAR with  $N$  variables) and a  $J$ -test with three degrees of freedom. In the application we also show  $J$ -test statistics for combinations of two covariances and for individual covariances. We suggest these as a set of diagnostics that might reveal that an instrument does not satisfy the exclusion restrictions (3b) or other misspecification.

In general, suppose there are  $N$  variables in the SVAR and  $K$  instruments, each correlated with one of the structural shocks and uncorrelated with the others. In total, the matrix  $\Theta$  contains  $N(N - 1)$  unknown parameters, because the diagonal elements are normalized to one. To fully identify the SVAR we require at least this many moment

conditions. Each instrument  $z_{k,t}$  identifies a column of  $\Theta$  using:

$$E[(\hat{u}_{j,t} - \Theta_{j,k}\hat{u}_{k,t})z_{k,t}] = 0 \quad \text{for } k \neq j, \quad (6)$$

which gives  $K(N-1)$  moments. There are also  $N(N-1)/2$  moments from the relationship between the structural shocks, when these are assumed to be uncorrelated:

$$E[\hat{\epsilon}_{j,t}\hat{\epsilon}'_{k,t}] = 0 \quad \text{for } k \neq j. \quad (7)$$

The shocks  $\epsilon_{j,t}$ , defined in equation (4), depend on  $\Theta$  which is how these moments can provide additional information about the model parameters. Together, the instruments and covariance restrictions give a total of  $K(N-1) + N(N-1)/2$  moment conditions. Identification of all columns of  $\Theta$  requires enough instruments to satisfy:

$$K(N-1) + \frac{N(N-1)}{2} \geq N(N-1), \quad (8)$$

which simplifies to  $K \geq N/2$ . Thus the covariance restrictions may allow one to fully identify the SVAR without a full set of instruments.

Equation (8) implies that the three-variable SVAR also would be overidentified using the covariance restrictions if we had only two instruments. In that case there are four IV moments and three covariance moments to identify six parameters (the off-diagonal elements of  $\Theta$ ). We do not focus on this case because the three instruments appear strong in this application. But section 4 briefly reports on an example in which we add a variable to the SVAR while using the same set of instruments.

To our knowledge, very few SVAR studies use a full set of instruments for identification (see Kilian and Lütkepohl, 2017). Angelini *et al* (2020, section 3.4) do use a full set. They combine those with short-run restrictions, within an SVAR that includes an auxiliary statistical model of the instruments, and so have overidentification. Our approach differs in that we base identification only on the instruments and the covariance restrictions, which we also test.

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.

### 3. Findings

We estimate the VAR with  $p = 4$  lags. This 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, 2014), and Liu and Williams (2019). We then consider two identifications. First, we just-identify the SVAR using the three instruments  $z_t$ . Second, we also impose the restriction that the shock covariances are zero.

#### 3.1 Instrument Relevance and Shock Properties

We first check on the relevance of each instrument. The first-stage  $F$ -statistics, equation-by-equation, for  $\{g_t, \tau_t, y_t\}$ , are: 9.45, 12.39, and 105.1. Thus each value is near or above 10. Above this cutoff, Montiel Olea, Stock, and Watson (2020) suggest that one can use standard methods for inference. 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  $z_{y,t}$  is obviously strong too.

Estimates and standard errors are calculated by iterated GMM. (Implementing a parametric bootstrap is challenging because the fiscal instruments have many zero observations, as seen in Figure 2.) 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 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  $J = 4.28$  yielding a  $P$ -value of 0.23. Thus there is no significant

evidence against the overidentified case. The top row of Table 1 contains this result. Further rows then show tests of the hypotheses that pairs of covariances are zero and then of hypotheses that individual covariances are zero. Though these tests are of course not independent, we suggest them as a set of diagnostics. For example, a rejection might suggest an exclusion restriction does not hold, and hence cast doubt on one or more instruments. Overall, though, in this application there is little evidence against these restrictions. Thus our next goal is to document the effect of this overidentification on the economic findings.

Table 2 shows the estimates  $\hat{\Theta}$  from the two identifications, along with standard errors. Notice that in both cases the estimated impact of a tax shock  $\hat{\epsilon}_{\tau,t}$  on  $g_t$  is positive and different from zero at the 5% level of significance. That finding is inconsistent with the Blanchard-Perotti timing restriction for example. There also are several differences between the two matrices but the most notable effect is the increase in precision. Moving from the IV estimation (in the upper panel) to the overidentified estimation with covariance restrictions (in the lower panel) the standard errors fall by an average of 39%.

Table 3 then reports the correlations between the three structural shocks that are calculated from  $\hat{\Theta}$  and the reduced-form shocks as in equation (4). The striking finding here is that there are correlations between the structural shocks from the traditional identification that uses one external instrument to identify each shock. Specifically, the shocks to the policy variables are correlated with shocks to output:  $\rho(\hat{\epsilon}_{g,t}, \hat{\epsilon}_{y,t}) = -0.50$  and  $\rho(\hat{\epsilon}_{\tau,t}, \hat{\epsilon}_{y,t}) = 0.24$ . 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. The lower panel of Table 3 then shows the remaining correlations once the covariance restrictions are imposed. They are very small, in keeping with the support for the restrictions from the  $J$ -test in Table 1.

We calculate the three structural shocks by inverting  $\hat{\Theta}$ , as shown in equation (4). One concern with this approach is that it may be sensitive to violations of the estimation assumptions because identification of the three shocks is no longer independent, even for IV estimation. This is because in general the inverse of a matrix depends on all elements of the matrix. If the exclusion restrictions were not satisfied for one of the instruments, then

the associated column of  $\hat{\Theta}$  would not be correct. These errors could affect all elements of the inverse of  $\hat{\Theta}$ , and by extension all of the shocks, even if the assumptions for the other two instruments are satisfied. If the shocks were instead uncovered individually then violations of the identifying assumptions for one shock should not influence the other shocks, which may result in lower correlations for at least some of the shocks. Stock and Watson (2018), who study the case when only one column of  $\Theta$  is identified, show that the respective shock—the government spending shock, for example—can be uncovered as:

$$\hat{\epsilon}_{g,t} = \frac{\hat{\Theta}'_g \hat{\Sigma}^{-1}}{(\hat{\Theta}'_g \hat{\Sigma}^{-1} \hat{\Theta}_g)} \hat{u}_t, \quad (9)$$

where  $\Theta_g$  is the column of  $\Theta$  associated with government spending shocks, which is the first column in our application. Since equation (9) depends only on a single column of  $\Theta$ , these shocks may be less susceptible to violations of the exclusion restrictions. The correlations of the shocks identified in this way are:  $\rho(\hat{\epsilon}_{g,t}, \hat{\epsilon}_{\tau,t}) = -0.17$ ,  $\rho(\hat{\epsilon}_{g,t}, \hat{\epsilon}_{y,t}) = 0.53$ ,  $\rho(\hat{\epsilon}_{\tau,t}, \hat{\epsilon}_{y,t}) = -0.29$ . The estimated correlations do depend on the way in which the shocks are calculated. For example, these three correlations change sign compared to the upper panel of Table 3. Notice, though, that all correlations are larger in absolute value when calculated with the Stock and Watson (2018) method, indicating that our finding of sizable correlations between the structural shocks is not sensitive to the method of calculation.

Table 4 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 4.

Angelini *et al* (2020) also study a fiscal SVAR for the US but find a significant correlation between  $\hat{\epsilon}_{\tau,t}$  and  $z_{y,t}$ . They 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. They then develop a toolkit for identification when condition (3b) does not hold. They also show that incorrectly imposing a zero correlation between the tax shock and the output instrument leads to a low tax multiplier, in keeping with the analyt-

ical results of Caldara and Kamps (2017). In the next sub-section we report a relatively small tax multiplier, even though this correlation in Table 4 is negligible.

### 3.2 Impulse Response Functions and Multipliers

Figure 3 shows impulse response functions from the two identifications for the effect of  $\hat{\epsilon}_g$  on  $y$  (in the upper panel) and the effect of  $\hat{\epsilon}_\tau$  on  $y$  (in the lower panel). The SVAR is in growth rates so the IRFs are cumulative and level off rather than decaying to zero. The figure includes 68% and 90% asymptotic confidence intervals (the dotted and dashed lines respectively) 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 impulse response functions.) The vertical axis scale is the same for each  $g$ -shock and then for each  $\tau$ -shock. The just-identified case is in black and the overidentified case is in red.

