CHAPTER 2

Macroeconomic Shocks and Their Propagation

V.A. Ramey

University of California, San Diego, CA, United States
NBER, Cambridge, MA, United States

Contents

Abstract

This chapter reviews and synthesizes our current understanding of the shocks that drive economic fluctuations. The chapter begins with an illustration of the problem of identifying macroeconomic shocks, followed by an overview of the many recent innovations for identifying shocks. It then reviews in detail three main types of shocks: monetary, fiscal, and technology. After surveying the literature, each section presents new estimates that compare and synthesize key parts of the literature. The penultimate section briefly summarizes a few additional shocks. The final section analyzes the extent to which the leading shock candidates can explain fluctuations in output and hours. It concludes that we are much closer to understanding the shocks that drive economic fluctuations than we were 20 years ago.

Keywords

Macroeconomic shocks, Monetary policy, Fiscal policy, Technology shocks, News, Identification, SVARs, DSGE estimation

JEL Classification Codes

E3, E5, E6

1. INTRODUCTION

At the beginning of the 20th century, economists began to recognize the importance of impulses and propagation mechanisms for explaining business cycle fluctuations. A key question was how to explain regular fluctuations in a model with dampened oscillations. In 1927, the Russian statistician Slutsky published a paper titled "The Summation of Random Causes as a Source of Cyclic Processes." In this paper, Slutsky demonstrated

the surprising result that moving sums of random variables could produce time series that looked very much like the movements of economic time series—"sequences of rising and falling movements, like waves…with marks of certain approximate uniformities and regularities."[a] This insight, developed independently by British mathematician Yule in 1926 and extended by Frisch (1933) in his paper "Propagation Problems and Impulse Problems in Dynamic Economics," revolutionized the study of business cycles. Their insights shifted the focus of research from developing mechanisms to support a metronomic view of business cycles, in which each boom created conditions leading to the next bust, to a search for the sources of the random shocks. Since then, economists have offered numerous candidates for these "random causes," such as crop failures, wars, technological innovation, animal spirits, government actions, and commodity shocks.

Research from the 1940s through the 1970s emphasized fiscal and monetary policy shocks, identified from large-scale econometric models or single equation analyses. The 1980s witnessed two important innovations that fundamentally changed the direction of the research. First, Sims' (1980a) paper "Macroeconomics and Reality" revolutionized the study of systems driven by random impulses by introducing vector autoregressions (VARs). Sims' VARs made the link between innovations to a linear system and macroeconomic shocks. Using his method, it became easier to talk about identification assumptions, to estimate impulse response functions, and to do innovation accounting using forecast error decompositions. The second important innovation was the expansion of the inquiry beyond policy shocks to consider important nonpolicy shocks, such as technology shocks (Kydland and Prescott, 1982).

These innovations led to a flurry of research on shocks and their effects. In his 1994 paper "Shocks," John Cochrane took stock of the state of knowledge at that time by using the by-then standard VAR techniques to conduct a fairly comprehensive search for the shocks that drove economic fluctuations. Surprisingly, he found that none of the popular candidates could account for the bulk of economic fluctuations. He proffered the rather pessimistic possibility that "we will forever remain ignorant of the fundamental causes of economic fluctuations" (Cochrane, 1994, abstract).

Are we destined to remain forever ignorant of the fundamental causes of economic fluctuations? Are Slutsky's "random causes" unknowable? In this chapter, I will summarize the new methodological innovations and what their application has revealed about the propagation of the leading candidates for macroeconomic shocks and their importance in explaining economic fluctuations since Cochrane's speculation.

The chapter progresses as follows. Section 2 begins by defining what a macroeconomic shock is. It then summarizes the many tools used for identifying macroeconomic shocks and computing impulse responses. It also highlights some of the complications and pitfalls, such as the effects of foresight and nonlinearities.

---

[a] Page 105 of the 1937 English version of the article published in *Econometrica*.

The topic of Section 3 is monetary shocks and their effects on the macroeconomy. The section summarizes the existing literature and the challenges to identification. It then explores the effects of several leading monetary shocks in a framework that incorporates some of the newer innovations.

Section 4 discusses fiscal shocks. It begins by summarizing results on government spending shocks and highlights the importance of anticipations. It estimates the effects of several leading identified shocks in a common framework. The second part of the section looks at tax shocks. It summarizes the literature on both unanticipated tax shocks and news about future tax changes and conducts some robustness checks.

Section 5 summarizes the literature on technology shocks, including total factor productivity (TFP) shocks, investment-specific technology (IST) shocks, and marginal efficiency of investment (MEI) shocks. It also discusses news about future technology. It compares a wide variety of identified shocks from the literature.

Section 6 briefly discusses four other candidate shocks: oil shocks, credit shocks, uncertainty shocks, and labor supply (or "wage markup") shocks.

Section 7 concludes by synthesizing what we have learned about shocks. It conducts a combined forecast error variance decomposition for output and hours to determine how much of the fluctuations can be accounted for by some of the leading shocks discussed in the earlier sections. It concludes that we have made substantial progress in understanding the shocks that drive the macroeconomy.

2. METHODS FOR IDENTIFYING SHOCKS AND ESTIMATING IMPULSE RESPONSES

2.1 Overview: What Is a Shock?

What, exactly, are the macroeconomic shocks that we seek to estimate empirically? There is some ambiguity in the literature about the definition because of some researchers' use of the term *shock* when they mean *innovation* (ie, the residuals from a reduced form VAR model) or *instrument*. Sims (1980a) equated innovations with macroeconomic shocks, despite claiming to be atheoretical. Others have used the word *shock* when they mean *instrument* (eg, Cochrane, 2004). In this chapter, I view shocks, VAR innovations, and instruments to be distinct concepts, although *identification assumptions* may equate them in many cases. Shocks are most closely related to the structural disturbances in a simultaneous equation system. I adopt the concept of shocks used by researchers such as Blanchard and Watson (1986), Bernanke (1986), and Stock and Watson (forthcoming). According to Bernanke (1986), the shocks should be *primitive* exogenous forces that are uncorrelated with each other and they should be *economically meaningful* (pp. 52–55).

I view the shocks we seek to estimate as the empirical counterparts to the shocks we discuss in our theories, such as shocks to technology, monetary policy, and fiscal policy. Therefore, the shocks should have the following characteristics: (1) they should be

exogenous with respect to the other current and lagged endogenous variables in the model; (2) they should be uncorrelated with other exogenous shocks; otherwise, we cannot identify the unique causal effects of one exogenous shock relative to another; and (3) they should represent either unanticipated movements in exogenous variables or *news* about future movements in exogenous variables. With regard to condition (2), one might counter with situations in which both fiscal and monetary policies respond to some event and argue that therefore the fiscal and monetary shocks would be correlated. I would respond that these are not primitive shocks, but rather the *endogenous* responses of policies to a primitive shock. A primitive shock may directly enter several of the equations in the system. For example, a geopolitical event might lead to a war that causes both fiscal and monetary policy to respond endogenously. The geopolitical event would be the primitive shock from the standpoint of our economic models (though it might be considered an endogenous response from the standpoint of a political science model).[b]

To match these theoretical shocks, we can link the innovations in a structural vector autoregression (SVAR) to these theoretical (structural) shocks, estimate them in a structural dynamic stochastic general equilibrium (DSGE) model, or measure them directly using rich data sources.

2.2 Illustrative Framework

In this section, I lay out a simple framework in order to discuss the problem of identification and to illustrate some of the leading identification methods. I begin with the problem of identifying shocks to fiscal policy in a simple model with no dynamics. I then generalize the model to a dynamic trivariate model.

Consider first a simple model of the link between fiscal variables and GDP in a static setting. Suppose the structural relationships are given by the following equations:

$$\begin{aligned}
\tau_t &= b_{\tau g} g_t + b_{\tau y} y_t + \varepsilon_{\tau t} \\
g_t &= b_{g\tau} \tau_t + b_{gy} y_t + \varepsilon_{gt} \\
y_t &= b_{y\tau} \tau_t + b_{yg} g_t + \varepsilon_{yt}
\end{aligned} \tag{1}$$

where τ is taxes, g is government spending, and y is GDP. The εs are the macroeconomic shocks we seek to identify. We assume that they are uncorrelated and that, in this simple example, each one affects only one equation. $\varepsilon_{\tau t}$ is the tax shock; it might represent legislation resulting from a change in political power. ε_{gt} might capture the sudden outbreak of war, which raises desired military spending. ε_{yt} might capture technological progress. The bs capture the usual interactions. For example, we would expect that government spending would raise output, while taxes would lower it, so $b_{yg} > 0$ and $b_{y\tau} < 0$. Because

---

[b] Of course, the war might be caused by something like rainfall, in which case the primitive shock would be the rainfall. This shock would enter even more equations, such as the equations for government spending, GDP, and productivity.

of automatic stabilizers, however, the fiscal variables might also respond to GDP, ie, $b_{gy} < 0$ and $b_{\tau y} > 0$. This means that a simple regression of GDP on government spending and taxes will not uncover b_{yg} and $b_{y\tau}$ because g_t and τ_t are correlated with the shock to GDP, ε_{yt}. For example, we might observe no correlation between GDP and government spending, but this correlation is consistent both with no structural relationship between GDP and government spending (ie, $b_{yg} = b_{gy} = 0$) and with b_{yg} and b_{gy} being large, but with opposite signs. Without further assumptions or data, we cannot identify either the parameters or the shocks.

Now let us move to a simple trivariate model with three endogenous variables, Y_1, Y_2, and Y_3 in which dynamics are potentially important.[c] In the monetary context, these variables could be industrial production, a price index, and the federal funds rate; in the fiscal context, they could be GDP, government purchases, and tax revenue; and in the technology shock context, they could be labor productivity, hours, and investment. Let $Y_t = [Y_{1t}, Y_{2t}, Y_{3t}]$ be the vector of endogenous variables. Suppose that the dynamic behavior of Y_t is described by the following structural model:

$$Y_t = B(L)Y_t + \Omega \varepsilon_t \qquad (2)$$

where $B(L) = B_0 + \sum_{k=1}^{p} B_k L^k$ and $E[\varepsilon_t \varepsilon_s'] = D$ if $t = s$, and 0 otherwise, where D is a diagonal matrix. The εs are the primitive structural shocks. Since a primitive shock can in principle affect more than one variable, I initially allow Ω to have nonzero off-diagonal elements.

The elements of B_0 are the same as the bs from Eq. (1), with $b_{jj} = 0$. Thus, the easiest way to address the dynamics is to recast the problem in terms of the *innovations* from a reduced form VAR:

$$A(L)Y_t = \eta_t \qquad (3)$$

where $A(L)$ is a polynomial in the lag operator and $A(L) = I - \sum_{k=1}^{p} A_k L^k$. $\eta_t = [\eta_{1t}, \eta_{2t}, \eta_{3t}]$ are the reduced form VAR innovations. We assume that $E[\eta_t] = 0$, $E[\eta_t \eta_t'] = \Sigma_\eta$ and that $E[\eta_t \eta_s'] = 0$ for $s \neq t$. We then can link the innovations η in the reduced form VAR equation (3) to the unobserved structural shocks, ε, in the structural equation (2) as follows:

$$\eta_t = B_0 \eta_t + \Omega \varepsilon_t \quad \text{or} \qquad (4a)$$

$$\eta_t = H \varepsilon_t, \quad \text{where } H = [I - B_0]^{-1}\Omega \qquad (4b)$$

I will now write out the system in Eq. (4a) explicitly in a way that incorporates a commonly used identification assumption and a normalization. These restrictions are (i) Ω is the identity matrix (meaning each shock enters only one equation); and (ii) the structural

[c] See chapter Stock and Watson (forthcoming) in this handbook for a more precise analysis of identification using SVARs.

shocks have unit effect (ie, the diagonal elements of H are unity).[d] The system can then be written as

$$\eta_{1t} = b_{12}\eta_{2t} + b_{13}\eta_{3t} + \varepsilon_{1t}$$
$$\eta_{2t} = b_{21}\eta_{1t} + b_{23}\eta_{3t} + \varepsilon_{2t} \qquad (5)$$
$$\eta_{3t} = b_{31}\eta_{1t} + b_{32}\eta_{2t} + \varepsilon_{3t}$$

This equation is the dynamic equivalent of Eq. (1). The only difference is that instead of writing the structural relationships in terms of the variables such as GDP, government spending, and taxes themselves, we now write them in terms of the reduced form VAR innovations. The interpretations of the bs, however, are the same if the structural relationships depend on contemporaneous interactions.

As discussed at the start of this section, we cannot identify the coefficients or the shocks without more restrictions. We require at least three more restrictions for identification of all three shocks, potentially fewer if we want to identify only one shock. Since a number of the common identification methods depend on contemporaneous restrictions, I will refer to the system of equations in Eq. (5) when discussing them.

2.3 Common Identification Methods

In this section, I briefly overview some of the most common methods for identification. This section is not meant to be comprehensive. See Stock and Watson (forthcoming) for more detailed treatments of the methods I summarize, as well as for a few other methods I do not summarize, such as set identification and identification through heteroscedasticity. I use the term "policy variable" for short, but it should be understood that it can represent any variable from which we want to extract a shock component.

