A Primer on the Empirical Identification of Government Spending Shocks*

Kristie M. Engemann  
Federal Reserve Bank of St. Louis  

Michael T. Owyang  
Federal Reserve Bank of St. Louis  

Sarah Zubairy  
Duke University

keywords: fiscal policy, structural VAR, Cholesky decomposition, defense spending

October 25, 2007

Abstract

The empirical literature on the effects of government spending lacks unanimity about the response of consumption and wages. Proponents of shocks identified by structural VARs find results consistent with new Keynesian models. On the other hand, proponents of the narrative approach find results consistent with neoclassical models. We review these two identifications. In addition, we consider some alternative data definitions to evaluate the robustness of the empirical conclusions. Our results confirm the differences across identifications and show them to be, for the most part, robust to alternative spending, output, and labor market data. However, we find some differences between the response to federal and state/local government spending. [JEL: C32, E62]

1 Introduction

Many textbook macroeconomic models contain predictions about the effects of fiscal policy. Unfortunately, this literature lacks unanimity about the response of some variables to surprise increases in government spending. For example, neoclassical models and neo-Keynesian models have opposing predictions regarding the direction of the effect of government spending shocks on consumption and real wages. In a neoclassical model, when a government spending shock hits the economy, households, facing higher taxes, experience a negative wealth effect. They respond by lowering

---

*The authors thank Riccardo DiCecio, Valerie Ramey, and Howard Wall for suggestions and useful conversations. Heidi Beyer-Powe provided research assistance. The opinions expressed herein do not reflect the opinions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.
their consumption and leisure. This increased labor supply from the household also leads to a fall in real wages for any given labor demand. Neo-Keynesian models instead predict that consumption and real wages rise in response to a positive government spending shock. These models often contain features that generate countercyclical markups (e.g., nominal price rigidities or deep habits), which in turn cause labor demand to shift up in response to a government spending shock. This results in rising wages and higher consumption for the households due to substitution effects or the presence of credit-constraints.

Perhaps more distressing is that the empirical literature has been unable to resolve this controversy. Depending on the nature of the identifying assumptions, the empirical literature finds disparate stylized facts regarding the responses of some variables to government spending shocks. The responses of consumption and wages, in particular, can take on different signs depending on the assumptions used to identify fiscal policy shocks. The structural VAR approach of Blanchard and Perotti (2002) and Fatás and Mihov (2001) to identify government spending shocks yields a positive response for output, consumption, and real wages. On the other hand, the narrative approach introduced by Ramey and Shapiro (1998) identifies fiscal policy episodes in the form of military spending that can be considered exogenous to the economy. This approach typically finds that, in response to these large military buildups, output rises but consumption and real wages fall.\footnote{In this growing literature, other researchers have considered the effect of government spending on different economic variables. For example, Tavares and Valkanov (2003) examine the effect on fiscal policy on asset prices and Favero and Giavazzi (2007) examine the effect on interest rates.}

A few recent papers (e.g., Ramey, 2006, and Perotti, 2007) have reenergized this debate regarding the responses of economic variables under different identification schemes. One concern is that the structural VAR approach may not be identifying exogenous innovations to fiscal policy. That is, the timing restrictions used in structural VARs may identify shocks that are anticipated by economic agents. This would confound the econometrician’s ability to disentangle the effects of fiscal policy. This results in responses that are biased by some omitted predictors.\footnote{We may view these biases as similar to the manner in which the price response to monetary shocks exhibits a “price puzzle” when commodity prices are omitted (Sims, 1992). Agents may have information about the economy that would be outside the VAR. Thus, the identified monetary shocks might be mixing exogenous shocks with an endogenous response to omitted variables.} Criticism can also be levied on the narrative approach. This methodology treats all of the “large” fiscal episodes equally rather than allowing for some variation in the size and shape of the response.
In this paper, we review some of the findings in this empirical literature on government spending. In addition, we are interested in distinguishing between shocks to total government spending and disaggregated measures such as federal government spending and state and local government expenditures. A third issue we address here is whether the responses of macro-variables are robust to alternative definitions, e.g., using personal income as a measure of real economic activity instead of real GDP or using employment instead of hours worked. Our findings suggest that the choice of macro-variables in the VAR is important. For instance, we find that employment is more responsive than hours to a government spending shock and that personal income has a smaller response than GDP to a total government spending shock.