There are three differences between the identifications. First, the overidentification improves precision (narrows the confidence intervals). Second, the overidentified case—on the right-hand side—provides some evidence of smaller effects of tax shocks on output relative to those in the just-identified case. Third, the two identifications differ dramatically in their findings for the effect of government spending shocks on output. In the just-identified case the effect is positive, peaking at a value of 0.3 at the 4-quarter and 5-quarter horizons. The 90% confidence interval lies above the zero axis. In contrast, the point estimates for the overidentified case are slightly negative, though both the 90% and 68% confidence intervals include zero. Thus the covariance restrictions lead to a sharp revision in the estimated impact of a government spending shock  $\hat{\epsilon}_{g,t}$ . In the just-identified case one would conclude spending shocks are expansionary. But using the overidentifying restrictions—and with the same level of confidence—one would conclude that spending shocks have no statistically significant effect on output.

Define  $\Psi_{y,g,h}$  as the cumulative response of output to a shock to variable  $g$  after  $h$  quarters, and  $\Psi_{y,\tau,h}$  as the corresponding measure for tax shocks. Because the variables are in log-differences this has the interpretation as the cumulative percentage change of

output after one of these shocks. We next calculate and report the spending and tax multipliers, which give the relative change to output to a given shock at some horizon. We follow Mountford and Uhlig (2009) and compute the cumulative multiplier for government spending. In response to a shock to 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. It is given by:

$$m_{y,g,h} = \frac{\bar{y} \sum_{j=0}^h (1 + \bar{i})^{-j} \Psi_{y,g,h}}{\bar{g} \sum_{j=0}^h (1 + \bar{i})^{-j} \Psi_{g,g,h}}. \quad (10)$$

where  $\bar{i}$ ,  $\bar{y}$ , and  $\bar{g}$  are the average 3-month treasury bill rate, output, spending, and tax revenue over the sample period. The tax multiplier  $m_{y,\tau,h}$  is computed similarly. Ramey (2019) discusses these and other multiplier expressions.

Figure 4 shows these spending and tax multipliers. The  $g$ -multipliers are in the upper panel and the  $\tau$ -multipliers are in the lower panel. Again the just-identified case is in black and the over-identified case is in red. The scale is the same for both  $g$ -multipliers and for both  $\tau$ -multipliers. The multipliers are shown with 68% and 90% confidence intervals (the dotted and dashed lines respectively). Perhaps the most noteworthy feature of Figure 4 is that the overidentification leads to greater precision/narrower confidence intervals.

In the just-identified case the  $g$ -multiplier starts at 2.1 and falls to 1.4 after 20 quarters. The confidence intervals are relatively wide, but always lie above the zero axis. In contrast, in the overidentified case the point estimates are slightly negative at all horizons and the 68% and 90% confidence intervals include zero. Thus the overidentification has a large effect on the  $g$ -multiplier estimates.

Ramey (2019, Table 1) reports that an updated Blanchard-Perotti identification gives a  $g$ -multiplier between 0.6 and 0.8. She also notes that time series studies overall yield multipliers in this range or perhaps as high as 1, whether measured at their peak or averaged over 5 years. Thus the just-identified  $g$ -multiplier in Figure 4 is larger than those typical of previous studies (though the difference may not be statistically significant) while the overidentified one is smaller.

Ramey notes that  $\tau$ -multipliers tend to increase in scale over time, and that is true in the lower panel of Figure 4 as well. In the just-identified case the multiplier peaks at -1.12

after 5 years while in the overidentified case it peaks at -0.59. These impacts are smaller than those in the studies surveyed by Ramey (2019, Table 2) which include the estimates of Mertens and Ravn (2013, 2014) who find values from -2.5 to -3. However, overall there is more variation across studies in the scale of  $\tau$ -multipliers than in the scale of  $g$ -multipliers. The point estimates in the lower panel of Figure 4 suggest comparatively small effects of revenue shocks that are more typical of those found in New Keynesian DSGE model than in SVAR or narrative studies. But there is considerable sampling variability even with overidentification.

### 3.3 Historical Decompositions and FEVDs

We next present historical decompositions that measure the cumulative effect on output growth of the current and previous 20 fiscal shocks, at each quarter in the sample period. Let  $\psi_{y,g,h}$  be the response of output to a government spending shock  $h$  periods after the shock occurs. The historical decomposition is the cumulative contribution of the current and previous 20 of these shocks:

$$\sum_{h=0}^{20} \psi_{y,g,h} \epsilon_{g,t-h}. \quad (11)$$

Figure 5 presents the results for  $g$ -shocks in the upper panel and for  $\tau$ -shocks in the lower panel. These measures reflect both the estimated effect of fiscal shocks (shown in Figure 3) and the shocks themselves (which the online appendix shows are similar for the two identifications). The just-identified case is in black and the overidentified one is in red. NBER-dated recessions are shown in grey. The horizontal axis begins in 1956 because we measure cumulative effects over the previous 5 years.

Recall that the overidentified model suggests smaller estimates of the effects of a spending shock. For  $g$ -shocks the upper panel of Figure 5 shows the effect of this difference: The historical effect of these shocks is estimated to be much smaller in the overidentified SVAR. In the overidentified case there also is no longer evidence of countercyclical government spending shocks during NBER recessions. The lower panel of Figure 5 shows the effects of  $\tau$ -shocks. Here the just-identified and overidentified models yield similar estimates, though the overall effect of tax shocks is smaller in the overidentified case.



These differences between the results under the two identifications are accompanied by a more general warning. In the just-identified case the impulse response functions and historical decompositions generally are misleading because they incorrectly treat the estimated shocks as uncorrelated. This is also true for the forecast error variance decompositions (FEVDs), as we now illustrate. Table 5 shows the FEVD of output at horizon zero, calculated as:

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

for each of the  $k = g, \tau, y$  structural shocks. Similar statistics could be calculated at further forecast horizons but these would additionally involve the parameter matrices  $B_i$ . Since those parameters do not depend on identification of the SVAR we focus on the estimates at horizon zero.

The upper row of Table 5 shows estimates of the statistic (12) for the just-identified case and the lower row shows the values for the overidentified case. The columns of Table 5 indicate the fraction of the variance of the forecast errors of output that can be explained by one of the estimated, structural shocks:  $\hat{\epsilon}_{g,t}$ ,  $\hat{\epsilon}_{\tau,t}$ , and  $\hat{\epsilon}_{y,t}$ . Because the forecast errors  $\hat{u}_t$  are just combinations of these shocks, together the shocks should explain all the variation in the forecast errors of output and the sum of equation (12) over  $k$  should equal one. The final column of Table 5 shows this sum.

The most striking feature of Table 5 is that the just-identified model suggests that the three shocks together explain more than 100% of the variation of the forecast errors. In fact, output shocks on their own can apparently explain more than the total variation of the forecast errors, according to this model. This is because the forecast error variance decomposition assumes that the structural shocks are orthogonal, whereas Table 3 documented large correlations between the shocks. For example, a positive shock  $\hat{\epsilon}_{g,t}$  is associated with a negative shock  $\hat{\epsilon}_{y,t}$ , so that one should draw from a multivariate shock density to study the findings in this case.

When the covariance restrictions are included as overidentifying moment conditions the total contribution of all shocks to the forecast errors of output is much closer to one, as it should be. But we now see that the contribution of fiscal policy shocks to output

decreases from around 30% to essentially zero. This is consistent with Figure 3 which showed that the impact effect of the two fiscal shocks is much lower in the overidentified case.

### 3.4 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 simply 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  $\Delta x_t = u_t$  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:

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

where  $v_{j,t}$  are independent standard normal variables. To ensure strong instruments with an  $F$ -statistic of roughly 100, we set  $\sigma_z = 0.316$ . We simulate 5000 Monte Carlo replications. The instruments thus satisfy assumptions (3a) and (3b).

Figure 6 shows the results. The upper 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. This absence of size distortions is one reason to build in zero correlations between the simulated, structural shocks, because sub-section 3.1 found the covariance restrictions could not be rejected in the historical application.

The central 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, and confidence intervals for IRFs also would be narrower in this case, as they are in Figure 3.

The lower panel shows the density of a representative shock correlation  $\rho(\hat{\epsilon}_{1,t}, \hat{\epsilon}_{2,t})$  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 in Table 3. Overall then there are large efficiency gains from using the covariance restrictions.

#### 4. Variations

This section reports on the effects of some other specifications, including those with election dummy variables, using different instruments, variables with missing instruments, and using levels rather than growth rates. In each case we summarize the findings while reporting the full results in the online appendix.