2.3.1 Cholesky Decompositions

The most commonly used identification method in macroeconomics imposes alternative sets of recursive zero restrictions on the contemporaneous coefficients. This method was introduced by Sims (1980a) and is also known as "triangularization." The following are two widely used alternatives:

A. The policy variable does not respond within the period to the other endogenous variables. This could be motivated by decision lags on the part policymakers or other adjustment costs. Let Y_1 be the policy variable and η_1 be its reduced form innovation. Then this scheme involves constraining $b_{12} = b_{13} = 0$ in Eq. (5), which is equivalent to ordering the policy variable first in the Cholesky ordering. For example, Blanchard and Perotti (2002) impose this constraint to identify the shock to government

---

[d] An alternative normalization to (ii) is the assumption that the structural shocks have unit standard deviation (ie, the variances of the εs are unity).

spending; they assume that government spending does not respond to the contemporaneous movements in output or taxes.[e]

B. The other endogenous variables do not respond to the policy shock within the period. This could be motivated by sluggish responses of the other endogenous variables to shocks to the policy variable. This scheme involves constraining $b_{21} = b_{31} = 0$, which is equivalent to ordering the policy variable last in the Cholesky ordering. For example, Bernanke and Blinder (1992) were the first to identify shocks to the federal funds rate as monetary policy shocks and used this type of identification.[f]

Several of the subsequent sections will discuss how these timing assumptions are not as innocuous as they might seem at first glance. For example, forward–looking behavior or superior information on the part of policymakers may invalidate these restrictions.

2.3.2 Other Contemporaneous Restrictions

Another more general approach (that nests the Cholesky decomposition) is what is known as a *structural VAR*, or SVAR, introduced by Blanchard and Watson (1986) and Bernanke (1986). This approach uses either economic theory or outside estimates to constrain parameters. Consider, for example, Blanchard and Perotti's (2002) identification of government spending and net tax shocks. Let Y_1 be net taxes, Y_2 be government spending, and Y_3 be GDP. They identify the shock to government spending using a Cholesky decomposition in which government spending is ordered first (ie, $b_{21} = b_{23} = 0$). They identify exogenous shocks to net taxes by setting $b_{13} = 2.08$, an outside estimate of the cyclical sensitivity of net taxes.[g] These three restrictions are sufficient to identify all of the remaining parameters and hence all three shocks.

2.3.3 Narrative Methods

Narrative methods involve constructing a series from historical documents to identify the reason and/or the quantities associated with a particular change in a variable. Friedman and Schwartz (1963) is the classic example of using historical information to identify policy shocks. Hamilton (1985) and Hoover and Perez (1994) used narrative methods to identify oil shocks. These papers isolated political events that led to disruptions in world

---

[e] To implement this identification using ordinary least squares (OLSs), one would simply regress government spending on p lags of all of the variables in the system and call the residual the government spending shock.

[f] To implement this identification using OLSs, one would regress the federal funds rate on contemporaneous values of the other variables in the system, as well as p lags of all of the variables, and call the residual the monetary policy shock.

[g] One way to implement the tax shock identification is to construct the variable $\eta_1 - 2.08\eta_3$ from the estimated reduced form residuals. One would then regress η_3 on η_1 and η_2, using $\eta_1 - 2.08\eta_3$ as the instrument for η_1. (Note that the assumption that $b_{21} = b_{23} = 0$ identifies η_2 as ε_{2t}, which is uncorrelated with ε_{3t} by assumption) This regression identifies b_{31} and b_{32}. The residual is the estimate of ε_{3t}.

oil markets. Other examples of the use of narrative methods are Poterba's (1986) tax policy announcements, Romer and Romer's (1989, 2004) monetary shock series based on FOMC minutes, Ramey and Shapiro (1998) and Ramey's (2011a) defense news series based on *Business Week* articles, and Romer and Romer's (2010) narrative series of tax changes based on reading legislative documents.

Until recently, these series were used either as exogenous shocks in sets of dynamic single equation regressions or embedded in a Cholesky decomposition. For example, in the framework above, we could set Y_1 to be the narrative series and constrain $b_{12} = b_{13} = 0$. As a later section details, recent innovations have led to additional methods for incorporating these series.

A cautionary note on the potential of narrative series to identify exogenous shocks is in order. Some of the follow-up research has operated on the principle that the narrative alone provides exogeneity. It does not. Shapiro (1994) and Leeper (1997) made this point for monetary policy shocks. Another example is in the fiscal literature. A series on fiscal consolidations, quantified by narrative evidence on the expected size of these consolidations, is not necessarily exogenous. If the series includes fiscal consolidations adopted in response to bad news about the future growth of the economy, the series cannot be used to establish a causal effect of the fiscal consolidation on future output.

2.3.4 High-Frequency Identification

Research by Bagliano and Favero (1999), Kuttner (2001), Cochrane and Piazzesi (2002), Faust et al. (2004), Gürkaynak et al. (2005), Piazzesi and Swanson (2008), Gertler and Karadi (2015), Nakamura and Steinsson (2015), and others has used high-frequency data (such as news announcements around FOMC dates) and the movement of federal funds futures to identify unexpected Fed policy actions. This identification is also based in part on timing, but because the timing is so high frequency (daily or higher), the assumptions are more plausible than those employed at the monthly or quarterly frequency. As I will discuss in the foresight section later, the financial futures data are ideal for ensuring that a shock is unanticipated.

It should be noted, however, that without additional assumptions the unanticipated shock is not necessarily exogenous to the economy. For example, if the implementation does not adequately control for the Fed's private information about the future state of the economy, which might be driving its policy changes, these shocks cannot be used to estimate a causal effect of monetary policy on macroeconomic variables.

2.3.5 External Instruments/Proxy SVARs

The "external instrument," or "proxy SVAR," method is a promising new approach for incorporating external series for identification. This method was developed by Stock and Watson (2008) and extended by Stock and Watson (2012) and Mertens and Ravn (2013). This approach takes advantage of information developed from "outside" the VAR, such

as series based on narrative evidence, shocks from estimated DSGE models, or high-frequency information. The idea is that these external series are noisy measures of the true shock.

Suppose that Z_t represents one of these external series. Then this series is a valid instrument for identifying the shock ε_{1t} if the following two conditions hold:

$$E[Z_t \varepsilon_{1t}] \neq 0 \tag{6a}$$

$$E[Z_t \varepsilon_{it}] = 0, \quad i = 2, 3 \tag{6b}$$

Condition (6a) is the instrument *relevance* condition: the external instrument must be contemporaneously correlated with the structural policy shock. Condition (6b) is the instrument *exogeneity* condition: the external instrument must be contemporaneously uncorrelated with the other structural shocks. If the external instrument satisfies these two conditions, it can be used to identify the shock ε_{1t}.

The procedure is very straightforward and takes place with the following steps[h]:

Step 1: Estimate the reduced form system to obtain estimates of the reduced form residuals, η_t.

Step 2: Regress η_{2t} and η_{3t} on η_{1t} using the external instrument Z_t as the instrument. These regressions yield unbiased estimates of b_{21} and b_{31}. Define the residuals of these regressions to be ν_{2t} and ν_{3t}.

Step 3: Regress η_{1t} on η_{2t} and η_{3t}, using the ν_{2t} and ν_{3t} estimated in Step 2 as the instruments. This yields unbiased estimates of b_{12} and b_{13}.

As an example, Mertens and Ravn (2014) reconcile Romer and Romer's (2010) estimates of the effects of tax shocks with the Blanchard and Perotti (2002) estimates by using the Romer's narrative tax shock series as an external instrument Z to identify the structural tax shock. Thus, they do not need to impose parameter restrictions, such as the cyclical elasticity of taxes to output. As I will discuss in Section 2.4, one can extend this external instrument approach to estimating impulse responses by combining it with Jordà's (2005) method.

2.3.6 Restrictions at Longer Horizons

Rather than constraining the contemporaneous responses, one can instead identify a shock by imposing long-run restrictions. The most common is an infinite horizon long-run restriction, first used by Shapiro and Watson (1988), Blanchard and Quah (1989), and King et al. (1991). Consider the moving average representation of Eq. (3):

$$Y_t = C(L)\eta_t \tag{7}$$

[h] This exposition follows Merten and Ravn (201_, online appendix). See Mertens and Ravn (2013a,b) and the associated online appendices for generalizations to additional external instruments and to larger systems.

where $C(L) = [A(L)]^{-1}$. Combining Eq. (4b) with Eq. (7), we can write the Ys in terms of the structural shocks:

$$Y_t = D(L)\varepsilon_t \tag{8}$$

where $D(L) = C(L)H$. Suppose we wanted to identify a technology shock as the only shock that affects labor productivity in the long run. In this case, Y_1 would be the *growth rate* of labor productivity and the other variables would also be transformed to induce stationary (eg, first-differenced). Letting $D^{ij}(L)$ denote the (i, j) element of the D matrix and $D^{11}(1)$ denote the lag polynomial with $L = 1$, we impose the long-run restriction by setting $D^{12}(1) = 0$ and $D^{13}(1) = 0$. This restriction constrains the unit root in Y_1 to emanate only from the shock that we are calling the technology shock. This is the identification used by Galí (1999).

An equivalent way of imposing this restriction is to use the estimation method suggested by Shapiro and Watson (1988). Let Y_1 denote the first difference of the log of labor productivity and Y_2 and Y_3 be the stationary transformations of two other variables (such as hours). Then, imposing the long-run restriction is equivalent to identifying the error term in the following equation as the technology shock:

$$Y_{1t} = \sum_{j=1}^{p} \omega_{11,j} Y_{1t-j} + \sum_{j=1}^{p-1} \omega_{12,j} \Delta Y_{2t-j} + \sum_{j=1}^{p-1} \omega_{13,j} \Delta Y_{3t-j} + \zeta_t \tag{9}$$

We have imposed the restriction by specifying that only the *first differences* of the other stationary variables enter this equation. Because the current values of those differences might also be affected by the technology shock, and therefore correlated with the error term, we use lags 1 through p of Y_2 and Y_3 as instruments for the terms involving the current and lagged values of those variables. The estimated residual is the identified technology shock. We can then identify the other shocks, if desired, by orthogonalizing the error terms with respect to the technology shock.

This equivalent way of imposing long-run identification restrictions highlights some of the problems that can arise with this method. First, identification depends on the relevance of the instruments. Second, it requires additional identifying restrictions in the form of assumptions about unit roots. If, for example, hours have a unit root, then in order to identify the technology shock one would have to impose that only the second difference of hours entered in Eq. (9).[i]

Another issue is the behavior of infinite horizon restrictions in small samples (eg, Faust and Leeper, 1997). Recently, researchers have introduced new methods that overcome these problems. Building on earlier work by Faust (1998) and Uhlig (2003, 2004), Francis

---

[i] To be clear, all of the Y variables must be trend stationary in this system. If hours have a unit root, then Y_2 must be equal to $\Delta hours_t$, so the constraint in Eq. (9) would take the form $\Delta^2 hours_t$.

et al. (2014) identify the technology shock as the shock that maximizes the forecast error variance share of labor productivity at some finite horizon *h*. A variation by Barsky and Sims (2011) identifies the shock as the one that maximizes the *sum* of the forecast error variances up to some horizon *h*. See those papers for details on how to implement these methods.

2.3.7 Sign Restrictions

A number of authors had noted the circularity in some of the reasoning analyzing VAR specifications in practice. In particular, whether a specification or identification method is deemed "correct" is often judged by whether the impulses they produce are "reasonable," ie, consistent with the researcher's priors. Faust (1998) and Uhlig (1997, 2005) developed a new method to incorporate "reasonableness" without undercutting scientific inquiry by investigating the effects of a shock on variable *Y*, where the shock was identified by sign restrictions on the responses of *other* variables (excluding variable *Y*). Work by Canova and De Nicolo (2002) and Canova and Pina (2005) introduced other variations.

The sign restriction method has been used in many contexts, such as monetary policy, fiscal policy, and technology shocks. Recently, there have been a number of new papers on sign restrictions using Bayesian methods. For example, Arias et al. (2015b) propose methods involving agnostic priors in one dimension and Baumeister and Hamilton (2015) propose methods involving agnostic priors in another dimension. Amir Ahmadi and Uhlig (2015) combine sign restrictions with Bayesian factor-augmented VARs (FAVARs). See Stock and Watson (forthcoming) for more discussion of sign restrictions as an identification method.

2.3.8 Factor-Augmented VARs

A perennial concern in identifying shocks is that the variables included in the VAR do not capture all of the relevant information. The comparison of price responses in monetary VARs with and without commodity prices is one example of the difference a variable exclusion can make. To address this issue more broadly, Bernanke et al. (2005) developed the FAVARs based on earlier dynamic factor models developed by Stock and Watson (2002) and others. The FAVAR, which typically contains over 100 series, has the benefit that it is much more likely to condition on relevant information for identifying shocks. In most implementations, though, it still typically relies on a Cholesky decomposition. Amir Ahmadi and Uhlig's (2015) new method using sign restrictions in Bayesian FAVARs is one of the few examples that does not rely on Cholesky decompositions. One shortcoming of FAVAR methods is that all variables must be transformed to a stationary form, which requires pretesting and its concomitant problems (eg, Elliott, 1998; Gospodinov et al., 2013). See Stock and Watson (forthcoming) for an in depth discussion of dynamic factor models.

2.3.9 Estimated DSGE Models

An entirely different approach to identification is the estimated DSGE model, introduced by Smets and Wouters (2003, 2007). This method involves estimating a fully specified model (a new Keynesian model with many frictions and rigidities in the case of Smets and Wouters) and extracting a full set of implied shocks from those estimates. In the case of Smets and Wouters, many shocks are estimated including technology shocks, monetary shocks, government spending shocks, wage markup shocks, and risk premium shocks. One can then trace out the impulse responses to these shocks as well as do innovation accounting. Other examples of this method appear in work by Justiniano et al. (2010, 2011) and Schmitt-Grohe and Uribe (2012). Christiano et al. (2005) take a different estimation approach by first estimating impulse responses to a monetary shock in a standard SVAR and then estimating the parameters of the DSGE model by matching the impulse responses from the model to those of the data.