The remainder of the paper is organized as follows. Section 2 reviews two of the identification procedures common to the literature on fiscal policy. We first outline the identification of the structural VAR based on timing restrictions. We then consider the dummy variable identification using the Ramey-Shapiro war dates. Section 3 reviews the data used in the estimation. In particular, we discuss differences across the government spending series and we suggest alternative measures of output and the labor market. Section 4 presents the results from the estimation using various identifications and specifications. Section 5 concludes.

2 Model and Identification

Since Sims (1980), the VAR has become a staple in the empirical literature on monetary, technology, and fiscal policy shocks. In each case, the fundamental issue in the use of VARs for policy centers around the assumptions used to transform the reduced-form residuals into structural innovations. With these structural innovations in hand, one can then determine the responses of non-policy variables (i.e., output, prices, etc.) to the shock in question. In this section, we consider two common identifying assumptions used in the empirical literature on fiscal policy: (1) timing restrictions (Blanchard and Perotti, 2002), which assume that innovations to government spending occur prior to the determination of other variables and (2) the use of military buildup dummy variables (Ramey and Shapiro, 1998) as a proxy for exogenous shocks to government spending. \(^3\)

---

\(^3\)Mountford and Uhlig (2005) and Pappa (2005) posit an alternative identifying assumption using sign restrictions. The findings for these identifications are similar to those of the structural VAR literature and are not considered here.
2.1 Timing Restrictions

Consider the following reduced-form $p$-order VAR

\[ y_t = B(L) y_{t-1} + \varepsilon_t, \]  

where $y_t$ is the $n \times 1$ vector of economic variables including government spending, $B(L)$ is a polynomial of lag operators, and $\varepsilon_t \sim N(0, \Sigma)$. The structural representation of the VAR can be written as

\[ A_0 y_t = A(L) y_{t-1} + \nu_t. \]

The objective, then, is to uncover the structural innovations $\nu_t$ defined by an orthonormal rotation of the reduced-form residuals

\[ A_0 \varepsilon_t = \nu_t, \]

where $A_0^{-1} \Omega A_0^{-1'} = \Sigma$, $\nu_t \sim N(0, \Omega)$, and the covariance matrix $\Omega$ of the structural innovations is diagonal. The well-known problem in the literature on structural VARs is that $A_0^{-1} \Omega A_0^{-1'} = \Sigma$ does not define a unique rotation. The matrix $A_0$ contains $n^2$ coefficients which need to be determined to identify a unique rotation. However, the system $A_0^{-1} \Omega A_0^{-1'} = \Sigma$ provides only $\frac{n(n+1)}{2}$ parameters to tie down elements of $A_0$. In order to identify the “true” structural innovations, one must place some restrictions on the system. If restrictions can be placed on $A_0$ itself, one would need only $\frac{n(n-1)}{2}$ binding restrictions for identification.\(^4\) Often, theory does not provide enough assumptions to identify the full complement of parameters in $A_0$. In such cases, one might choose to place fewer restrictions on the system and identify only a particular shock (i.e., a single row of $A_0$).

These restrictions can be of several forms. Exclusion restrictions assume that some variables do not respond contemporaneously to the shock. These restrictions are implemented by setting elements of $A_0$ to zero and generally imply a causal ordering across the variables (i.e., the federal funds rate responds to innovations in output but not vice versa). Sign restrictions define the shock by the direction of the impulse response of certain variables at predefined horizons. For example,\[^4\]Weaker identifying assumptions may require more than $\frac{n(n-1)}{2}$ restrictions (see Paustian, 2007).
contractionary monetary policy innovations raise the federal funds rate and lower inflation at short
horizons.\footnote{See Uhlig (2005) for implementation.} Forecast-error-variance restrictions identify the shocks through their relative power in
explaining fluctuations in certain economic variables. These types of restrictions are particularly
useful if theory indicated the structural innovation should be neutral or dominant at long horizons.\footnote{For example, Faust (1998) assumes long-run neutrality on output to identify monetary shocks. Francis, Owyang, and Roush (2007) assume long-run dominance of technology shocks on productivity to identify technology shocks.}