##### 4.1 Election Timing

Drazen (2008) argues that there is little evidence of an electoral cycle in budget aggregates for the US. But in these data and VAR we find that the identified structural shocks are predictable using lags of an indicator for general elections. Specifically, the pattern suggests a reduction in spending and later a reduction in revenue after an election. Thus we added to the deterministic terms the current value and four lags of a dummy variable indicating the quarters in which general elections occur. Thus spending or revenue (or output) may respond to a recent or upcoming election. When we include these elements in  $w_t$  tests suggest (a) they are jointly significant in the revenue equation (with a  $P$ -value of 0.007) and (b) they are not simply capturing remaining seasonality. To our knowledge this is a new feature in this VAR. However, the main findings are not affected by the inclusion of the electoral dummy variable.

##### 4.2 Additional Instruments

Next we consider a further source of overidentification: additional instruments. While valid instruments typically are scarce in the macroeconomic context, several additions are possible in this VAR.

First we consider the instrument for unanticipated tax changes developed by Mertens and Ravn (2013) and extended by Liu and Williams (2019). When we include this in addition to the Mertens and Montiel Olea (2018) instrument the first-stage  $F$ -statistic falls from 12.39 to 6.18 (and the incremental  $F$ -statistic is only 0.017). On its own, this instrument has a first-stage  $F$ -statistic of 3.8. Thus it does not provide significant, additional information in this specific VAR.

Second, we consider instrumenting the shock to GDP growth with the quarterly series of utilization-adjusted total factor productivity constructed by Fernald (2014). The Fernald instrument has a first-stage  $F$ -statistic on its own of 10.58, but an incremental  $F$ -statistic of 0.15 when combined with the innovation to the temporary productivity shock from the FRBNY DSGE model. Thus we do not report findings from adding this instrument either.

Third, we add the innovation in the permanent component of the productivity shock (from the FRBNY DSGE model) as an additional instrument for the GDP shock. When combined with the innovation in the temporary component, this has an incremental  $F$ -statistic of 18.86 with an overall  $F$ -statistic of 44.13. The  $J$ -test of the overidentifying restrictions resulting both from this additional instrument and the covariance restrictions has a  $P$ -value of 0.196 so that the restrictions are not rejected at any traditional significance level. The additional instrument narrows the confidence intervals for IRFs slightly, but the findings are very similar to those in Figure 3 and so again are in the online appendix.

### 4.3 Missing Instruments and Additional Variables

One possible explanation for correlated shocks is an omitted variable. In the application of Section 3 the correlation was not significant, in that the zero-covariance restrictions passed the  $J$ -test. Nevertheless, we next indicate an added use for these restrictions: As noted in Section 2 they can allow identification of the SVAR with fewer external instruments than variables. (Angelini and Fanelli (2019) provide a general discussion of identification with fewer instruments than shocks.)

As mentioned in Section 2, we can identify all  $N$  structural shocks if we have  $K \geq N/2$  instruments. Thus we could drop one or more suspect instruments and use the covariance

restrictions to identify the columns of  $\Theta$  associated with the now uninstrumented variables. Or we could add a variable. Because the covariance restrictions introduce additional overidentifying moment conditions, applying our approach in this context amounts to using these conditions to identify the column of  $\Theta$  associated with the new variable. This could be an attractive approach for researchers interested in studying the causal effect of a variable without a known, valid instrument.

To illustrate, we use the original three instruments but add a fourth variable to the SVAR: consumption,  $c_t$ , measured as the log of real per capita personal consumption expenditures, seasonally adjusted. Adding  $\Delta c_t$  to the VAR naturally changes the residuals  $u_t$  but all three instruments remain strong. This addition adds 3 parameters to  $\Theta$  so that with 6 covariance restrictions there are 3 overidentifying restrictions. The  $J$ -test statistic is 3.36 with 3 degrees of freedom, yielding a  $P$ -value of 0.339. Thus the overidentifying restrictions are not rejected and so we present some of the findings from the overidentified case.

Figure 7 shows some IRFs from this case. The upper row shows the effects of a  $g$ -shock on output and consumption. The effect on output remains small (as in Figure 3) with confidence bands again including zero at each horizon. The effect on consumption is negative but only significantly so after 2 quarters and at the 68% confidence level.

Blanchard and Perotti (2002) and Fisher and Peters (2010) find positive effects of  $g$ -shocks on both output and consumption, of course with different identification methods. Ramey (2016, p 113) cites studies with similar findings, while noting that Mountford and Uhlig (2009) (who used sign restrictions) “found only weak effects on GDP and no significant effect on consumption.” Our findings are similar to Mountford and Uhlig’s.

Figure 7 also shows that a  $\tau$ -shock has a contractionary effect on output that is similar to that shown for the benchmark, 3-variable SVAR in Figure 3. A  $\tau$ -shock also leads to a significant decline in consumption and a significant increase in government spending. Mertens and Ravn (2012) find the same sign for the effect of  $\tau$ -shocks on consumption in US history.

The last entry in Figure 7 (in the lower right panel) shows one can identify a large effect of consumption (or ‘preference’) shocks on output, with both 68% and 90% confidence

intervals lying above the zero axis. Note that these preference shocks are identified even though we have no instrument for them.

#### 4.4 Levels

Our benchmark VAR (1) includes three variables specified in first differences. Some studies instead include variables in log-levels and remove the growth trend by including a quadratic time trend in the VAR. We favor a difference specification in our application because all three instruments appear strong, which demonstrates that correlated shocks can arise even when instruments appear sufficiently strong according to a standard benchmark.

When we estimate the VAR in levels (with a quadratic time trend) we find that the government spending instrument has a first-stage  $F$ -statistic of only 2.80, well below the conventional threshold of 10. Otherwise, our main results hold. The just-identified model finds that spending shocks are expansionary and tax shocks contractionary but also evidence of significant correlations between the shocks. Imposing the covariance restrictions as overidentifying moments, spending shocks become mildly contractionary (although not statistically different from zero) and the effect of tax shocks becomes attenuated. A test of the overidentifying restrictions has a  $J$ -statistic of 4.3 with a  $P$ -value of 0.23, so that the restrictions are not rejected at conventional levels.

### 5. Conclusion

An appealing feature of adopting an external instrument (proxy variable) is that it allows one to identify a column of  $\Theta$  and hence impulse response functions for the associated shock, without fully identifying the SVAR. For the  $\{g_t, \tau_t, y_t\}$  SVAR using a full set of instruments (carefully constructed by several researchers) we find empirically that this method leads to structural shocks that are correlated. For example, the correlation between spending and output shocks is -0.5. However, when we impose zero-covariance restrictions on the shocks those restrictions are not rejected, so that this correlation disappears and the shocks can be interpreted individually. The overidentification yields greater precision in estimating IRFs and multipliers. Monte Carlo simulations also illustrate the efficiency gains and the possibility of finding large shock correlations when the restrictions are not imposed.

The overidentification also suggests much smaller (indistinguishable from zero) effects of government spending shocks on output. As a result the historical contributions of these shocks to output growth and their counter-cyclical roles during recessions are estimated to have been much smaller.

These findings are not intended as criticism of the method of using a single instrument, for finding additional instruments can be challenging. And our example also is not meant as a decisive word on the effects of government spending shocks. Measuring their impacts also depends on the time period, the specification of the VAR, and the instruments. The findings simply illustrate the applied econometrics of this interesting SVAR. They suggest that researchers draw on additional instruments and then covariance restrictions where possible, both to add to precision and because doing so may change the economic findings.

## References

- Alesina, A., Favero, C. and Giavazzi, F. (2019a) Effects of austerity: Expenditure and tax-based approaches. *Journal of Economic Perspectives* 33:2, 141–162.
- Alesina, A., Favero, C. and Giavazzi, F. (2019b) *Austerity: When it Works and When it Doesn't*. Princeton University Press.
- 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.
- Arias, J.E., J.F. Rubio-Ramírez, and D. F. Waggoner (2018) Inference in Bayesian proxy-SVARs. Federal Reserve Bank of Atlanta working paper 2018-16.
- Bernanke, B. (1986) Alternative explanations of the money-income correlation. *Carnegie-Rochester Conference Series on Public Policy* 25:1, 49–99.
- 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.
- Cúrdia, V. and Reis, R. (2010) Correlated disturbances and U.S. business cycles. NBER working paper 15774.
- 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
- Drazen, A. (2008) Political budget cycles, in S.N. Durlauf and L.E. Blume, eds., in *The New Palgrave Dictionary of Economics* (2nd edition), London and New York: Macmillan Palgrave.
- Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco, Working Paper No. 2012-19.