These models achieve identification by imposing structure based on theory. It should be noted that identification is less straightforward in these types of models. Work by Canova and Sala (2009), Komunjer and Ng (2011), and others highlights some of the potential problems with identification in DSGE models. On the other hand, this method overcomes some of the potential problems of unrestricted VARs highlighted by Fernandez-Villaverde et al. (2007).

2.4 Estimating Impulse Responses

Suppose that one has identified the economic shock through one of the methods discussed earlier. How do we measure the effects on the endogenous variables of interest? The most common way to estimate the impulse responses to a shock uses nonlinear (at horizons greater than one) functions of the estimated VAR parameters. In particular, estimation of the reduced form system provides the elements of the moving average representation matrix $C(L) = [A(L)]^{-1}$ in Eq. (7) and identification provides the elements of B_0. Recalling that $D(L) = C(L)H$, we write out $D(L) = D_0 + D_1 L + D_2 L^2 + D_3 L^3 + \cdots$, and denoting $D_h = [d_{ijh}]$, we can express the impulse response of variable Y_i at horizon $t + h$ to a shock to ε_{jt} as:

$$\frac{\partial Y_{i,t+h}}{\partial \varepsilon_{j,t}} = d_{ijh} \qquad (10)$$

These d_{ijh} parameters are nonlinear functions of the reduced form VAR parameters.

If the VAR adequately captures the data-generating process, this method is optimal at all horizons. If the VAR is mispecified, however, then the specification errors will be compounded at each horizon. To address this problem, Jordà (2005) introduced a *local projection* method for estimating impulse responses. The comparison between his procedure and the standard procedure has an analogy with direct forecasting vs iterated

forecasting (eg, Marcellino et al., 2006). In the forecasting context, one can forecast future values of a variable using either a horizon–specific regression ("direct" forecasting) or iterating on a one-period ahead estimated model ("iterated" forecasting). Jordà's method is analogous to the direct forecasting, whereas the standard VAR method is analogous to the iterated forecasting method. Chang and Sakata (2007) introduce a related method they call *long autoregression* and show its asymptotic equivalence to Jordà's method.

To see how Jordà's method works, suppose that ε_{1t} has been identified by one of the methods discussed in the previous section. Then, the impulse response of Y_i at horizon h can be estimated from the following single regression:

$$Y_{i,t+h} = \theta_{i,h} \cdot \varepsilon_{1t} + \text{control variables} + \xi_{t+h} \tag{11}$$

$\theta_{i,h}$ is the estimate of the impulse response of Y_i at horizon h to a shock ε_{1t}. The control variables need not include the other Ys as long as ε_{1t} is exogenous to those other Ys. Typically, the control variables include deterministic terms (constant, time trends), lags of the Y_i, and lags of other variables that are necessary to "mop up"; the specification can be chosen using information criteria. One estimates a separate regression for each horizon and the control variables do not necessarily need to be the same for each regression. Note that except for horizon $h=0$, the error term ξ_{t+h} will be serially correlated because it will be a moving average of the forecast errors from t to $t+h$. Thus, the standard errors need to incorporate corrections for serial correlation, such as a Newey–West (1987) correction.

Because the Jordà method for calculating impulse response functions imposes fewer restrictions, the estimates are often less precisely estimated and are sometimes erratic. Nevertheless, this procedure is more robust than standard methods, so it can be very useful as a heuristic check on the standard methods. Moreover, it is much easier to incorporate state dependence with this method (eg, Auerbach and Gorodnichenko, 2013).

One can extend the Jordà method in several ways that incorporates some of the new methodology. First, one can incorporate the advantages of the FAVAR method (see Section 2.3.8) by including estimated factors as control variables. Second, one can merge the insights from the external instrument/proxy SVAR method (see Section 2.3.5). To see this, modify Eq. (11) as follows:

$$Y_{i,t+h} = \theta_{i,h} \cdot Y_{1,t} + \text{control variables} + \zeta_{t+h} \tag{12}$$

where we have replaced the shock ε_{1t} with $Y_{1,t}$. As discussed earlier, an OLS regression of Y_i on Y_1 cannot capture the structural effect if Y_1 is correlated with ζ_{t+h}. We can easily deal with this issue, however, by estimating this equation using the external instrument Z_t as an instrument for $Y_{1,t}$. For example, if Y_i is real output and $Y_{1,t}$ is the federal funds rate, we can use Romer and Romer's (2004) narrative-based monetary shock series as an instrument. As I will discuss later, in some cases there are multiple potential external instruments. We can readily incorporate these in this framework by using multiple

instruments for Y_1. In fact, these overidentifying restrictions can be used to test the restrictions of the model (using a Hansen's J-statistic, for example).

2.5 The Problem of Foresight

The problem of foresight presents serious challenges to, but also opportunities for, the identification of macroeconomic shocks.[j] There are two main foresight problems: (i) foresight on the part of private agents; and (ii) foresight on the part of policymakers. I will discuss each in turn.

It is likely that many changes in policy or other exogenous shocks are anticipated by private agents in advance. For example, Beaudry and Portier (2006) explicitly take into account that news about future technology may have effects today even though it does not show up in current productivity. Ramey (2011a) argues that the results of Ramey and Shapiro (1998) and Blanchard and Perotti (2002) differ because most of the latter's identified shocks to government spending are actually anticipated. Building on work by Hansen and Sargent (1991), Leeper et al. (2013) work out the econometrics of "fiscal foresight" for taxes, showing that foresight can lead to a nonfundamental moving average representation. The growing importance of "forward guidance" in monetary policy means that many changes in policy rates may be anticipated.

Consider the following example, based on Leeper et al. (2013), of a simple growth model with a representative household with log utility over consumption, discount factor β, and a production function $Y_t = A_t K_{t-1}^{\alpha}$, with $\alpha < 1$. The government taxes output Y at a rate τ_t and there are i.i.d. shocks, $\hat{\tau}_t$, to the tax rate relative to its mean τ. Shocks to technology, ε_{At}, are also i.i.d. Suppose that agents potentially receive news in period t of what the tax rate will be in $t + q$, so that $\hat{\tau}_t = \varepsilon_{\tau, t-q}$. If the shocks are unanticipated ($q=0$), the rule for capital accumulation is:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t}$$

which reproduces the well-known result that unanticipated i.i.d. tax rate shocks have no effect on capital accumulation. If the tax rate shock is anticipated two periods in advance ($q=2$), however, then optimal capital accumulation is:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - k\{\varepsilon_{\tau,t-1} + \theta\varepsilon_{\tau,t}\}$$

where $\theta = \alpha\beta(1-\tau) < 1$ and $k = (1-\theta)\dfrac{\tau}{1-\tau}$. Can we uncover the tax shocks by regressing capital on its own lags? No, we cannot. Because $\theta < 1$, this representation is not invertible in the current and past ks; we say that $\{\varepsilon_{\tau,t-j}\}_{j=0}^{\infty}$ is not *fundamental* for

---

[j] The general problem was first recognized and discussed decades ago. For example, Sims (1980a) states: "It is my view, however, that rational expectations is more deeply subversive of identification than has yet been recognized."

$\left\{k_{t-j}\right\}_{j=0}^{\infty}$. If we regress k_t on its own lags and recover the innovations, we would be recovering the discounted sum of tax news observed at date t and earlier, ie, "old" news. Adding lagged taxes to the VAR does not help.

Beaudry et al. (2015) develop a diagnostic to determine whether nonfundamentalness is quantitatively important. They argue that in some cases the nonfundamental representation is close to the fundamental representation.

The second foresight problem is foresight on the part of policymakers. Sometimes policymakers have more information about the state of the economy than private agents. If this is the case, and we do not include that information in the VAR, part of the identified shock may include the endogenous response of policy to expectations about the future path of macroeconomic variables. Consider the "price puzzle" in monetary VARs, meaning that some identified monetary policy shocks imply that a monetary contraction raises prices in the short run. Sims (1992) argued that the "price puzzle" was the result of typical VARs not including all relevant information for forecasting future inflation. Thus, the identified policy shocks included not only the exogenous shocks to policy but also the endogenous policy responses to forecasts of future inflation. In the fiscal context, governments may undertake fiscal consolidations based on private information about declining future growth of potential GDP. If this is not taken into account, then a finding that a fiscal consolidation lowers output growth may be confounding causal effects with foresight effects.

The principal methods for dealing with the problem of foresight are measuring the expectations directly, time series restrictions, or theoretical model restrictions. For example, Beaudry and Portier (2006) extracted news about future technology from stock prices; Ramey (2011a) created a series of news about future government spending by reading *Business Week* and other periodicals; Fisher and Peters (2010) created news about government spending by extracting information from stock returns of defense contractors; Poterba (1986) and Leeper et al. (2012) used information from the spread between federal and municipal bond yields for news about future tax changes; and Mertens and Ravn (2012) decomposed Romer and Romer's (2010) narrative tax series into one series in which implementation was within the quarter (unanticipated) and another series in which implementation was delayed (news). In the monetary shock literature, many papers use high-frequency financial futures prices to decompose the anticipated and unanticipated components of interest rates changes (eg, Rudebusch, 1998; Bagliano and Favero, 1999; Kuttner, 2001; Gürkaynak et al., 2005).

The typical way that news has been incorporated into VARs is by adding the news series to a standard VAR, and ordering it first. Perotti (2011) has called these "EVARs" for "Expectational VARs." Note that in general one cannot use news as an external instrument in Mertens and Ravn's proxy SVAR framework. The presence of foresight invalidates the interpretation of the VAR reduced form residuals as prediction errors,

since the conditioning variables may not span the information set of forward-looking agents (Mertens and Ravn, 2013, 2014).

2.6 The Problem of Trends

Most macroeconomic variables are nonstationary, exhibiting behavior consistent with either deterministic trends or stochastic trends. A key question is how to specify a model when many of the variables may be trending. Sims et al. (1990) demonstrate that even when variables might have stochastic trends and might be cointegrated, the log levels specification will give consistent estimates. While one might be tempted to pretest the variables and impose the unit root and cointegration relationships to gain efficiency, Elliott (1998) shows that such a procedure can lead to large size distortions in theory. More recently, Gospodinov et al. (2013) have demonstrated how large the size distortions can be in practice.

Perhaps the safest method is to estimate the SVAR in log levels (perhaps also including some deterministic trends) as long as the imposition of stationarity is not required for identification. One can then explore whether the imposition of unit roots and cointegration lead to similar results but increase the precision of the estimates. For years, it was common to include a linear time trend in macroeconomic equations. Many analyses now include a broken trend or a quadratic trend to capture features such as the productivity slowdown in 1974 or the effect of the baby boom moving through the macroeconomic variables (eg, Perron, 1989; Francis and Ramey, 2009).

2.7 Some Brief Notes on Nonlinearities

In the previous sections, we have implicitly assumed that the relationships we are trying to capture can be well approximated with linear functions. There are many cases in which we believe that nonlinearities might be important. To name just a few possible nonlinearities, positive shocks might have different effects from negative shocks, effects might not be proportional to the size of the shock, or the effect of a shock might depend on the state of the economy when the shock hits.

A thorough analysis of nonlinearities is beyond the scope of this chapter, so I will mention only three items briefly. First, Koop et al. (1996) provide a very useful analysis of the issues that arise when estimating impulse responses in nonlinear models. Second, if one is interested in estimating state-dependent models, the Jordà (2005) local projection method is a simple way to estimate such a model and calculate impulse response functions. Auerbach and Gorodnichenko (2013) and Ramey and Zubairy (2014) discuss this application and how it relates to another leading method, smooth transition VARs.

The third point is a cautionary note when considering the possibility of asymmetries. Many times researchers posit that only positive, or only negative, shocks matter. For example, in the oil shock literature, it is common to assume that only oil price *increases* matter and to include a variable in the VAR that captures increases but not decreases.

Kilian and Vigfusson (2011) demonstrate the serious biases and faulty inference that can result from this specification. Their explanation is simple. Suppose Y is a linear function of X, where X takes on both negative and positive values. If one imposes the restriction that only positive values matter, one is in essence setting all of the negative values of X to zero. Figure 1 of Kilian and Vigfusson's (2011) paper demonstrates how this procedure that truncates on the X variable produces slope coefficients that are biased upward in magnitude. Thus, one would incorrectly conclude that positive Xs have a greater impact than negative Xs, even when the true relationship is linear. To guard against this faulty inference, one should always make sure that the model nests the linear case when one is testing for asymmetries. If one finds evidence of asymmetries, then one can use Kilian and Vigfusson's (2011) methods for computing the impulse responses correctly.

2.8 DSGE Monte Carlos

Much empirical macroeconomics is linked to testing theoretical models. A question that arises is whether shocks identified in SVARs, often with minimal theoretical restrictions, are capable of capturing the true shocks. Fernandez-Villaverde et al. (2007) study this question by comparing the state-space representation of a theoretical model with the VAR representation. They note that in some instances an invertibility problem can arise and they offer a method to check whether the problem is present.

Erceg et al. (2005) were perhaps the first to subject an SVAR involving long-run restrictions to what I will term a "DSGE Monte Carlo." In particular, they generated artificial data from a calibrated DSGE model and applied SVARS with long restrictions to the data to see if the implied impulse responses matched those of the underlying model.

This method has now been used in several settings. Chari et al. (2008) used this method to argue against SVARs' ability to test the real business cycle (RBC) model, Ramey (2009) used it to show how standard SVARs could be affected by anticipated government spending changes, and Francis et al. (2014) used this method to verify the applicability of their new finite horizon restrictions method. This method seems to be a very useful tool for judging the ability of SVARs to test DSGE models. Of course, like any Monte Carlo, the specification of the model generating the artificial data is all important.