The timing identification involves the assumption that government spending is determined be-
fore the realizations of output and any other economic variables in quarterly data. Essentially,
this assumption presumes that all other variables have no contemporaneous impact on government
spending. This is accomplished by ordering government spending first in the VAR and identifying
the matrix $A_0^{-1}$ as the Cholesky decomposition of $\Sigma$. The fiscal shock, then, is represented by the
first row of the rotation matrix $A_0$.

\subsection*{2.2 Defense Spending Dummies}

Identification of fiscal policy via a narrative approach is conducted in a similar framework as the
timing restrictions outlined above.\footnote{Ramey and Shapiro (1998) estimate a two-variable version of the fiscal policy dummy. Edelberg, Eichenbaum, and Fisher (1999) extend the analysis to a VAR framework.} In this case, however, the VAR in (1) is augmented by a series
of dummy variables representing the timing of large military buildups:

\[ y_t = B(L)y_{t-1} + C(L)d_{t-1} + \varepsilon_t, \]

where $d_t$ is the period $t$ realization of the military spending dummy having effect on the variables
of interest at, say, $q$ lags and both $B(L)$ and $C(L)$ are lag polynomials of possibly different order.

The effect of fiscal policy is then computed as the impulse response to the military spending dummy.
The $k$--step ahead response to the military shock is defined by the coefficient of $L^k$ in the expansion
of $(I - B(L)L)^{-1}C(L)$.

Figure 1 about here

Ramey and Shapiro (1998) construct the series of military spending dummies much in the way
that Romer and Romer (1994) construct their series of monetary policy innovations. While Romer
and Romer consulted the transcripts from the FOMC meetings, Ramey and Shapiro use historical accounts and *Business Week* to identify periods in which the private sector revised upward their forecasts of future government spending.\(^8\) Given their definition of fiscal spending shocks, four episodes in the post-War period qualify as exogenous shocks to government spending.\(^9\) The bold vertical lines in Figure 1 depict these dates, along with the growth rate in government spending and the NBER-defined recessions for the period 1947:I to 2007:II.\(^10\)

### 3 Data

In addition to assumptions pertaining to identification, the choice of variables to be included in the regression can affect the results. In particular, the choice of the government spending variable can significantly alter the conclusions that can be drawn. In the previous section, we discussed differences in the identification when shocks to government spending are isolated as increases in defense spending. Here, we discuss alternative government spending series. In addition, we outline some of the possible economic data that can be used to measure the effect of fiscal policy.

#### 3.1 Government Spending

Figure 2 about here

Not surprisingly, the effect of an innovation to government spending depends on one’s definition of government spending. The government spending series frequently used in the literature with timing restrictions is real per capita government consumption expenditures and gross investment. This series includes federal, state, and local expenditures. Figure 2 shows how the latter two components have increased in recent years both in levels and as a proportion to total government spending. In addition, the figure depicts the differences in tax and transfer payment policies over time.

---

\(^8\)Romer and Romer (2007) employ a similar “narrative” method to identify fiscal shocks, namely tax shocks. They examine news, speeches of elected officials, and other government documents.


\(^10\)Other papers have used slightly different versions of the war dummies. For example, Burnside, Eichenbaum, and Fisher (2004) relax the restriction of responses to the war dummy in all episodes to have the same size and allow the fiscal episodes to have different intensities.
While this might suggest a rise in the importance of state and local spending, a few caveats must be observed. First, many states and municipalities have balanced budget requirements. While the degree to which these requirements are enforced varies across states (Wagner and Sobel, 2006), this may suggest a different level of fiscal flexibility for state/local versus federal spending. Second, innovations to state and local government spending may be more likely to be anticipated. This is the one of the primary motivations for Ramey and Shapiro’s use of military buildups as a proxy for unanticipated fiscal shocks. Third, shocks to state and local government spending may interfere constructively or destructively with each other if disaggregate business cycles are asynchronous (see Owyang, Piger, and Wall, 2005, and Owyang, Piger, Wall, and Wheeler, 2007, for more on disaggregate business cycles).