- Fisher, J.D.M. and Peters, R. (2010) Using stock returns to identify government spending shocks. *Economic Journal* 120, 414–436.
- Gertler, M. and P. Karadi, P. (2015) Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics* 7, 44–76.
- Giacomini, R., T. Kitagawa, and M. Read (2020) Robust Bayesian inference in proxy SVARs. mimeo, Department of Economics, University College London.
- Guay, A. and Normandin, M. (2018) Identification of structural vector autoregressions through higher unconditional moments. mimeo, HEC Montréal.
- Kilian, L. and Lütkepohl, H. (2017) *Structural Vector Autoregressive Analysis*. Cambridge University Press.
- 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, Daniel 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., 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.
- Rigobon, R. (2003) Identification through heteroskedasticity. *Review of Economics and Statistics* 85, 777–792.
- Rigobon, R. and Sack, B. (2003) Measuring the reaction of monetary policy to the stock market. *Quarterly Journal of Economics* 118, 639–669.
- 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.
- Weber, E. (2010) Structural conditional correlation. *Journal of Financial Econometrics* 8, 392–407.

**Table 1: *J*-Tests**

$\rho(\epsilon_g, \epsilon_\tau) = 0$	$\rho(\epsilon_g, \epsilon_y) = 0$	$\rho(\epsilon_\tau, \epsilon_y) = 0$	$J$ ( $df$ )	$P$ -value
✓	✓	✓	4.28 (3)	0.23
✓	✓		3.65 (2)	0.16
✓		✓	1.20 (2)	0.55
	✓	✓	4.27 (2)	0.12
✓			0.14 (1)	0.71
	✓		3.64 (1)	0.06
		✓	0.93 (1)	0.33

Notes: The table reports  $J$ -statistics and their degrees of freedom and  $P$ -values for each combination of covariance restrictions shown with checkmarks. For example the top line shows the test of the hypothesis that all three covariances are zero so that the three shocks are mutually uncorrelated.

**Table 2: Estimates  $\hat{\Theta}$** 

IV			
	$g$	$\tau$	$y$
$g$	1.000	0.237** (0.099)	0.678* (0.403)
$\tau$	-0.244 (0.327)	1.000	0.414 (0.340)
$y$	0.191* (0.107)	-0.056 (0.048)	1.000

Overidentified GMM			
	$g$	$\tau$	$y$
$g$	1.000	0.192*** (0.070)	0.543 (0.346)
$\tau$	-0.376** (0.156)	1.000	0.847*** (0.203)
$y$	-0.026 (0.042)	-0.017 (0.031)	1.000

Notes: The table reports estimates  $\hat{\Theta}$  from  $u_t = \Theta \epsilon_t$  that are first just identified with instruments  $z_t$  and then overidentified with covariance restrictions. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level.

**Table 3: Structural Shock Correlations**

IV			
	$\hat{\epsilon}_g$	$\hat{\epsilon}_\tau$	$\hat{\epsilon}_y$
$\hat{\epsilon}_g$	1.00	0.02	-0.50
$\hat{\epsilon}_\tau$	0.02	1.00	0.24
$\hat{\epsilon}_y$	-0.50	0.24	1.00

Overidentified GMM			
	$\hat{\epsilon}_g$	$\hat{\epsilon}_\tau$	$\hat{\epsilon}_y$
$\hat{\epsilon}_g$	1.00	0.08	0.09
$\hat{\epsilon}_\tau$	0.08	1.00	0.06
$\hat{\epsilon}_y$	0.09	0.06	1.00

Notes: The table reports correlations between estimated structural shocks first from just-identified IV and then with covariance restrictions also imposed.

**Table 4: Shock-Instrument Correlations**

IV			
	$\hat{\epsilon}_{g,t}$	$\hat{\epsilon}_{\tau,t}$	$\hat{\epsilon}_{y,t}$
$z_{g,t}$	0.1846	0.0019	-0.0005
$z_{\tau,t}$	-0.0012	0.2689	0.0008
$z_{y,t}$	0.0004	0.0009	0.4749

Overidentified GMM			
	$\hat{\epsilon}_{g,t}$	$\hat{\epsilon}_{\tau,t}$	$\hat{\epsilon}_{y,t}$
$z_{g,t}$	0.1956	-0.0140	0.1039
$z_{\tau,t}$	0.0154	0.2749	-0.0252
$z_{y,t}$	0.0525	-0.0772	0.5783

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

**Table 5: FEVDs for Output**

	$\hat{\epsilon}_{g,t}$ -share	$\hat{\epsilon}_{\tau,t}$ -share	$\hat{\epsilon}_{y,t}$ -share	total
IV	0.283	0.019	1.418	1.720
Overidentified GMM	0.004	0.002	1.010	1.016

Notes: The table reports forecast-error variance decompositions (FEVDs) for output at horizon 0. Each entry is the share of the output innovation variance attributed to the shock in that column. The last column then gives the sum of these shares. The upper row shows results from just-identified IV while the lower row shows results from overidentified GMM with covariance restrictions also imposed.

## Notes to Figures

### Figure 1: Growth Rates

Notes: The figure shows quarterly growth rates for 1950–2019 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 variables are expressed as changes in logarithms after being deflated by the GDP deflator and expressed in per-capita terms.

### Figure 2: Instruments

Notes:  $z_{g,t}$  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 across brackets. We convert their annual series to quarterly by assigning the tax change to the quarter it took effect.  $z_{y,t}$  is the innovation in the temporary productivity shock in the FRBNY DSGE model. Documentation is in the references given in the text or at <https://github.com/FRBNY-DSGE/DSGE.jl>

### Figure 3: Impulse Response Functions

Notes: The upper panel shows the effect of  $\hat{\epsilon}_g$  on  $y$  while the lower panel shows the effect of  $\hat{\epsilon}_\tau$  on  $y$ . The left-hand side shows the just-identified case (in black) while the right-hand side shows the overidentified case (in red). The SVAR is in growth rates so the IRFs are cumulative and level off rather than decaying to zero. Dotted lines and dashed lines show 68% and 90% asymptotic confidence intervals respectively. The vertical axis scale is the same for each  $g$ -shock and then for each  $\tau$ -shock.

### Figure 4: Multipliers

Notes: The  $g$ -multipliers are in the upper panel and the  $\tau$ -multipliers are in the lower panel. The just-identified case is on the left (in black) and the overidentified case is on the right (in red). The scale is the same for both  $g$ -multipliers and for both  $\tau$ -multipliers. The multipliers are shown with 68% and 90% confidence intervals (the dotted and dashed lines respectively).



Figure 5: Historical Decompositions

Notes: The graphs measure the cumulative effect on output growth of the current and previous 20 fiscal shocks, at each quarter in the sample period. The upper panel gives results for  $g$ -shocks while the lower panel gives those for  $\tau$ -shocks. The just-identified case is in black and the overidentified one is in red. The horizontal axis begins in 1956 because we measure cumulative effects over the previous 5 years. NBER recession periods are shown in grey.