3. MONETARY POLICY SHOCKS

Having discussed the definition of macroeconomic shocks and the leading methods for identifying them, I now turn to the first of the candidate shocks that will be discussed in detail: monetary policy shocks. In this section, I review the main issues and results from the empirical literature seeking to identify and estimate the effects of monetary policy shocks. I begin with a brief overview of the research before and after Christiano et al.'s (1999) *Handbook of Macroeconomics* chapter on the subject. I revisit Christiano,

Eichenbaum and Evan's specification and then focus on two leading types of externally identified monetary policy shocks, Romer and Romer's (2004) narrative/Greenbook shock and Gertler and Karadi's (2015) recent high-frequency identification (HFI) shocks identified using federal funds futures. I focus on these two types of shocks in part because they both imply very similar effects of monetary policy on output, despite using different identification methods and different samples.

Before beginning, it is important to clarify that the "shocks" identified in the monetary shock literature are not always the empirical counterparts to the shocks from our theoretical models, as discussed in Section 2.1. Because monetary policy is typically guided by a rule, most movements in monetary policy instruments are due to the *systematic* component of monetary policy rather than to deviations from that rule.[k] We do not have many good economic theories for what a structural monetary policy shock should be. Other than "random coin flipping," the most frequently discussed source of monetary policy shocks is shifts in central bank preferences, caused by changing weights on inflation vs unemployment in the loss function or by a change in the political power of individuals on the FOMC. A few papers explicitly link the empirically identified shocks to shifts in estimated central bank preferences (eg, Owyang and Ramey, 2004; Lakdawala, 2015), but most treat them as innovations to a Taylor rule, with no discussion of their economic meaning.[1]

If many macroeconomists now believe that monetary policy shocks themselves contribute little to macroeconomic outcomes, why is there such a large literature trying to identify them? The reason is that we want to identify nonsystematic movements in monetary policy so that we can estimate *causal* effects of money on macroeconomic variables. As Sims (1998) argued in his response to Rudebusch's (1998) critique of standard VAR methods, we need instruments in order to identify key structural parameters. Analogous to the supply and demand framework where we need demand shift instruments to identify the parameters of the supply curve, in the monetary policy context we require deviations from the monetary rule to identify the response of the economy to monetary policy. Thus, much of the search for "shocks" to monetary policy is a search for instruments rather than for primitive macroeconomic shocks.

3.1 A Brief History Through 1999

The effect of monetary policy on the economy is one of the most studied empirical questions in all of macroeconomics. The most important early evidence was Friedman and Schwartz's path-breaking 1963 contribution in the form of historical case studies

---

[k] Friedman argued, however, that most fluctuations in monetary instruments before 1960 were due to non-systematic components of monetary policy.

[1] Christiano et al. (1999) discuss a few other possibilities, such as measurement error in preliminary data (pp. 71–73).

and analysis of historical data. The rational expectations revolution of the late 1960s and 1970s highlighted the importance of distinguishing the part of policy that was part of a rule vs shocks to that rule, as well as anticipated vs unanticipated parts of the change in the policy variable. Sims (1972, 1980a,b) developed modern time series methods that allowed for that distinction while investigating the effects of monetary policy. During the 1970s and much of the 1980s, shocks to monetary policy were measured as shocks to the stock of money (eg, Sims, 1972; Barro, 1977, 1978). This early work offered evidence that (i) money was (Granger–) causal for income; and (ii) that fluctuations in the stock of money could explain an important fraction of output fluctuations. Later, however, Sims (1980b) and Litterman and Weiss (1985) discovered that the inclusion of interest rates in the VAR significantly reduced the importance of shocks to the money stock for explaining output, and many concluded that monetary policy was not important for understanding economic fluctuations.[m]

There were two important rebuttals to the notion that monetary policy was not important for understanding fluctuations. The first rebuttal was by Romer and Romer (1989), who developed a narrative series on monetary policy shocks in the spirit of Friedman and Schwarz's (1963) work. Combing through FOMC minutes, they identified dates at which the Federal Reserve "attempted to exert a contractionary influence on the economy in order to reduce inflation" (p. 134). They found that industrial production decreased significantly after one of these "Romer Dates." The Romer and Romer series rapidly gained acceptance as an indicator of monetary policy shocks.[n] A few years later, though, Shapiro (1994) and Leeper (1997) showed that Romer and Romer's dummy variable was, in fact, predictable from lagged values of output (or unemployment) and inflation. Both argued that the narrative method used by Romer and Romer did not adequately separate *exogenous* shocks to monetary policy, necessary for establishing the strength of the causal channel, from the *endogenous* response of monetary policy to the economy.[o]

The second rebuttal to the Sims and Litterman and Weiss argument was by Bernanke and Blinder (1992). Building on an earlier idea by McCallum (1983), Bernanke and Blinder turned the money supply vs interest rate evidence on its head by arguing that interest rates, and in particular the federal funds rate, were *the* key indicators of monetary policy. They showed that both in Granger-causality tests and in variance decompositions of forecast errors, the federal funds rate outperformed both M1 and M2, as well as the 3-month Treasury bill and the 10-month Treasury bond for most variables.

---

[m] Of course, this view was significantly strengthened by Kydland and Prescott's (1982) seminal demonstration that business cycles could be explained with technology shocks.

[n] Boschen and Mills (1995) also extended the Romer and Romer's dummy variables to a more continuous indicator.

[o] See, however, Romer and Romer's (1997) response to Leeper.

The 1990s saw numerous papers that devoted attention to the issue of the correct specification of the monetary policy function. These papers used prior information on the monetary authority's operating procedures to specify the policy function in order to identify correctly the shocks to policy. For example, Christiano and Eichenbaum (1992) used nonborrowed reserves, Strongin (1995) suggested the part of nonborrowed reserves orthogonal to total reserves, and Bernanke and Mihov (1998) generalized these ideas by allowing for regime shifts in the type of monetary instrument that is targeted.[P] Another issue that arose during this period was the "Price Puzzle," a term coined by Eichenbaum (1992) to describe the common result that a contractionary shock to monetary policy appeared to raise the price level in the short-run. Sims (1992) conjectured that the Federal Reserve used more information about future movements in inflation than was commonly included in the VAR. He showed that the price puzzle was substantially reduced if commodity prices, often a harbinger of future inflation, were included in the VAR.

Christiano et al.'s (1999) *Handbook of Macroeconomics* chapter "Monetary policy shocks: What have we learned and to what end?" summarized and explored the implications of many of the 1990s innovations in studying monetary policy shocks. Their benchmark model used a particular form of the Cholesky decomposition in which the first block of variables consisting of output, prices, and commodity prices was assumed not to respond to monetary policy shocks within the quarter (or month). They called this identification assumption the "recursiveness assumption." On the other hand, they allowed contemporaneous values of the first-block variables to affect monetary policy decisions. Perhaps the most important message of the chapter was the robustness of the finding that a contractionary monetary policy shock, whether measured with the federal funds rate or nonborrowed reserves, had significant negative effects on output. On the other hand, the price puzzle continued to pop up in some specifications.

3.2 Some Alternatives to the Standard Model

Not all research on monetary policy shocks has been conducted in the canonical time-invariant linear SVAR model. In this section, I briefly highlight some of the research that generalizes the linear models or uses completely different methods.

3.2.1 Regime-Switching Models

In addition to the switch between interest rate targeting and nonborrowed reserve targeting (discussed by Bernanke and Mihov, 1998), several papers have estimated regime-switching models of monetary policy. The idea in these models is that monetary policy is driven not just by shocks but also by changes in the policy parameters. In an early

---

[P] An important part of this literature was addressed to the "liquidity puzzle," that is, the failure of some measures of money supply shocks to produce a negative short-run correlation between the supply of money and interest rates.

contribution to this literature, Owyang and Ramey (2004) estimated a regime-switching model in which the Fed's preference parameters could switch between "hawk" and "dove" regimes. They found that the onset of a dove regime leads to a steady increase in prices, followed by decline in output after approximately a year. Primiceri (2005) investigated the roles of changes in systematic monetary policy vs shocks to policy in the outcomes in the last 40 years. While he found evidence for changes in systematic monetary policy, he concluded that they are not an important part of the explanation of fluctuations in inflation and output. Sims and Zha (2006a) also considered regime-switching models and found evidence of regime switches that correspond closely to changes in the Fed chairmanship. Nevertheless, they also concluded that changes in monetary policy regimes do not explain much of economic fluctuations.

3.2.2 Time-Varying Effects of Monetary Policy

In their summary of the monetary policy literature in their chapter in the *Handbook of Monetary Economics*, Boivin et al. (2010) focus on time variation in the estimated effects of monetary policy. I refer the reader to their excellent survey for more detail. I will high-light two sets of results that emerge from their estimation of a FAVAR, using the standard Cholesky identification method. First, they confirm some earlier findings that the responses of real GDP were greater in the pre-1979Q3 period than in the post-1984Q1 period.[q] For example, they find that for the earlier period, a 100 basis point increase in the federal funds rate leads to a decline of industrial production of 1.6% troughing at 8 months. In the later period, the same increase in the funds rate leads to a −0.7% decline troughing at 24 months. The second set of results concerns the price puzzle. They find that in a standard VAR the results for prices are very sensitive to the specification. Inclusion of a commodity price index does not resolve the price puzzle, but inclusion of a measure of expected inflation does resolve it in the post-1984:1 period. In contrast, there is no price puzzle in the results from their FAVAR estimation. Boivin et al. (2010) discuss various reasons why the monetary transmission mechanism might have changed, such as changes in the regulatory environment affecting credit and the anchoring of expectations.

    Barakchian and Crowe (2013) estimate many of the standard models, such as by those by Bernanke and Mihov (1998), CEE (1999), Romer and Romer (2004), and Sims and Zha (2006b), splitting the estimation sample in the 1980s and showing that the impulse response functions change dramatically. In particular, most of the specifications estimated from 1988 to 2008 show that a positive shock to the federal funds rate *raises* output and prices in most cases.

    Another source of time variation is state-dependent or sign-dependent effects of monetary shocks on the economy. Cover (1992) was one of the first to present evidence

---

[q] See, for example, Faust (1998), Barth and Ramey (2002), and Boivin and Giannoni (2006).

that negative monetary policy shocks had bigger effects (in absolute value) than positive monetary shocks. Follow-up papers such as by Thoma (1994) and Weise (1999) found similar results. Recent work by Angrist et al. (2013) finds related evidence that monetary policy is more effective in slowing economic activity than it is in stimulating economic activity. Tenreyro and Thwaites (forthcoming) also find that monetary shocks seem to be less powerful during recessions.

Olivei and Tenreyro (2007) estimate important seasonality in the effects of monetary shocks that is well explained by sticky wage models. They find that monetary shocks that take place in the first two quarters of the year have sizeable, but temporary, effects on output, whereas shocks that take place in the third and fourth quarters of the year have little effect on output. They explain these results with evidence on uneven staggering of labor contracts over the year: a shock that hits near the end of the year has little effect because the bulk of wage contracts is reset then, so wages can adjust immediately.

Since fall 2008, the federal funds rate has been near the zero lower bound. Thus, a key question that has arisen is how to measure shocks in light of this nonlinear constraint. Wu and Xia (2016) use a multifactor Shadow Rate Term Structure Model to estimate a shadow federal funds rate. This shadow rate can capture additional features, such as quantitative easing. Wu and Xia find that unconventional monetary policy has a noticeable impact on the macroeconomy.

3.2.3 Historical Case Studies
Given the important impact of Friedman and Schwartz's (1963) case study of monetary policy during the Great Depression, it is surprising that more case studies have not been conducted. Romer and Romer (1989)'s first narrative monetary analysis was designed to be a quasi-case study in the spirit of Friedman and Schwartz. Their dummy variable series was assigned to episodes in which the Fed decided to risk a recession in order to reduce inflation.

Velde (2009) presents one of the most striking case studies of monetary nonneutrality, based on an episode in 1724 France. In that year, the French government cut the money supply three times, resulting in a cumulative drop of 45%! The action was taken for a variety of reasons, such as long-term price targeting and worries that soldiers and creditors of the state were being hurt by the rise in prices during the previous 6 years. Velde finds that while prices on foreign exchange markets adjusted instantly, other prices adjusted slowly and incompletely and industrial output fell by 30%. The circumstances of that episode are unusually clean for a historical case study, so his evidence of monetary nonneutrality is quite compelling.

3.3 Main Identification Challenges
Several parts of Section 2 discussed some of the challenges to identification in general. Here, I review the issues that are particular important for the identification of monetary policy shocks.

3.3.1 The Recursiveness Assumption

A key assumption used by Christiano et al. (1999) was the "recursiveness assumption." Consider the trivariate model from Eq. (5) in the last section, rewritten here for convenience:

$$
\begin{aligned}
\eta_{1t} &= b_{12}\eta_{2t} + b_{13}\eta_{3t} + \varepsilon_{1t} \\
\eta_{2t} &= b_{21}\eta_{1t} + b_{23}\eta_{3t} + \varepsilon_{2t} \\
\eta_{3t} &= b_{31}\eta_{1t} + b_{32}\eta_{2t} + \varepsilon_{3t}
\end{aligned}
\tag{13}
$$

CEE include more than three variables in the system, so we should think of each η_t as representing a block of variables: η_{1t} includes output, a general price index, and a commodities price index; η_{2t} is the federal funds rate; and η_{3t} contains a monetary stock measure such as M1 or M2, nonborrowed reserves, and total reserves. CEE interpret the equation for η_{2t} as the Fed's feedback rule and ε_{2t} as the monetary policy shock. They assume that current values of the η_{1t} enter the Fed's rule, so $b_{21} \neq 0$, but the money stock and reserves do not enter the rule, so $b_{23} = 0$. These are still not enough assumptions to identify the monetary policy shock because if the monetary policy shock can affect output, etc., within the period, η_{1t} will be correlated with ε_{2t} so we cannot use OLSs. CEE thus add the additional recursiveness assumptions that none of the η_{1t} variables (output and prices) is affected by the monetary policy shock or the monetary aggregates within the period, ie, $b_{12} = b_{13} = 0$. In practice, this is just a Cholesky decomposition generalized to blocks of variables. Since CEE focused only on the monetary policy shock, they did not need to make more assumptions to identify shocks within the first and third block.