In light of these issues, we consider a few alternative government spending series. The first is real per capita government consumption expenditures and gross investment (G1) as the literature suggests. We also consider G1 net of gross investment (per capita real government current expenditures, G2) and G2 net of state and local spending (per capita federal government real current expenditures, G3).

3.2 Economic Variables

We include a number of variables that typically appear in both empirical and theoretical models. These include output, consumption, investment, hours, and the real wage. Real GDP, real non-durables and services consumption, real private fixed investment, and real personal income; all of those series were divided by the GDP implicit price deflator. To obtain per capita numbers, we divided by the total civilian labor force (from the Bureau of Labor Statistics). The nonfarm business sector hours index, the components of the nonfarm business sector real product hourly compensation (compensation per hour and implicit price deflator), nonfarm payroll employment, and nonfarm private employment also come from the Bureau of Labor Statistics. We used the log of all data, which are seasonally adjusted.

The real wage variable used here is the real product wage, which is deflated by its own deflator. The consumption wage, which may be substituted in other studies, is the nominal wage deflated by the GDP deflator.
nondurables and services consumption, and real private fixed investment are expressed in per capita
terms. All four of the latter variables reflect private sector contributions; government contributions
to these variables are either modeled explicitly or are embedded in government spending. We then
perform some additional experiments replacing hours with employment.\textsuperscript{16}

Much of the disparity between the two identification schemes can be highlighted by a compar-
ison of the responses of consumption and the real wage to various identified spending shocks. In
addition to these two responses, we measure the effect of government spending shocks on various
components of output. While most other papers focus on breaking down output into consumption
and investment, we will also consider an alternative decomposition including private personal in-
come (PI), corporate taxes, and government transfers. Simple GDP accounting yields the following:

\[
\text{GDP} = \text{PI} - \text{transfers} + \begin{bmatrix}
\text{corporate taxes} \\
\text{indirect business tax} \\
\text{retained earnings} \\
\text{social security} \\
\text{depreciation}
\end{bmatrix} - \begin{bmatrix}
\text{net income from abroad} \\
\text{net interest}
\end{bmatrix}.
\]

This decomposition may allow us to determine how differences in the composition of the government
spending shock affect components of the output portfolio.

\section{Empirical Results}

In this section, we report the resulting impulse responses to fiscal innovations identified from the
different methods described above. Each empirical model with timing restrictions is a VAR(4)
estimated at a quarterly frequency with the logs of the variables listed. In the dummy-variable
approach, each equation includes four lags of the endogenous variables and the Ramey-Shapiro
dummy variables are entered with lags 0 to 6. The data range from 1954:III to 2007:II.\textsuperscript{17} Each
case varies slightly as we alter the decomposition of output. The point values of the impulse

\textsuperscript{16}Hours and employment have some obvious differences in the nature of their business cycle fluctuations. However, the
hours series may be unavailable for some applications (e.g., transition country analyses).

\textsuperscript{17}For regressions using personal income and the Ramey-Shapiro dummy variables, the sample extends through
2007:II.
responses are accompanied by their corresponding 95 percent confidence intervals.

4.1 Baseline Results

Figure 3: Baseline results about here

Figure 3 depicts the baseline case estimated with GDP, hours, consumption, and real wages. The first three columns of responses correspond to different identification for fiscal spending using the timing restrictions. The first column shows the responses of these variables to a one-standard-deviation increase to per capita real total government expenditures (G1). The second column presents the impulse responses to a similar shock to per capita real total government current expenditures (G2). The third column presents the responses to federal current expenditures in the baseline model. In this case, we are interested in isolating the impact of federal versus state and local spending (G3). The last column depicts the impulse responses to the war dummy using per capita real total government expenditures (G1) as a measure of government spending.