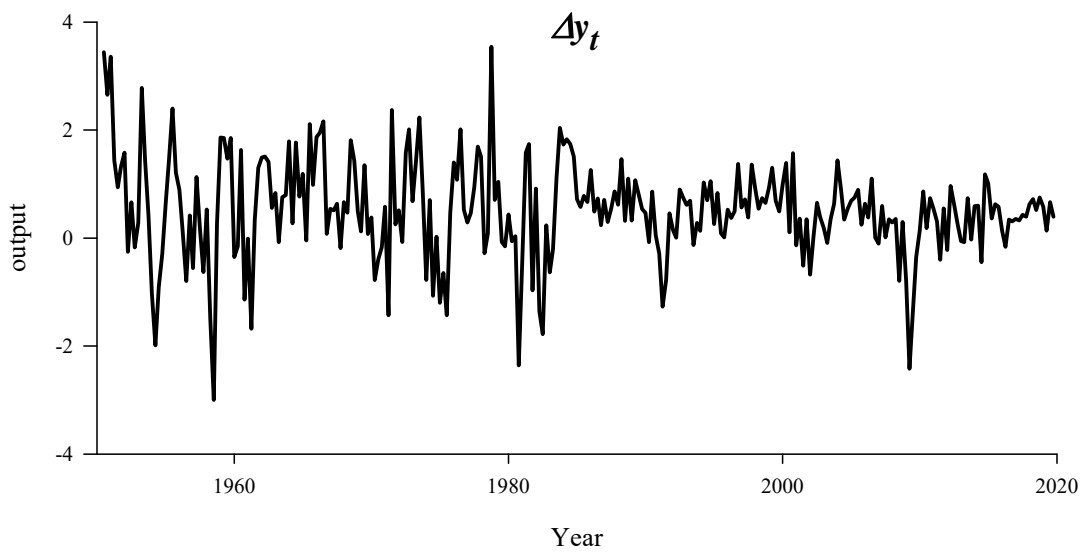
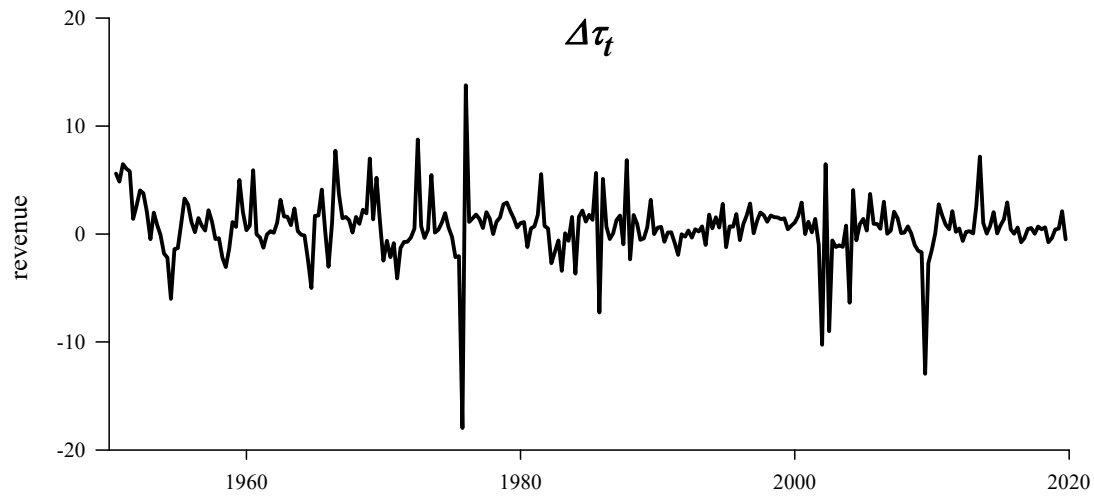
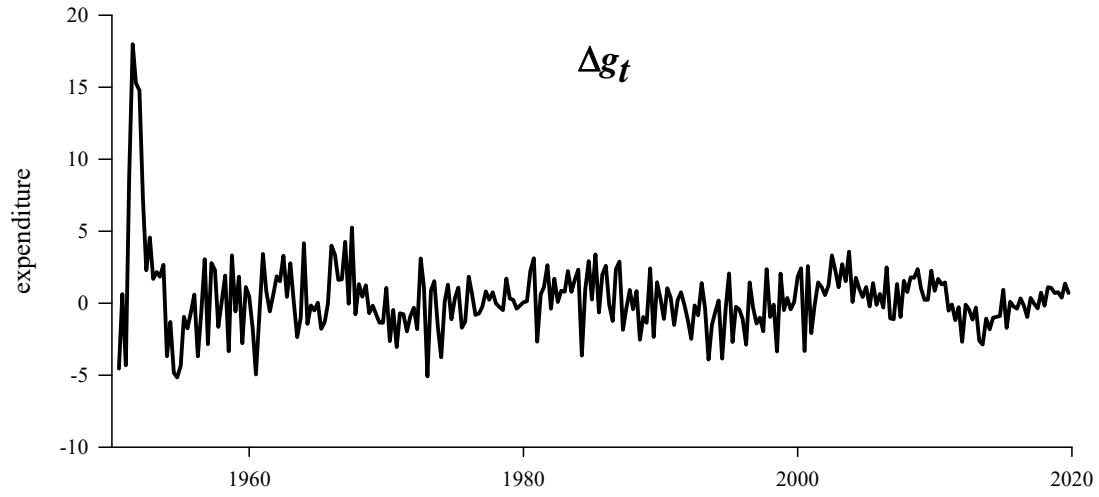
Figure 6: Monte Carlo Densities

Notes: The upper panel shows the density function 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 functions of  $\hat{\Theta}_{12}$ . The lower panel shows the density functions for the correlation between the two estimated structural shocks:  $\rho(\hat{\epsilon}_{1t}, \hat{\epsilon}_{2t})$ . The just-identified case is shown in black. The overidentified case is shown in red. The bandwidth is  $1.06sr^{-0.2}$ , where  $s$  is the standard deviation.

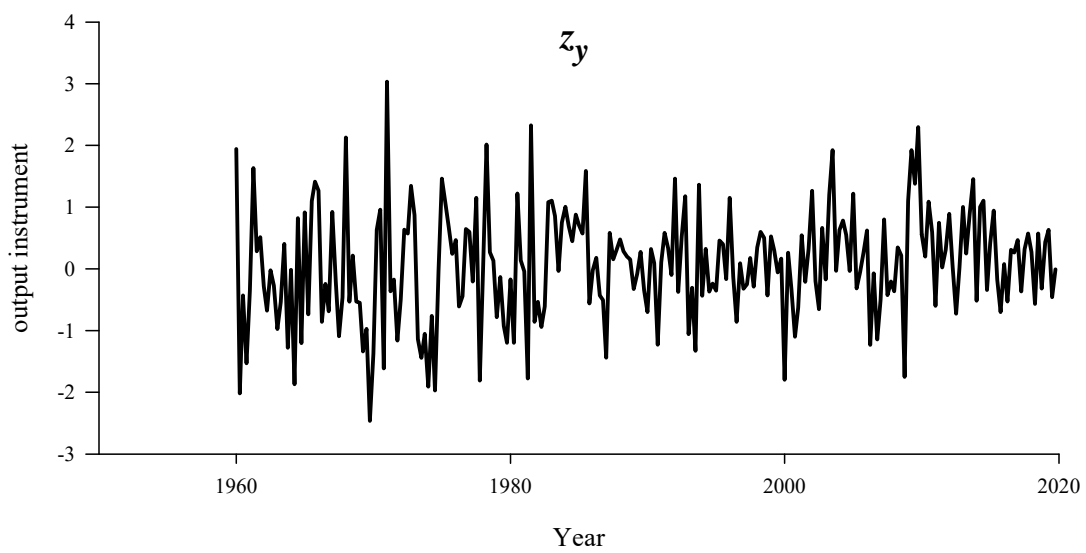
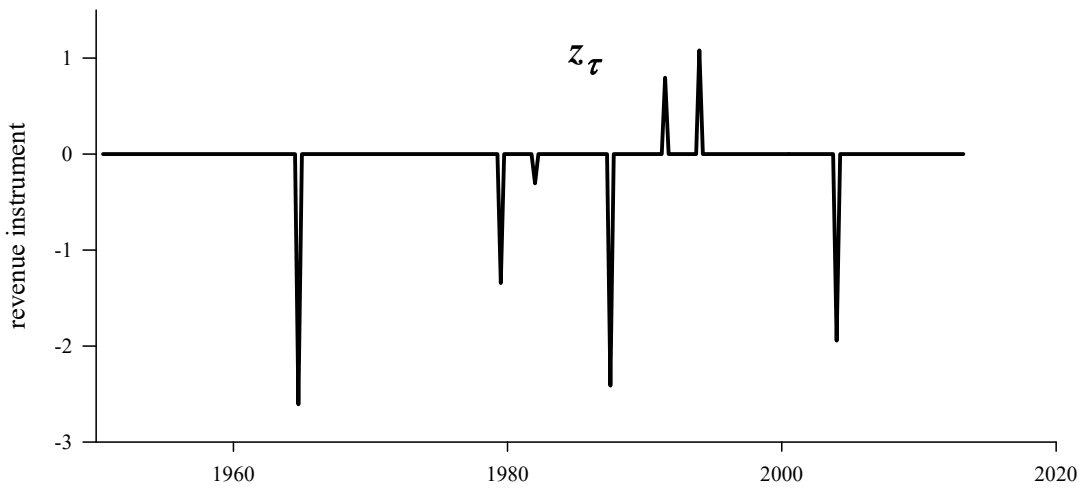
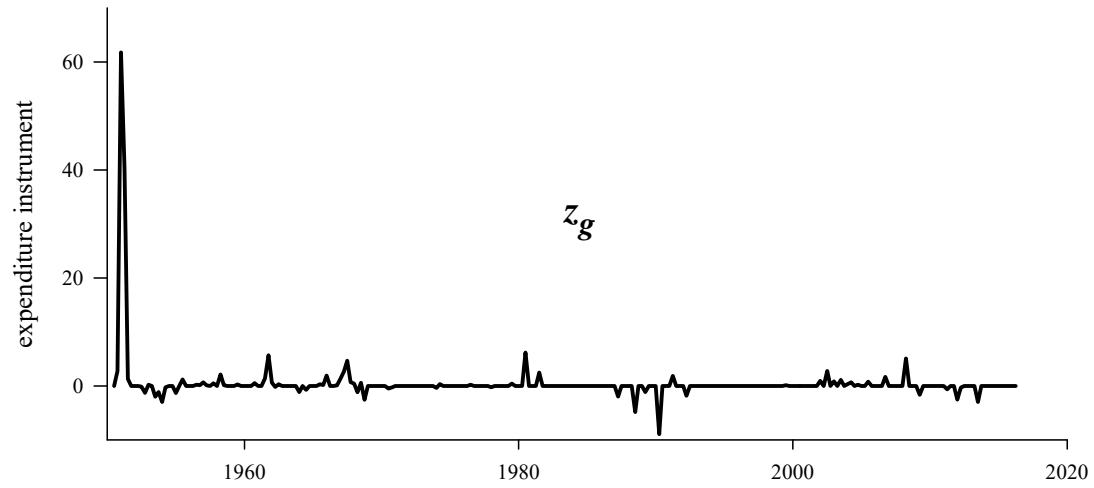
Figure 7: Impulse Response Functions (with Consumption)