It is important to emphasize, however, the importance of the recursiveness assumption for identification. All of CEE's results depend on setting $b_{12} = 0$, meaning that output and prices are not allowed to respond to changes in the federal funds rate within the period. Note that this assumption is at odds with some later estimated New Keynesian DSGE models. For example, Smets and Wouters' (2007) estimated model implies that output, hours, and inflation should respond immediately to the monetary policy shock (see figure 6 of their paper). The estimated DSGE model of CEE (2005) does not imply an immediate response, but only because they assume that no agents can react to the monetary policy shock within the period. They make this theoretical assumption because they estimate their model parameters to match the impulse responses of their VARs which identify the monetary policy shock with the recursiveness assumption.

Even research that develops external instruments typically uses the recursiveness assumption. For example, Romer and Romer (2004) develop a new measure of monetary policy shocks using narrative methods and Greenbook forecasts, but when they study the effects on output and prices, they impose the additional constraint that $b_{12} = b_{13} = 0$. They do so because they do not view their estimated shock as being pure, and thus also use the recursivity assumption as "exogeneity insurance." Coibion's (2012) generalization of the Romer and Romer procedures also imposes the constraint. Barakchian and

Crowe (2013) use HFI from fed funds futures, but nevertheless invoke the recursiveness assumption in their VARs. The typical FAVAR models, such as those by Bernanke et al. (2005) and Boivin et al. (2010), use the recursiveness assumption as well.

Some of the few papers that do not use the recursiveness assumption are those that use sign restrictions. Uhlig (1997, 2005), Faust (1998), Faust et al. (2004), Arias et al. (2015a), and Amir Ahmadi and Uhlig (2015) are able to avoid imposing the zero restriction associated with the recursiveness assumption by instead using sign restriction, also known as "set identification" or partial identification. For example, Uhlig (1997, 2005) imposes the restriction that contractionary monetary policy shocks cannot raise prices. Faust et al. (2004) constrain $-0.1 \leq b_{12} \leq 0$ for the output and price equations. The sign restriction papers can often yield confidence sets that imply possibly positive effects of contractionary monetary policy on output (eg, Uhlig, 2005; Faust et al., 2004; Amir Ahmadi and Uhlig, 2015).

In Section 3.4, I will investigate the importance of the recursiveness assumption in more detail.

3.3.2 Foresight Problems

Section 2.5 discussed how two types of foresight could create problems in identifying shocks and their effects. Both types of foresight are particularly important for monetary policy, and significant progress has been made recently both in appreciating their importance and in developing methods for addressing them.

The first type of foresight problem is foresight on the part of policymakers. As an illustration of the problem, suppose the Fed follows a simple policy rule:

$$ff_t = \alpha_1 E_t(\Delta_h y_{t+h}) + \alpha_2 E_t(\Delta_h \pi_{t+h}) + \varepsilon_{ft} \tag{14}$$

where ff is the federal funds rate, y is the log output, and π is the inflation. Δ_h is the change in the variable from t to $t+h$. The Fed sets interest rates based on its expectations of the future path of output and inflation because it is aware of the lags in the effects of monetary policy. I have modeled this simply as expectations of changes from t to $t+h$, but the argument applies for more general notions of expectations about the path.

The usual SVAR specification assumes that the Fed's expectations about the future paths of output and inflation are adequately captured by the current and lagged values of the (typically) few macroeconomic variables included in the SVAR. This is a strong assumption. The idea that identified monetary shocks might be incorrectly mixing systematic responses to the Fed's expectations was first highlighted by Sims (1992), who argued that the price puzzle was due to the Fed observing more information on inflation. He advocated the incorporation of sensitive commodity prices to address the problem. However, this inclusion does not always get rid of the price puzzle. Bernanke et al.'s (2005) FAVARs are another method for incorporating more information. The FAVAR,

which typically contains over one hundred series, has the benefit that it is much more likely to condition on relevant information for identifying shocks. The FAVAR method nonetheless relies on the assumption that linear combinations of publicly available series fully capture the Fed's expectations.

In a 2000 paper, Romer and Romer presented evidence suggesting that the Fed had superior information when constructing inflation forecasts compared to the private sector. To see the problem this asymmetric information presents, rewrite Eq. (14) as:

$$\mathit{ff}_t = \alpha_1 E_t^p(\Delta_h y_{t+h}) + \alpha_2 E_t^p(\Delta_h \pi_{t+h}) + \alpha_1 \left[E_t^f(\Delta_h y_{t+h}) - E_t^p(\Delta_h y_{t+h}) \right]$$
$$+ \alpha_2 \left[E_t^f(\Delta_h \pi_{t+h}) - E_t^p(\Delta_h \pi_{t+h}) \right] + \varepsilon_{ft} \tag{15}$$

In this equation, E_t^p denotes expectations based on private agent information and E_t^f denotes expectations based on the Fed's information. If information is symmetric and publicly available, then the two terms in square brackets will be zero and methods that incorporate sufficient amounts of public information should be able to identify the monetary policy shock ε_{ft} correctly. If, on the other hand, the Fed has superior information, the terms in brackets will not be zero and an SVAR or an FAVAR will produce an incorrectly identified monetary policy shock, $\widetilde{\varepsilon}_{ft}$ that consists of two components, the true shock as well as a component based on the informational superiority of the Fed:

$$\widetilde{\varepsilon}_{ft} = \varepsilon_{ft} + \alpha_1 \left[E_t^f(\Delta_h y_{t+h}) - E_t^p(\Delta_h y_{t+h}) \right] + \alpha_2 \left[E_t^f(\Delta_h \pi_{t+h}) - E_t^p(\Delta_h \pi_{t+h}) \right] \tag{16}$$

Barth and Ramey (2002) suggested that the problem might be corrected by controlling for Fed forecasts in VARs. They augmented their monetary VARs with Greenbook forecasts of inflation and output in order to determine whether controls for the Fed's superior information would make the price puzzle disappear in their early sample from 1965 to 1979. They found that even with the controls for the Greenbook forecasts, the price puzzle was still very strong in this early sample (see figure 7 of their paper).

Romer and Romer (2004) (R&R) combined the use of Greenbook forecasts with narrative methods to construct a new measure of monetary policy shocks. First, they derived a series of *intended* federal funds rate changes during FOMC meetings using narrative methods. Second, in order to separate the endogenous response of policy to information about the economy from the exogenous shock, they regressed the intended funds rate change on the current rate and on the Greenbook forecasts of output growth and inflation over the next two quarters. They then converted the estimated residuals based on the FOMC meeting frequency data to monthly and used them in dynamic regressions for output and other variables. They found very large effects of these shocks on output.

John Cochrane's (2004) NBER EFG discussion of the Romer and Romer paper highlights how their method can identify movements in monetary policy instruments that are exogenous to the error term of the model. If the Greenbook forecast of future GDP growth contains all of the information that the FOMC uses to make its decisions,

then that forecast is a "sufficient statistic." Any movements in the target funds rate that are not predicted by the Greenbook forecast of future *output* can be used as an instrument to identify the causal effect of monetary policy on *output*. Analogously, any movements in the target funds rate that are not predicted by the Greenbook forecast of *inflation* can be used as an instrument to identify the causal effect of monetary policy on *inflation*. The idea is that if the Fed responds to a shock for reasons other than its effect on future output or future inflation, that response can be used as an instrument for output or inflation. Cochrane states the following proposition in his discussion:

> Proposition 1: To measure the effects of monetary policy on output, it is enough that the shock is orthogonal to output forecasts. The shock does not have to be orthogonal to price, exchange rate, or other forecasts. It may be predictable from time t information; it does not have to be a shock to the agent's or the Fed's entire information set (Cochrane, 2004).

Cochrane's conceptualization of the issue of identifying movements in monetary policy that are exogenous to the error term in the equation is an important step forward. Note, however, that what Cochrane calls a "shock" is not the same as the definition of shock that I use in this chapter. Cochrane's notion of a shock is not a primitive structural shock, but rather a useful instrument for estimating the effect of monetary policy on output, etc.

The possibility of asymmetric information between the Fed and the private sector leads to a further complication, though. If the Federal Reserve has superior information, then any action or announcement by the Fed presents a signal extraction problem for private agents. Private agents observe $\widetilde{\varepsilon}_{ft}$ in Eq. (14), but they know that it is composed of the true shock as well as the systematic component of the Fed's rule based on the Fed's informational advantage. The problem can easily be extended to include the possibility that $\widetilde{\varepsilon}_{ft}$ also contains time-varying inflation or output targets that are unobserved by the public. Gürkaynak et al. (2005) argue that their estimated negative effects of an unanticipated rise in the federal funds rate on long-term forward rates can be explained as the response to information revealed by the Fed action about inflation targets.

The second type of foresight problem is news about future policy actions. Campbell et al. (2012) argue that the Fed has been using forward guidance since the early 1990s. This means that many changes in the federal funds rate are in fact *anticipated* in advance. As discussed in Section 2.5 on foresight, foresight about future movements in policy variables can lead to a nonfundamental moving average representation. This would imply that standard VARs typically cannot be used to identify the shocks.

Fortunately, the monetary literature has developed excellent methods for identifying news shocks. As discussed in Section 2.3.4, research by many, such as Kuttner (2001), Cochrane and Piazzesi (2002), Gürkaynak et al. (2005), Piazzesi and Swanson (2008), Barakchian and Crowe (2013), Gertler and Karadi (2015), and Nakamura and

Steinsson (2015), has used the movements of federal funds and other interest rate futures in small windows around FOMC announcements to identify unexpected Fed policy actions. Exploiting the information in interest rate futures is an ideal way to construct "news" series. D'Amico and King (2015) study the effects of anticipated monetary policy by combining information on expectations, as in Campbell et al. (2012), with sign restrictions in an SVAR. In particular, they identify a monetary news shock by restricting the responses of the expected short term rate to move in the opposite direction of expected inflation and expected output.

3.4 Summary of Recent Estimates

Table 1 summarizes some of the main results from the literature on the impact of the identified monetary shock on output, the contribution of monetary shocks to output fluctuations, and whether the price puzzle is present. Rather than trying to be encyclopedic in listing all results, I have chosen leading examples obtained with the various identifying assumptions. In addition, I attempted to standardize the results by normalizing the peak of response of the federal funds rate to 100 basis points; this standardization does not control for differences in persistence of the response as I discuss later.

As Table 1, the standard Christiano et al. (1999) SVAR, the Faust et al. (2004) HFI, Uhlig's (2005) and Amir Ahmadi and Uhlig's (2015) sign restrictions, Smets and Wouters' (2007) estimated DSGE model, and Bernanke et al.'s (2005) FAVAR all produce rather small effects of monetary policy shocks. Also, most are plagued by the price puzzle to greater or lesser degree. On the other hand, Romer and Romer (2004), Coibion (2012), Barakchian–Crowe (2013), and Gertler–Karadi (2015) all find larger impacts of a given shock on output.

I will also summarize briefly the effects on other variables from some of the leading analyses. A particularly comprehensive examination for many variables is conducted by Boivin et al.'s (2010) with their FAVAR. Recall that they obtained different results for the pre- vs post–1980 period. For the period from 1984m1 to 2008m12, they found that a positive shock to the federal funds rate leads to declines in a number of variables, including employment, consumption expenditures, investment, housing starts, and capacity utilization.

3.5 Explorations with Three Types of Monetary Shocks

I now explore the robustness of the effects of monetary policy shocks using some of the new methods introduced in the literature to deal with both the foresight problems and the recursiveness assumption. For reference, I begin by estimating the classic Christiano et al.'s (1999) type of specification and then move on to the Romer and Romer (2004) shock and Gertler and Karadi's (2015) HFI shock.

Table 1 Summary of monetary policy shock effects on output and prices

Paper	Method, sample	Trough effect of 100 basis point funds peak	% of output explained by shock	Price puzzle?
Christiano et al. (1999)–FFR identification	SVAR, 1965q3–1995q3	−0.7% at 8 quarters	4.4% at 2 years	Yes, but very small
Faust et al. (2004)	HFI, 1991m2–2001m7	−0.6% at 10 months		
Romer and Romer (2004)	Narrative/Greenbook 1970m1–1996m12	−4.3% at 24 months	Major part	No, but prices do not change for 22 months
Uhlig (2005)	Sign restrictions, 1965m1–1996m12	Positive, but not statistically different from 0	5–10% at all horizons.	No (by construction)
Bernanke et al. (2005)	FAVAR, 1959m1–2001m7	−0.6% at 18 months	5% at 5 years	Yes
Smets–Wouters (2007)	Estimated DSGE model, 1966Q1–2004Q4	−1.8 at 4 quarter trough	10% at 1 year (trough)	No
Boivin et al. (2010)	FAVAR, 1962m1–1979m9, 1984m1–2008m12	−1.6% at 8 months in early period −0.7% at 24 months in later period		Only in the early period
Coibion (2012)	"Robust" Romer–Romer methods, 1970m1–1996m12	−2% at 18 months	"Medium" part	Yes, sometimes
Barakchian–Crowe (2013)	HFI, Romer hybrid VAR, 1988m12–2008m6	−5% at 23 months	50% at 3 years	Yes
Gertler–Karadi (2015)	HFI-proxy SVAR, 1979m7–2012m6 (1991m1–2012m6 for instruments)	−2.2% at 18 months	?	No
Amir Ahmadi and Uhlig (2015)	Bayesian FAVAR with sign restrictions, 1960m2–2010m6	−1.3% at 9 months	7% at 24 months	No (by construction)

3.5.1 The Christiano et al. (1999) Benchmark

Christiano et al. (1999) presented estimates based on shocks identified using the recursiveness assumption and showed robust results that were generally consistent with conventional views on the effect of monetary shocks. Here I study how the results change when the sample is updated.