Regardless of the identification, the fiscal policy shock raises the output measure and government spending on impact. The hours and investment responses are essentially negligible at all horizons. The recent literature has focused on the difference between the consumption and real wage responses across identifications. A shock to any of the three spending variables produces a positive response in consumption and a weakly positive response in the real wage. On the other hand, the responses of consumption and the real wage to the war dummy are statistically insignificant but the point estimates are negative on impact.

4.2 Hours v. Employment

The canonical macroeconomic models typically contain predictions about hours. However,

---

18 These results may differ slightly across the literature depending on which investment series is used. Ramey (2006), for example, uses gross domestic investment and finds a stronger response to government spending. Fatás and Mihov (2001) estimate disaggregated investment responses. They find residential investment rises and nonresidential falls in response to government spending shock in a structural VAR.

19 The point estimate of the real wage response is positive but statistically insignificant for the government spending identifications. This result is consistent with other studies.

20 Note that our data sample spans 1954:III–2007:II. Thus, we are omitting the largest post-World War II military buildup episode, the Korean War, from our set of dummy variables. Results, however, are consistent with other studies which include the Korean War episode.
in some cases of interest to econometricians, hours data may be unavailable.\footnote{For example, hours data for some transition countries could be unavailable.} One series that can be used as an alternative to hours data is employment. To demonstrate the similarities and differences between the empirical responses of these two series, we estimate a VAR replacing hours with employment in the baseline specification.\footnote{Using employment instead of hours may change the theoretical conclusions of some macro models. Differences in the predictions might be attributed to differences between labor usage at the extensive versus the intensive margins.} We utilize the same four identifications of fiscal policy and the same set of economic variables described in the previous specification. The responses, of these other variables, as shown in Figure 4, remain qualitatively similar to those from the baseline specification with hours. The response of employment, however, differs slightly from that of hours. While the responses of hours and employment are statistically indistinguishable, employment does seem more responsive to both government spending and the war dummies based on the point response.\footnote{One difference between the two series is that hours measures action at the intensive margin while employment measures the extensive margin (see Fang and Rogerson, 2007).} Also, both employment and hours seem relatively more responsive to federal current government expenditures than the other measures.

Figure 5: Private employment results about here

We are interested in determining whether the sensitivity of employment to government spending is attributable to differences in government versus private hires. We can accomplish this by replacing employment with private employment. These results are shown in Figure 5. Here, the responses of private employment more closely match the responses displayed in Figure 4. It seems that the major component of the weak increase in employment seen in Figure 4 is the rise in private not government hires.

4.3 Personal Income v. GDP

Figure 6: Personal Income results about here

The output measure used in most fiscal policy regressions is GDP. However, in some cases, an equivalent measure of output may not be available at a high frequency (e.g., for U.S. states). One alternative to GDP is personal income. Figure 6 shows the responses of the baseline model in
which GDP has been replaced by real per capita personal income (PI). Of note are the responses to total government spending (G1) and federal spending (G3). In the first case depicted in the first column of Figure 6, the response of PI is qualitatively similar to GDP but shifted downward. That is, the impact response of PI to a shock to total government spending is “smaller” than that for GDP. However, when isolating a shock to federal spending, the impact responses of GDP and PI have similar magnitudes. In response to an innovation in the dummy variable, the response of PI has shape similar to that of GDP. The magnitude of the PI response, however, is muted.

4.4 The Effect of Local Spending

In some of the regressions above, we analyzed the effect of an increase in total government spending and the effect of an increase in federal government spending. The next logical step is to determine the effect of an increase in state and local spending. To accomplish this, we re-estimate the VAR with both federal and the state/local spending, along with the previously mentioned economic variables. The series for state and local spending is the sum of all quarterly state and local spending. The VAR identification orders federal spending first and state spending second. Figure 7 reports the responses of the variables to shocks to both federal and state/local spending.