Notes: A fourth variable, the growth in real per capita consumption, now is included in the VAR. The top row shows the effects of  $\hat{\epsilon}_g$  on output and consumption. The centre row shows the effects of  $\hat{\epsilon}_\tau$  on output and consumption. The last row also shows a further effect of  $\hat{\epsilon}_\tau$ , on government spending. The lower right panel shows the effect of  $\hat{\epsilon}_c$  on  $y$ . Dotted lines and dashed lines show 68% and 90% asymptotic confidence intervals respectively.

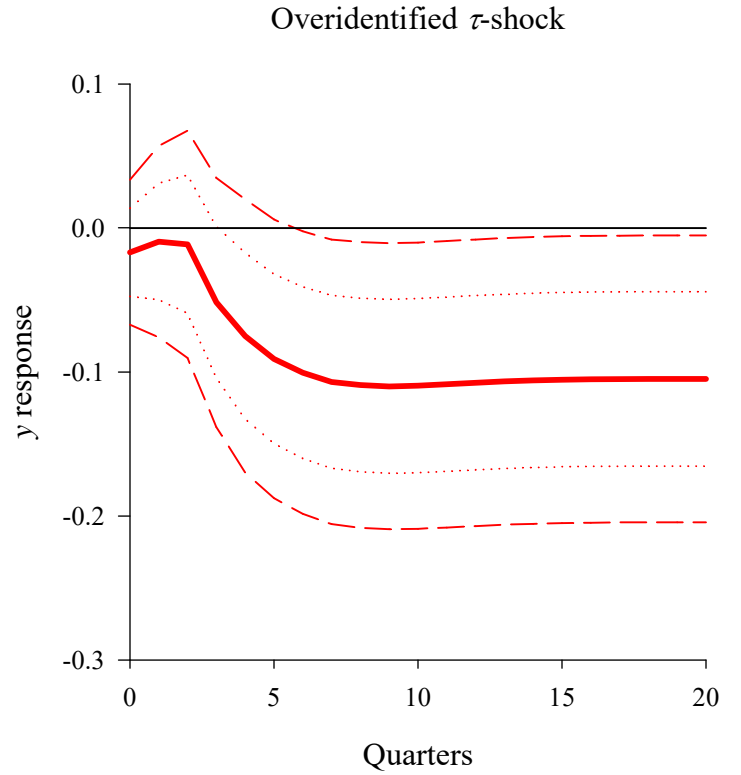
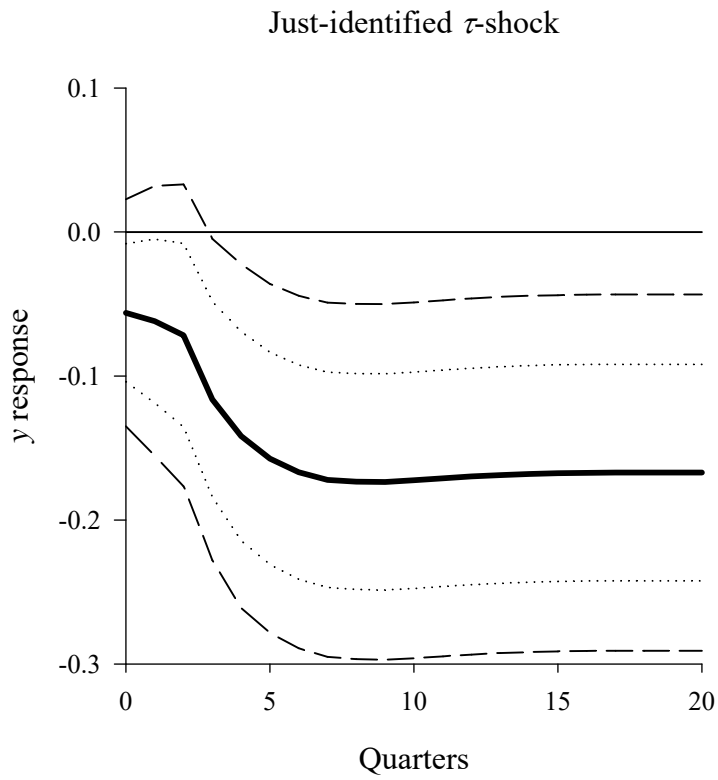
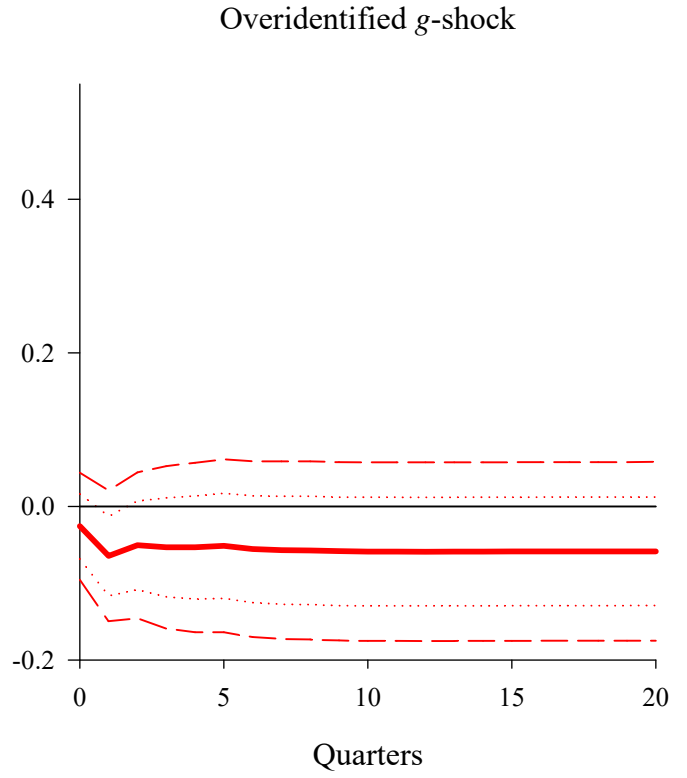
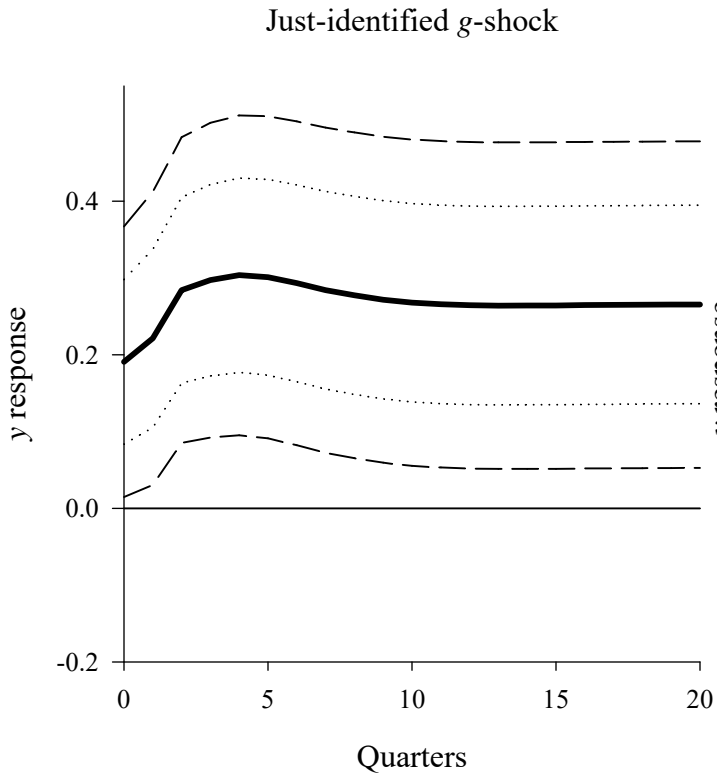
**Figure 1: Growth Rates**