I estimate a specification similar to Christiano et al.'s (1999) specification but use Coibion's (2012) macroeconomic variables for the first block. In particular, I use monthly data and include log industrial production, the unemployment rate, the log of the CPI, and the log of a commodity price index in the first block. The second block consists of the federal funds rate. The third block consists of the logs of nonborrowed reserves, total reserves, and M1. Thus, the innovation to the federal funds rate (orthogonal to contemporaneous values of the first block variables and lags of all of the variables) is identified as the monetary policy shock.

Fig. 1 shows the estimated impulses for this SVAR. The solid black line and shaded areas are the point estimates and 90% bootstrap confidence bands for the system estimated over CEE's sample period, 1965m1–1995m6. The responses look like classic effects of monetary policy shocks. The Federal funds rate jumps up temporarily but then falls back to 0 by 6 months. This temporary blip in the funds rate, however, sets off a prolonged recession. Industrial production begins to fall in the next month and troughs 21 months later. Unemployment rises and peaks around 23 months later. Both unemployment and industrial production return to normal after 4 years. Prices rise slightly for the first few months, but then follow a steady path down, settling at the new lower level after 4 years. Nonborrowed reserves, total reserves, and M1 fall and then recover after 3 years. Nonborrowed reserves display some unusual oscillations, though. For the most part, these responses look very similar to the ones shown in figure 3 of CEE (1999).

The blue short dashed lines in the same figures show the responses for the sample from 1983m1 to 2007m12. The sample stops in 2007 both to exclude the financial crisis and for the practical reason that nonborrowed reserves became negative starting in 2008. The results change dramatically and imply that increases in the federal funds rate lower the unemployment rate. These results echo those of Barakchian and Crowe (2013), who show that the leading specifications imply expansionary effects in the sample from 1988 through 2007.

The red long-dashed lines show the results of a simplified model for the sample 1983m1–2007m12. This model omits M1, nonborrowed reserves, and total reserves. In this specification, there is still a small amount of expansionary effect on output and unemployment at the beginning, but then the more standard contractionary effects take hold. Prices never fall, however.

Table 2 shows various measures of the importance of monetary policy shocks for industrial production in CEE's specification. The first column shows the trough effect on output of a shock that raises the funds rate to a peak of 100 basis points. Even in CEE's original sample, the effects are very modest, less than a −0.5% fall. When estimated over the period
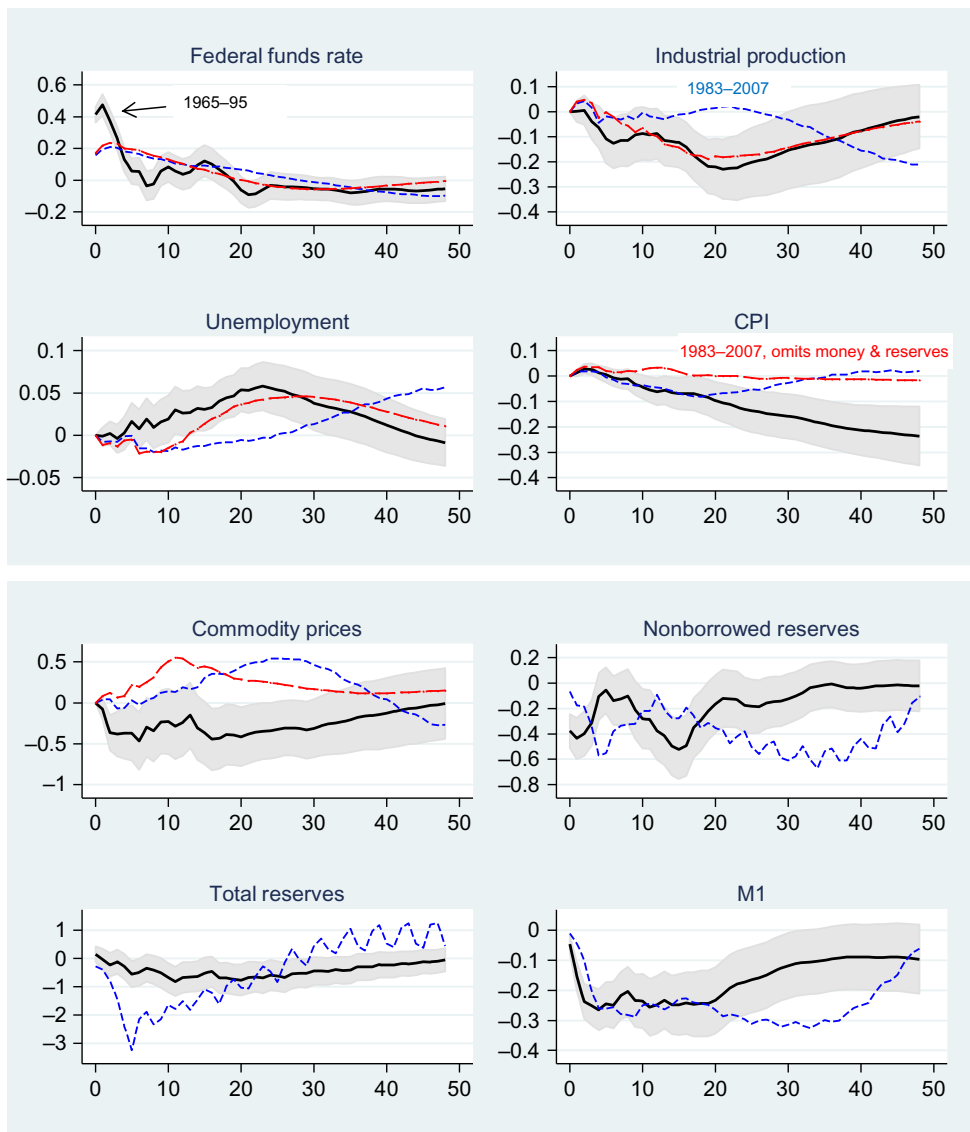
Fig. 1 Christiano et al. (1999) identification. 1965m1–1995m6 full specification: *solid black lines*; 1983m1–2007m12 full specification: *short dashed blue* (*dark gray* in the print version) *lines*; 1983m1–2007m12, omits money and reserves: *long-dashed red* (*gray* in the print version) *lines*. Light gray bands are 90% confidence bands.

1959 through 2007, the effects are less than half. The other columns show the forecast error variance decompositions at 24 months. These indicate that monetary policy shocks account for less than 7% of the forecast error variance in the original sample and less than 1% in the longer sample. A reasonable interpretation of the decline in the contribution of monetary shocks to output volatility is improved, less erratic monetary policy.

Table 2 Effects of Monetary Policy Shocks on Industrial Production: My Estimates

Method and sample	Trough effect of 100 basis point funds peak (%)	Forecast error variance decompositions (%) 24 months
CEE: 1965m1–1996m6	−0.48	6.6
CEE: 1959m1–2007m12	−0.20	0.5
R&R VAR: 1969m3–1996m12	−1.38	8.8
R&R VAR: 1969m3–2007m12	−0.83	2.7
R&R, Jordà method: 1969m3–1996m12	−0.83	
R&R, Jordà method, no recursiveness assumption: 1969m3–1996m12	−0.90	
Gertler–Karadi, proxy SVAR: 1990m1–2012m6	−2.2	
Gertler–Karadi Jordà method: 1990m1–2012m6	−1, but then rises to +4	

Notes: See text for the description of the CEE (Christiano et al., 1999), R&R (Romer and Romer, 2004) VARs, and Gertler–Karadi proxy SVARs.

3.5.2 Greenbook/Narrative Identification of Shocks

Next, I consider the effects of the shocks identified by Romer and Romer (2004). As discussed in Section 3.3.2, Romer and Romer (2004) (R&R) sought to overcome the problem of superior Federal Reserve information by regressing the federal funds target rate on the Greenbook forecasts at each FOMC meeting date and using the residual as the monetary policy shock. They find much larger effects of monetary policy than CEE do. Coibion (2012) explores many possible reasons for the differences and provides very satisfactory and revealing answers. In particular, he finds that R&R's main results, based on measuring the effect of their identified shock using a single dynamic equation, are very sensitive to the inclusion of the period of nonborrowed reserves targeting, 1979–82, and the number of lags (the estimated impact on output is monotonically increasing in the number of lags included in the specification). In addition, their large effects on output are linked to the more persistent effects of their shock on the funds rate. In contrast, R&R's hybrid VAR specification, in which they substituted their (cumulative) shocks for the federal funds rate (ordered last) in a standard VAR, produces results implying that monetary policy shocks have "medium" effects. Coibion (2012) goes on to show that the hybrid model results are consistent with numerous other specifications, such as GARCH estimates of Taylor Rules (as suggested by Hamilton, 2010; Sims–Zha, 2006a) and time-varying parameter models as in Boivin (2006) and Coibion and Gorodnichenko (2011). Thus, he concludes that monetary policy shocks have "medium" effects. In particular, a 100 basis point rise in the federal funds rate leads industrial production to fall 2–3% at its trough at around 18 months.

In my explorations, I use the Coibion version of the R&R hybrid VAR, a monthly VAR with the log of industrial production, the unemployment rate, the log of the CPI, the log of a commodity price index in the first block, and the cumulative Romer and Romer shock replacing the federal funds rate. This specification uses the recursiveness assumption as well, placing the funds rate last in the ordering and thus assuming that the monetary shock cannot affect the macroeconomic variables within the month.

Fig. 2A shows the results, with the solid lines and confidence bands estimated using the original R&R shocks on the original R&R sample of 1969m3–1996m12. These results also match the classic monetary policy effects. Output falls within a month or two, while unemployment rises. Prices remain constant until around 9 months, when they fall steadily until they bottom out during the 4th year after the shock. A qualitative difference with the CEE results is that the response of the federal funds rate is more persistent in the R&R VAR.

The short dashed blue lines show the responses based on the sample from 1983m1 to 2007m12. I constructed new R&R shocks by reestimating their FOMC meeting regression for the later sample, using the updates by Wieland and Yang (2015). I converted these shocks to monthly and then used them in the VAR estimated over the same later sample. The results are similar to those found by Barakchian and Crowe (2013): contractionary monetary policy shocks appear to be expansionary.

The long-dashed red lines show the responses based on the sample from 1969m3 to 2007m12. The R&R shock is based on reestimating their FOMC meeting regression for the entire sample. The results for the full sample look more like those for the original R&R sample, but with more muted effects on output and unemployment.

Rows 3 and 4 of Table 2 show the trough effects and variance decompositions for the R&R VAR. In their original sample, the trough effect on output is −1.38, which is substantially larger than the results using CEE.[r] The forecast error variance decomposition implies that monetary policy shocks account for 9% of the variance for horizons at 24 months.[s] As with most monetary shock specifications, however, the effects are considerably less if we include more recent periods in the sample.

An odd feature of the impulse responses shown in Fig. 2A is the robust rebound of industrial production after 2 years of recession. The peak of industrial production at 48 months is the same magnitude as the trough at 13 months. One possible explanation is that

---

[r] My numbers are slightly different from those of Coibion's for the original sample because he normalized by the impact effect on the funds rate rather than the peak response of the funds rate. Emi Nakamura has suggested that it might be better to compare the integral of the output response to the integral of the funds rate response because this measure incorporates persistence. I found that this measure sometimes behaved oddly because of the tendency of some of the variables to oscillate around zero.