![Figure 7: Federal and State/Local Separately about here](image)

The impulse responses of the system of variables remain qualitatively similar to the baseline model when federal and local shocks are explicitly modeled. The relevant comparison is the first column here to the third column of Figure 3. We include the response of local spending, which begins weakly negative but rises over time. A notable exception to the similarities is the response of investment. In this case, the response of investment is weakly positive but still insignificant.

The second column reports the response of the system to a shock to the sum of state and local spending. While federal spending rises after a state-and-local shock, GDP, investment, and hours all fall. Consumption is unaffected.

---


25 Obviously, the war dummy identification cannot be reproduced for state and local government spending.

26 In a separate paper, Owyang and Zubairy (2007) consider the effect of state and local spending shocks on their respective regions.

27 This implies that state and local spending may respond to federal spending contemporaneously but not vice versa. We could further assume that state and local spending and federal spending are determined simultaneously.
While looking at the effects of an increase in federal government spending versus the effects of an increase in state and local spending, it is important to consider the composition of the two. The major component of federal spending is defense spending; meanwhile, state and local expenditures are primarily towards education and various public services. Therefore, federal spending shocks might be considered more exogenous and not as prone to being anticipated by the economy. On the other hand, with state and local spending, the shocks might be anticipated a quarter or two before spending actually goes up.

The two different spending aggregates might also be capturing different effects on the economy. Because state and local spending goes mostly towards public services, a positive local spending shock may not have as large a negative wealth effect. Thus, people may not feel as compelled to increase their labor supply, reducing the magnitudes of the responses of hours and output in Figure 7.

5 Conclusions

This paper re-establishes the conflicting predictions of identifying fiscal shocks using structural VARs versus using the Ramey-Shapiro war dates. Both identification schemes have strengths and weaknesses. The narrative approach of Ramey and Shapiro (1998) identifies dates of military buildup and does not require any additional identifying restrictions. However, this approach relies heavily on only four episodes in the post-World War II era. On the other hand, the structural VAR approach raises questions about whether the identified changes in government spending might have been anticipated a few quarters before they actually occur.

We find these conflicting results to be generally robust to changes in the data definitions. Using employment rather than hours produces a stronger labor market response. Using personal income rather than GDP produces a weaker output response. This weaker response is mitigated if the government spending shock focuses on federal expenditures.

Finally, we find that disaggregating the government spending shock may be important. While the responses to federal spending shocks are, for the most part, preserved, shocks to state and local spending produce very different responses. In fact, the response of output to a state/local

\[28\] Only three episodes occur within the data sample that we consider.
spending shock is negative. These results suggest that the state and local shocks may have important compositional and locational differences from the federal shocks.
References


NOTE: Black lines indicate Ramey-Shapiro military buildup dates; gray shaded areas indicate NBER recession dates.
The government spending series used is real government consumption expenditures & gross investment (SAAR).
Figure 2: Government Spending

NOTE: Black lines indicate Ramey-Shapiro military buildup dates; gray shaded areas indicate NBER recession dates.
Figure 3: Impulse Responses from Baseline Model

NOTE: The dotted lines indicate 95% confidence intervals.
Figure 4: Impulse Responses Using Employment Instead of Hours

NOTE: The dotted lines indicate 95% confidence intervals.
Figure 5: Impulse Responses Using Private Employment Instead of Hours

NOTE: The dotted lines indicate 95% confidence intervals.
Figure 6: Impulse Responses Using Personal Income Instead of GDP

NOTE: The dotted lines indicate 95% confidence intervals.
Figure 7: Impulse Responses Using Federal and State/Local Government Spending

Shock to federal gov. sp.

FEDERAL GOV. SPENDING

STATE/LOCAL GOV. SPENDING

GDP

CONSUMPTION

INVESTMENT

HOURS

PRODUCT WAGE

Shock to state/local gov. sp.

FEDERAL GOV. SPENDING

STATE/LOCAL GOV. SPENDING

GDP

CONSUMPTION

INVESTMENT

HOURS

PRODUCT WAGE

NOTE: The dotted lines indicate 95% confidence intervals.