**Figure 2: Instruments**

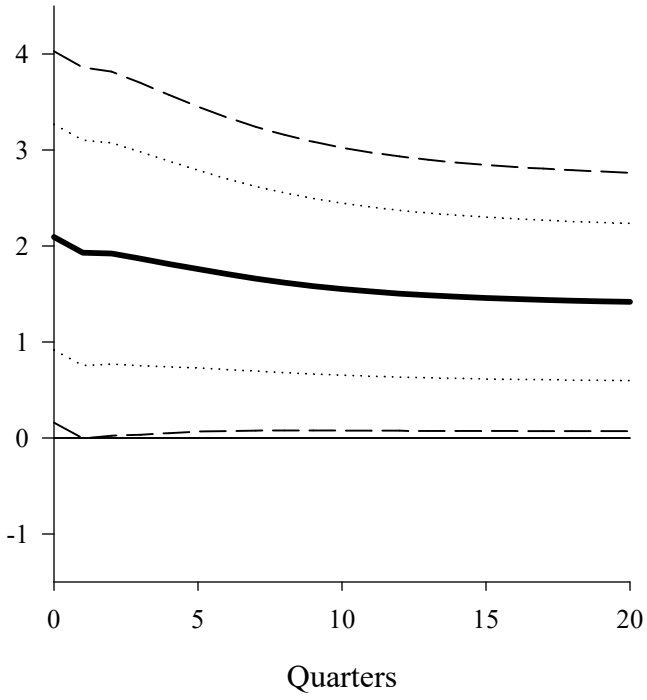


**Figure 3: Impulse Response Functions**

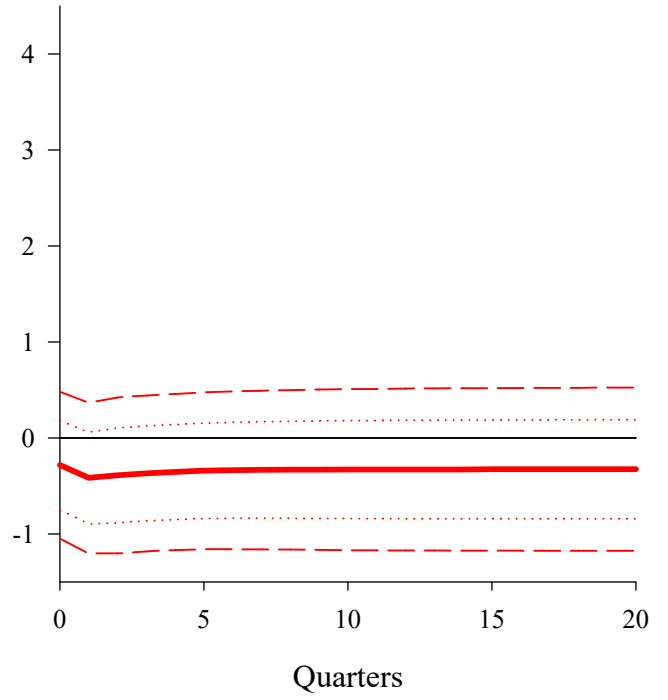


**Figure 4: Multipliers**

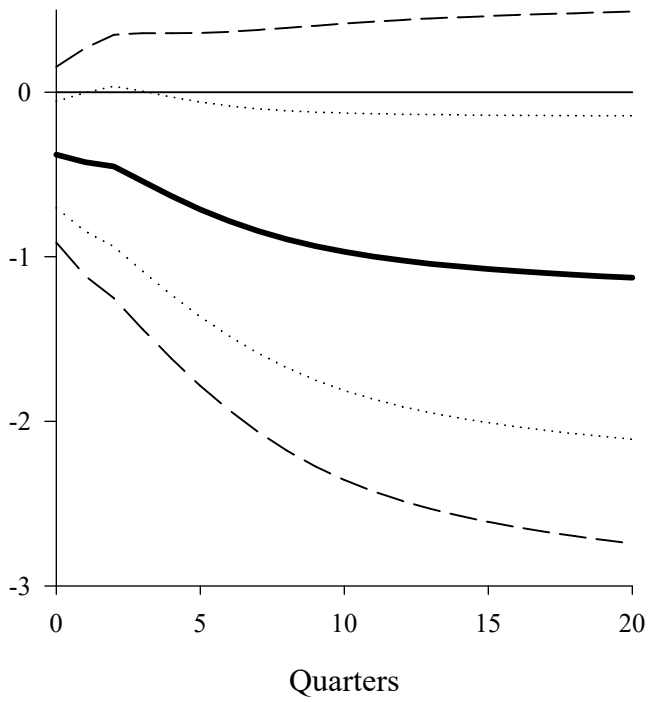
Just-identified  $g$ -shock



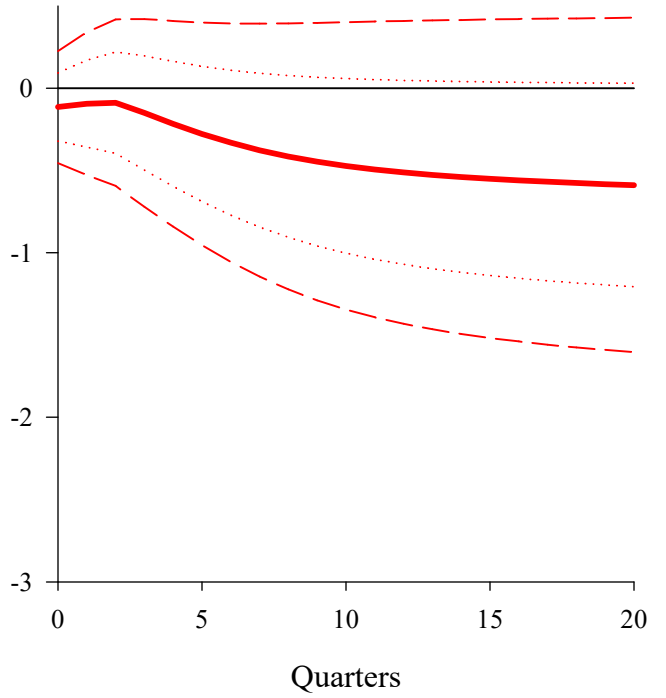
Overidentified  $g$ -shock



Just-identified  $\tau$ -shock

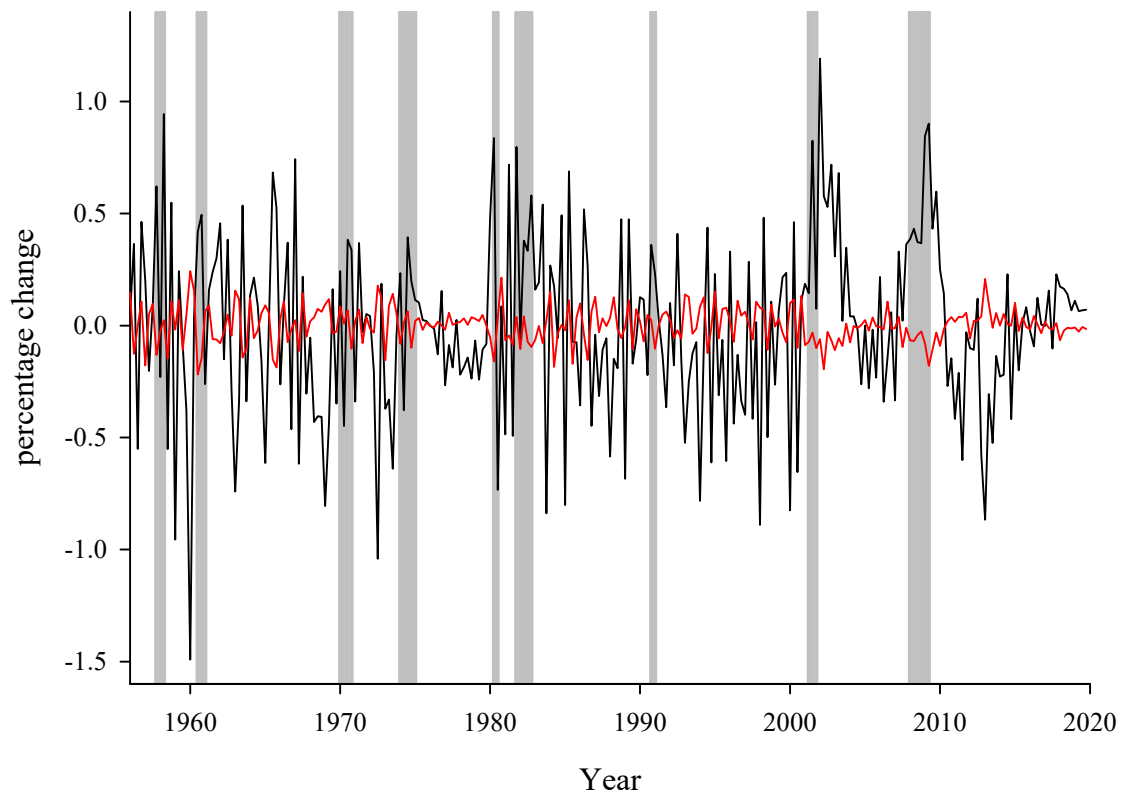