[s] Neither Romer and Romer (2004) nor Coibion (2012) conducted forecast error variance decompositions. Their claim of large or "medium" effects was based on comparing the actual paths of output to the predicted paths implied by the estimated monetary policy shocks.
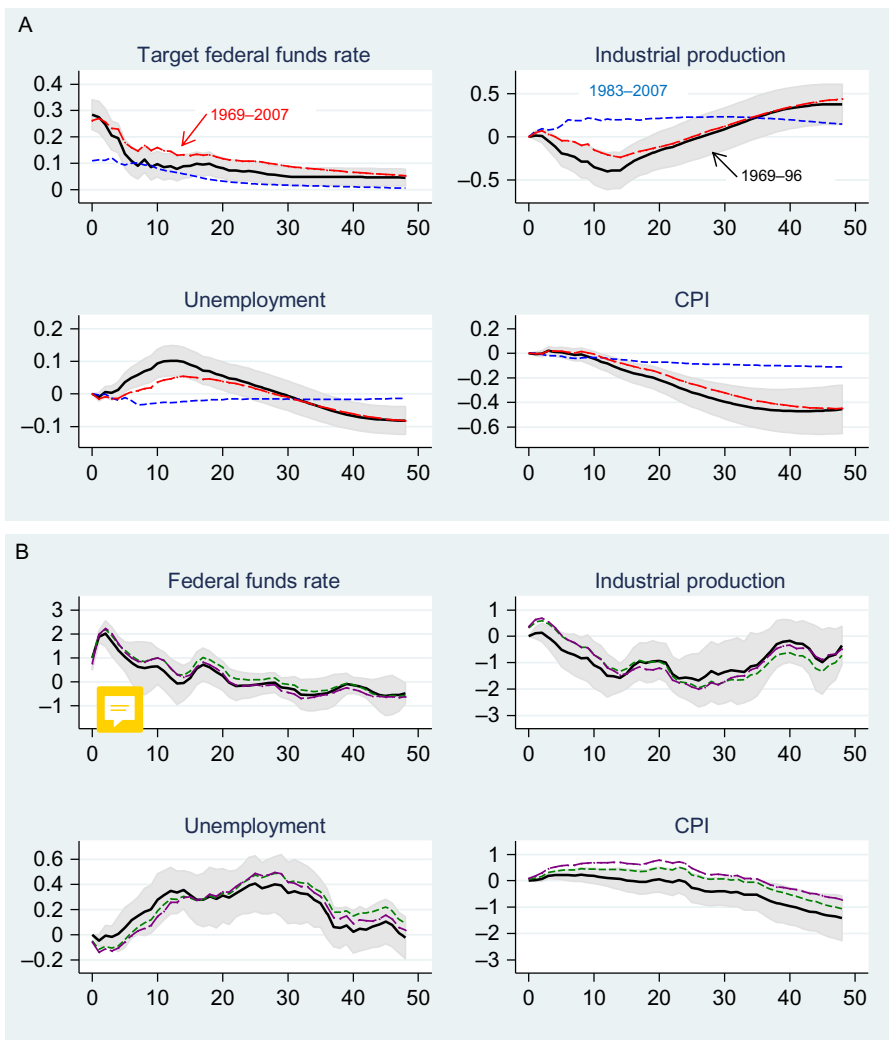
Fig. 2 Romer and Romer monetary shock. (A) Coibion VAR 1969m3–1996m12: *solid black lines*; 1983m1–2007m12: *short dashed blue* (*dark gray* in the print version) *lines*; 1969m3–2007m12: *long-dashed red* (*gray* in the print version) *lines*. (B) Jordà local projection method, 1969m3–1996m12 recursiveness assumption: *solid black lines*; no recursiveness assumption: *short dashed green* (*dark gray* in the print version) *lines*; no recursiveness assumption, FAVAR controls: *long dashed purple* (*dark gray* in the print version) *lines*. Light gray bands are 90% confidence bands.
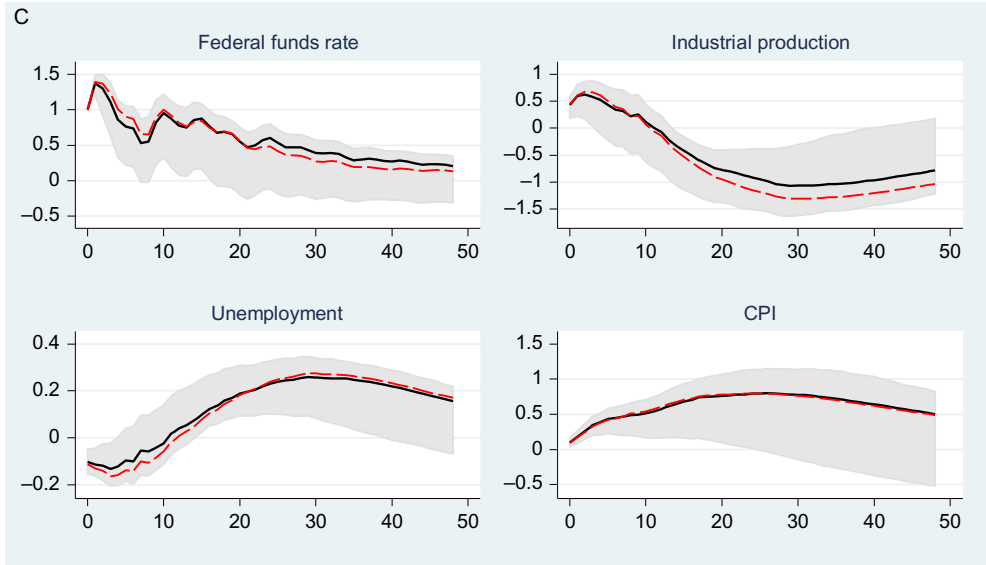
Fig. 2—Cont'd (C) Proxy SVAR, 1969m3–1996m12: *solid black lines*; 1969m3–2007m12: *long-dashed red (gray in the print version) lines.*

misspecification of the VAR is distorting the estimated impulse responses. One way to assess this hypothesis is to use Jordà's (2005) local projection method. As discussed in Section 2.4, the Jordà method puts fewer restrictions on the impulse responses. Rather than estimating impulse responses based on nonlinear functions of the reduced form parameters, the Jordà method estimates regressions of the dependent variable at horizon $t + h$ on the shock in period t and uses the coefficient on the shock as the impulse response estimate.

To investigate the results of this less restrictive specification, I estimate the following series of regressions:

$$z_{t+h} = \alpha_h + \theta_h \cdot \text{shock}_t + \text{control variables} + \varepsilon_{t+h} \tag{17}$$

The z is the variable of interest. The control variables include two lags of the R&R shock, the federal funds rate, the log of industrial production, the unemployment rate, the log of the CPI, and the log of the commodity price index.[t] In addition, to preserve the recursiveness assumption, I include *contemporaneous* values of the log of industrial production, unemployment rate, and the logs of the two price indices. The coefficient θ_h gives the response of z at time $t + h$ to a shock at time t. As discussed in Section 2, ε_{t+h} will be serially correlated, so the standard errors must incorporate a correction, such as Newey–West.

Fig. 2B shows the impulse responses estimated using the Jordà method on R&R's original sample 1969m3–1996m12. The relevant impulse responses are indicated by

---

[t] The point estimates are similar if more lags are included.

the black solid lines. While the responses are somewhat more erratic, they display more coherent dynamics. In particular, rather than the swing from recession to boom for industrial production implied by the VAR estimates, the Jordà estimates imply a more persistent decline in output (and rise in unemployment) that slowly returns to normal.

As discussed earlier, the R&R VAR still imposes the recursiveness assumption. If one believes that the Greenbook forecasts incorporate all relevant information used by the Fed, then one does not need to impose the additional CEE recursiveness assumption. To determine the effect of removing the restriction that output and prices cannot respond to the shock within the month, I reestimate the Jordà regressions omitting the contemporaneous values of all variables other than the R&R shock. The results of this estimation are shown by the green short dashed lines in Fig. 2B. Many aspects of the responses are similar to those obtained with the recursiveness assumptions. However, there are several key differences. First, the estimates imply that a shock that raises the funds rate is initially expansionary: industrial production rises and the unemployment rate falls for the first several months, and the point estimates are statistically different from zero (not shown). Second, there is a pronounced price puzzle for the first 2 years, and most of those estimates are statistically different from zero.

One possible explanation for these puzzles is a failure of the Greenbook forecasts to capture all of the information the Federal Reserve uses. To examine whether expanding the information set helps eliminate the price puzzle, I reestimate the nonrecursive Jordà model with the following alternative controls: two lags each of the R&R shock, the federal funds rate, the dependent variable, and updates of Boivin et al.'s (2010) five FAVAR factors.[u] The results are shown as the long dashed purple (*dark gray* in the print version) line in Fig. 2B. In this case, the price puzzle is even worse and the initial expansionary effects on output and unemployment are no better. Thus, including FAVAR–type factors does not reproduce the results obtained using the recursiveness assumption.

The proxy SVAR is another method that can be used to relax the recursiveness assumption. Kliem and Kriwoluzky (2013) use this method to reconcile VAR monetary shocks with the Romers' narrative shocks. They do not, however, explore effects on output, prices, or other variables. To investigate the results using this method, I estimate the reduced form of Coibion's system with the federal funds rate substituted for the cumulative R&R shock and with R&R's monetary policy shock as an external instrument following Stock and Watson's (2012) and Mertens and Ravn's (2013) proxy SVAR method (see Section 2 for a description).

Fig. 2C shows the results for the original sample (1969–1996) and the full sample from 1969 through 2007. The shaded areas are 90% confidence bands using Mertens and Ravn's methods for the original sample estimates. In both samples, a shock to monetary policy raises the federal funds rate, which peaks at 1.4% by the month after the shock and

---

[u] I am indebted to Shihan Xie for providing her updates of the Boivin et al. estimated factors.

falls slowly to 0% thereafter. The response of industrial production (shown as the red long-dashed line in Fig. 2) is different from the one obtained using the hybrid VAR. In particular, industrial production now rises above normal for about 10 months and then begins falling, hitting a trough at about 29 months. Normalized by the funds rate peak, the results imply that a shock that raises the funds rate to a peak of 100 basis points first raises industrial production by 0.5% at its peak a few months after the shock and then lowers it by −0.9% by 29 months. The unemployment rate exhibits the same pattern in reverse. After a contractionary monetary policy shock, it falls by 0.1 percentage points in the first year and then begins rising, hitting a peak of about 0.2 percentage points at month 30. The behavior of the CPI shows a pronounced, statistically significant prize puzzle.

In sum, relaxing the recursiveness assumption imposed by Romer and Romer's hybrid VAR leads to several puzzles. A contractionary monetary policy shock is now expansionary in its first year and the price puzzle is very pronounced.

The most obvious explanation for these results is that the FOMC responds to more information than even the Greenbook forecast and FAVAR factors capture, and making the R&R shock orthogonal to current output and prices (ie, the recursiveness assumption) helps cleanse the shock of these extra influences. However, this means that even with the R&R shock, one is forced to make the recursiveness assumption, which does not have a solid economic basis. As discussed earlier, leading New Keynesian models, such as Smets and Wouters (2007), imply immediate effects of monetary policy shocks on output and prices.

This exploration highlights the importance of additional restrictions imposed in standard monetary models, as well as the importance of the sample period when estimating the effects of monetary shocks. Without the additional recursiveness assumption, even narrative methods can produce puzzling results. Furthermore, as highlighted by Barakchian and Crowe (2013), many of the methods that produce classic monetary shock results in samples through the mid-1990s produce puzzles when estimated over later samples. In particular, contractionary monetary shocks seem to have expansionary effects in the first year and the price puzzle is pervasive. A plausible explanation for the breakdown in results in the later sample is an identification problem: because monetary policy has been conducted so well in the last several decades, true monetary policy shocks are rare. Thus, it is difficult to extract meaningful monetary shocks that are not contaminated by problems with foresight on the part of the monetary authority.

3.5.3 HFI Shocks

As discussed in the previous sections, numerous papers have used HFI methods to deal with possible foresight about monetary policy changes. Part of the literature focuses only on the effects on interest rates and asset prices (eg, Krishnamurthy and Vissing-Jorgensen, 2011; Hanson and Stein, 2015). Nakamura and Steinsson (2015) link their estimated interest rate changes to output effects by calibrating a New Keynesian model. The

strength of the effect, however, depends crucially on the assumed intertemporal elasticity of substitution. For this reason, it is also important to estimate direct links in the data as well.

A recent paper by Gertler and Karadi (2015) combines HFI methods with proxy SVAR methods to investigate the effects on macroeconomic variables. They have two motivations for using these methods. First, they seek to study the effect of monetary policy on variables measuring financial frictions, such as interest rate spreads. The usual Cholesky ordering with the federal funds rate ordered last imposes the restriction that no variables ordered earlier respond to the funds rate shocks within the period. This is clearly an untenable assumption for financial market rates. Second, they want to capture the fact that over time the Fed has increasingly relied on communication to influence market beliefs about the future path of interest rates (forward guidance).

In the implementation, Gertler and Karadi estimate the residuals using monthly data from 1979 to 2012, but then execute the proxy SVAR from 1991 to 2012 since the instruments are only available for that sample. Fig. 3A replicates the results from Gertler and Karadi's baseline proxy SVAR for figure 1 of their paper.[v] This system uses the 3-month ahead fed funds futures (ff4_tc) as the shock and the 1-year government bond rate as the policy instrument. The other variables included are log of industrial production, log CPI, and the Gilchrist and Zakrajšek (2012) excess bond premium spread. The results show that a shock raises the 1-year rate, significantly lowers industrial production, does little to the CPI for the first year, and raises the excess bond premium. In order to put the results on the same basis as other results, I also estimated the effect of their shock on the funds rate. The results imply that a shock that raises the federal funds rate to a peak of 100 basis points (not shown) lowers industrial production by about -2%.

I explore the robustness of the results by estimating the effects of their shocks in a Jordà local projection framework. The control variables are two lags of the shock itself, the interest rate on 1-year government bonds, industrial production, the CPI, and the Gilchrist–Zakrajsek (2012) excess bond premium spread. I do not include current values of these other variables, so I am not imposing the recursiveness assumption.

Fig. 3B shows the results. The impulse responses look very different from those using the proxy SVAR method. The interest rate rises more slowly, but then remains high for a much longer time. Output does not respond for a year, but then rises. Prices respond little for the first 30 months, but then finally fall.