Overidentified  $\tau$ -shock

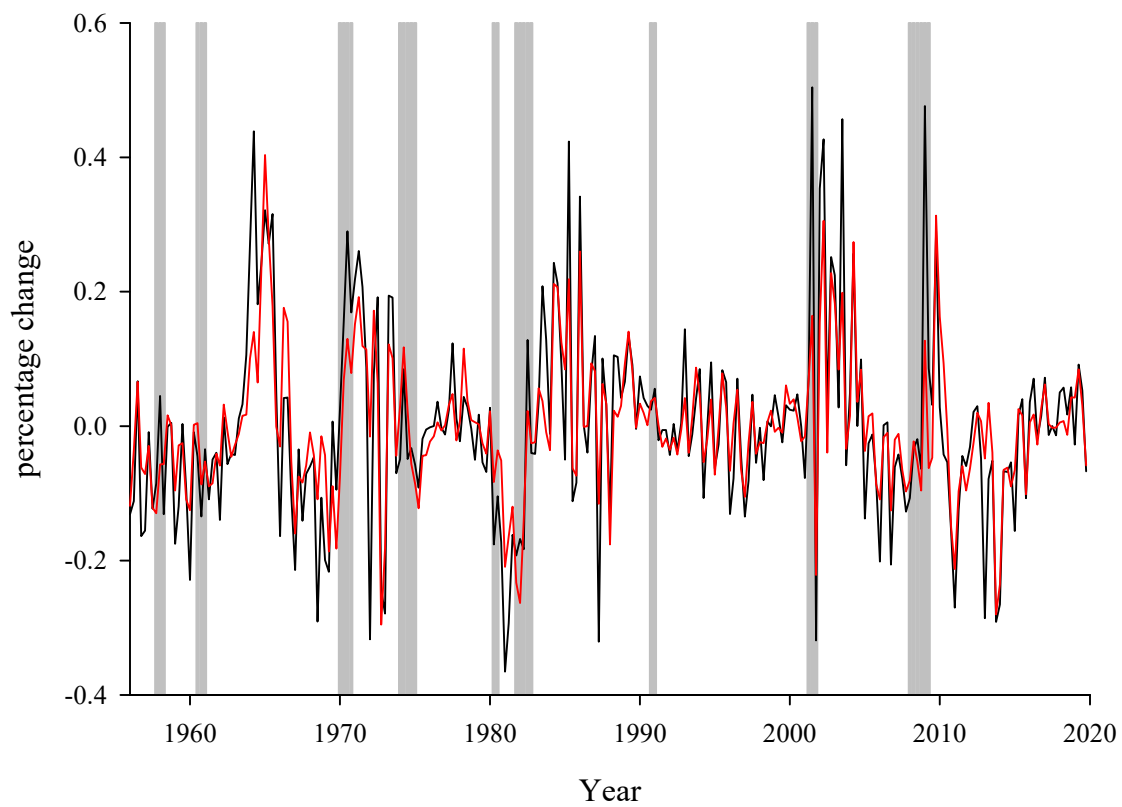


**Figure 5: Historical Decompositions**

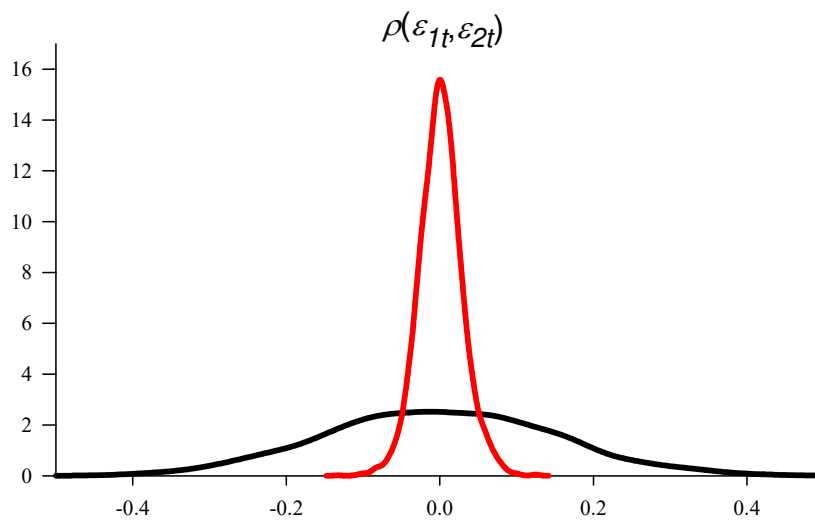
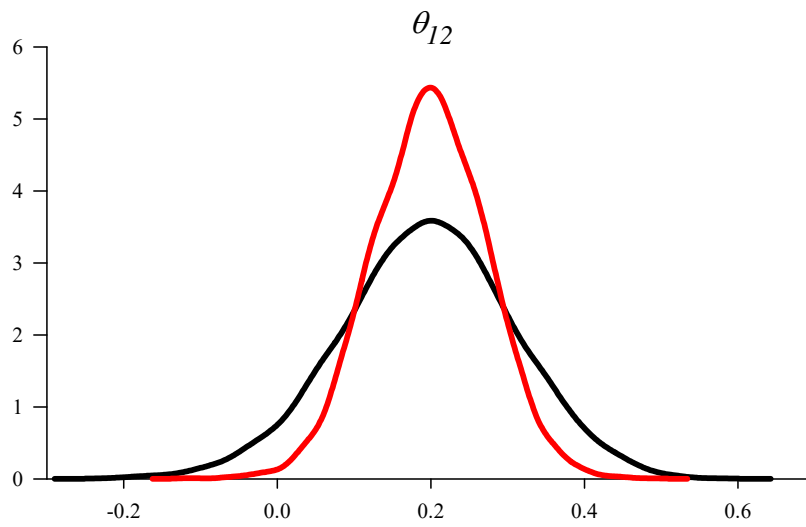
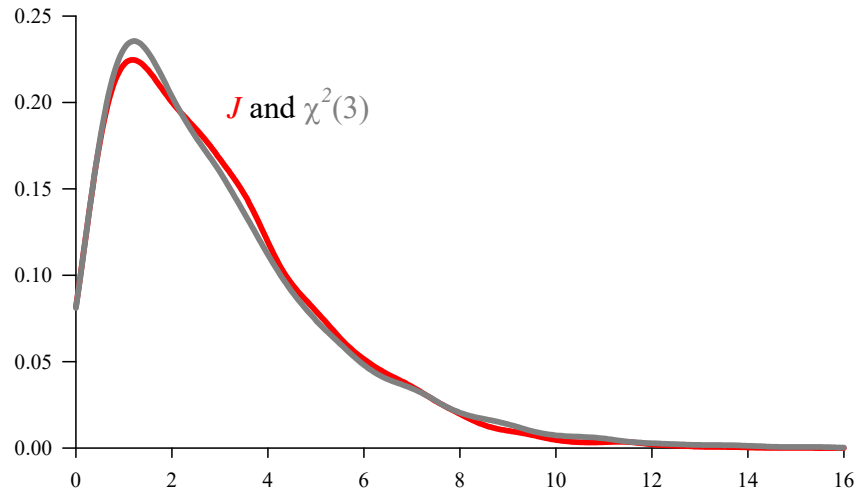
**Response of output to  $g$ -shocks**



**Response of output to  $\tau$ -shocks**



**Figure 6: Monte Carlo Densities**



**Figure 7: Impulse Response Functions (with Consumption)**