I then conducted some further investigations of the Gertler-Karadi shock. Several features emerge. First, the shock is not zero mean. The mean is -0.013 and is statistically different from zero. Second, it is serially correlated; if I regress it on its lagged value, the coefficient is 0.31 with a robust standard error of 0.11. This is not a good feature since it is

---

[v] The only difference is that I used 90% confidence intervals to be consistent with my other graphs.
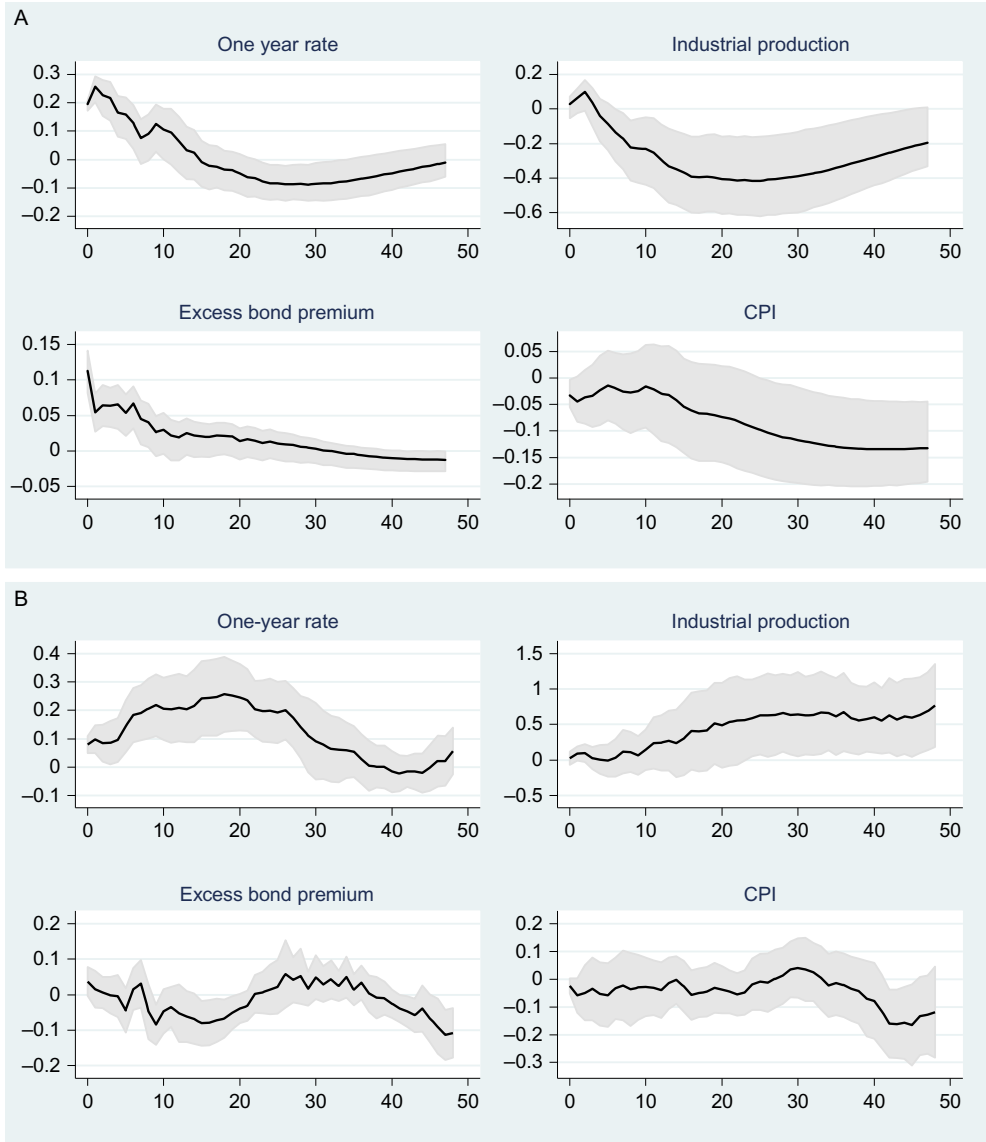
Fig. 3 Gertler–Karadi's monetary shock. (A) Gertler–Karadi's monetary proxy SVAR, VAR from 1979m7 to 2012m6, instrument from 1991m1 to 2012m6. (B) Gertler–Karadi monetary shock, Jordà 1990m1–2012m6. Light gray bands are 90% confidence bands.

supposed to capture only unanticipated movements in interest rates. In my local projec‐ tion framework implementation, I include lagged values of the shock, so my procedure purges the shock of this serial correlation. I discovered that the serial correlation is induced by the method that Gertler and Karadi use to convert the announcement day

shocks to a monthly series.[w] Third, if I used FOMC-frequency data to regress the shock on all of the Greenbook variables that the Romers used to create their shock, the R-squared of the regression is 0.21 and I can reject that the coefficients are jointly zero with a p-value of 0.027.[x] Thus, the Gertler–Karadi variable is predicted by Greenbook projections. Gertler and Karadi also worried about this issue, but they performed a robustness check based only on the *difference* between private forecasts and Greenbook forecasts. They found a much lower R-squared (see their table 4). When they use their purged measure, they find greater falls in industrial production. I explored the effect of using a version of their measure that was (i) orthogonal to the Romer Greenbook variables; and (ii) converted to a monthly basis the same way that the Romer's converted their data, in the Jordà framework. The results (not shown) were still puzzling.

Why does the Jordà method give such different estimates from the proxy SVAR? One possible explanation is the different method and sample used to estimate the impulse response function. Gertler and Karadi's impulse responses functions are constructed as nonlinear functions of the reduced form VAR parameters estimated on data from 1979 through 2012; the Jordà method estimates are for the 1991–2012 sample and are direct projections rather than functions of reduced form VAR parameters. Since the estimates of the impact effects on industrial production are near zero for both methods, the entire difference in the impulse responses is due to the differences in the dynamics implied by Gertler and Karadi's reduced form VAR parameter estimates. A second possible explanation for the difference is that the rising importance of forward guidance starting in the mid-1990s means that the VAR underlying the proxy SVAR is misspecified. As discussed in Section 2.5, anticipations of future policy actions can lead to the problem of a nonfundamental moving average representation. Gertler and Karadi's fed funds futures variable captures news well, but they do not include it directly in the SVAR; they only use it as an instrument.

3.6 Summary of Monetary Shocks

When Christiano et al. (1999) wrote their Handbook chapter, they provided what became a benchmark framework for identifying monetary policy shocks and tracing their effects. As long as the recursiveness assumption was incorporated, the results were quite robust. Since then, the literature has incorporated new methods and faced new challenges. Researchers now take instrument identification and relevance much more seriously when estimating monetary policy shocks. New methods, such as FAVARs and Greenbook forecasts, have improved the conditioning set for estimating monetary policy shocks. Structural VARs (SVARs), sign restrictions, and regime-switching models have provided alternatives to the usual Cholesky decomposition. Moreover, new measures of

[w] See footnote 11 of Gertler and Karadi (2015) for details.
[x] I am indebted to Peter Karadi for sharing with me the announcement date series.

monetary shocks have been developed using rich external data, such as narrative data, Greenbook projections, and high-frequency information from financial markets. Recently published work using shocks estimated with external data results in similar conclusions. In particular, Coibion's (2012) reconciliation of the Romer results with the VAR results suggests that a 100 basis point rise in federal funds rate lowers industrial production by about −2% at 18 months. Those results are based on data from 1969 through 1996. Gertler and Karadi's (2015) research uses HFI from fed funds futures and external instruments/proxy SVAR methods to find very similar results for a later sample.

This rosy reconciliation picture disappears, however, when the specifications are subjected to some robustness tests. My explorations have highlighted several potential issues, some of which were already noted in the literature. First, the original Christiano et al. (1999) specification, as well as many other specifications, does not hold up well in later samples. Second, lifting the recursiveness assumption can lead to estimates that imply expansionary effects of contraction monetary policy in the short run. Third, one needs to be very careful when estimating models in samples where anticipation effects may be important. For example, it is not clear that HFI shocks should be used as external instruments for otherwise standard VARs.

How should we interpret these results? I would argue that the most likely reason for the breakdown of many specifications in the later sample is simply that we can no longer identify monetary policy shocks well. Monetary policy is being conducted more systematically, so true monetary policy shocks are now rare. It is likely that what we now identify as monetary policy shocks are really mostly the effects of superior information on the part of the Fed, foresight by agents, and noise. While this is bad news for econometric identification, it is good news for economic policy.

What, then, are we to conclude about the output effects of monetary shocks? I would argue that the best evidence still remains the historical case studies, such as Friedman and Schwarz, and the times series models estimated on samples that exclude recent decades. Of course, one worries that the structure of the economy may have changed in the last few decades, but we simply do not have enough information to produce estimates with any great certainty. Monetary policy can have big effects, but it is likely that monetary shocks are no longer an important source of macro instability.

4. FISCAL SHOCKS

This section reviews the main identification methods and summarizes existing results from the empirical literature seeking to identify and estimate the effects of fiscal policy shocks. It also presents some new results comparing several leading identified shocks.

In contrast to a monetary policy shock, a fiscal shock is a more straightforward economic concept. Because the legislative and executive branches of government often make tax and spending decisions based on concerns that are orthogonal to the current

state of the macroeconomy, the notion of regularly occurring fiscal policy shocks is more plausible than regularly occurring monetary policy shocks.

Measuring the empirical effects of changes in government spending and taxes on aggregate GDP and its components was an active research area for a number of decades. The large Keynesian models of the 1960s included fiscal variables and numerous academic papers estimated their effects in behavioral equations. For several decades afterward, though, research on the aggregate effects of tax and spending shocks experienced a lull, punctuated by only a few papers. Most empirical research on shocks during the 1980s, 1990s, and 2000s instead focused on monetary policy. With the onset of the Great Recession and the zero lower bound, however, research energy immediately shifted to the effects of fiscal policy. The recent literature has built on and extended the strides made by the few authors working on the topic during the long dormant period.

The following sections will discuss some of the literature since 1990 that has sought to analyze the effects of fiscal shocks. I will begin by considering government spending shocks and then discuss tax shocks.

4.1 Government Spending Shocks

In this section, I discuss shocks to government spending. When I use the term *government spending*, I mean government *purchases*, ie, G in the NIPA identity. In common parlance, however, government spending typically refers to government *outlays*, which include both government purchases and transfer payments. Economists usually consider transfer payments to be negative taxes. Thus, I will include a discussion of transfer payments in the section on the effects of tax shocks.

4.1.1 Summary of Identification Methods

Many of the identification methods summarized in Section 2 are used in the literature that analyzes the effects of shocks to government spending. These methods include SVARs with contemporaneous restrictions, sign restrictions, medium-horizon restrictions, narrative methods, and estimated DSGE models.

Perhaps the first example of what looks like a VAR-type analysis of the effects of fiscal shocks is Rotemberg and Woodford's (1992) analysis of the effects of military spending and employment on macroeconomic variables. Their purpose was to provide evidence in favor of their countercyclical markup model, showing that a "demand" shock would lead to countercyclical markups. To do this, they estimated systems with military spending, military employment, and a macroeconomic variable of interest (such as private value added and private hours worked). They included lags of the variables in the system, but restricted the VAR so that there was no feedback of the macroeconomic variables onto the military variables. In their system, identification was achieved as follows. To identify government purchases shocks that were exogenous to the economy, they followed Hall (1980, 1986) and Barro (1981) who argued that defense spending is driven

by military events rather than by macroeconomic events. To identify *unanticipated* shocks, they regressed the military variables on their own lags and used the residuals. This identification strategy assumes that all relevant information for predicting military spending and employment is contained in lags of military spending and employment. They showed that their identified shocks to defense spending raised real wages.

In a paper analyzing the effects of sectoral shifts in the presence of costly mobility of capital across sectors, Ramey and Shapiro (1998) used narrative techniques to create a dummy variable capturing major military buildups. They read through *Business Week* in order to isolate the political events that led to the buildups in order to create a series that was exogenous to the current state of the economy. They also used this narrative approach to ensure that the shock was unanticipated. They stated: "We believe this approach gives a clearer indicator of unanticipated shifts in defense spending than the usual VAR approach, since many of the disturbances in the VAR approach are due solely to timing effects on military contracts and do not represent unanticipated changes in military spending" (Ramey and Shapiro, 1998, p. 175). Ramey and Shapiro (1998) estimated the effects of "war dates" by regressing each variable of interest on current values and lags of the war dates and lags of the left-hand side variable. A number of follow-up papers embedded the war dates in VARs, ordered first in the Cholesky decomposition, creating "EVARs", a term coined by Perotti (2011). These papers include Edelberg et al. (1999) and Burnside et al. (2004). Most applications typically found that while government spending raised GDP and hours, it lowered investment, consumption, and real wages. Most of these papers did not specifically estimate a multiplier, though one can typically back out the implied multiplier from the impulse responses.

In contrast, Blanchard and Perotti (2002) used an SVAR to identify both government purchases and tax shocks. They assumed that government purchases were predetermined within the quarter, and identified the shock to government purchases using a standard Cholesky decomposition with government spending ordered first. They found that government purchases shocks raised not only GDP but also hours, consumption, and real wages. Follow-up work, such as by Fatás and Mihov (2001), Perotti (2005), Pappa (2009) and Galí et al. (2007), found similar results. Mountford and Uhlig (2009) used sign restrictions and found only weak effects on GDP and no significant effect on consumption.

In Ramey (2011a), I sought to reconcile why the war dates were producing different results from the SVARs that used Cholesky decompositions. I argued that most government spending is anticipated at least several quarters in advance, so that the standard SVAR method was not identifying unanticipated shocks. In support of this idea, I showed that the shocks from an SVAR were indeed Granger-caused by the Ramey and Shapiro (1998) war dates. To create a richer narrative variable to capture the "news" part of government spending shocks, I read *Business Week* starting in 1939

and created a quantitative series of estimates of changes in the expected present value of government purchases, caused by military events. I then embedded the news series in a standard VAR, with the news ordered first in the Cholesky decomposition. In that work, I found results that were broadly consistent with the estimates based on the simple war dates.

In follow-up work, Owyang et al. (2013) and Ramey and Zubairy (2014) extended the military news series back to 1889. The military news variable tends to have low instrument relevance for samples that begin after the Korean War, though. In Ramey (2011a), I augmented my analysis by also considering shocks that were orthogonal to professional forecasts of future government purchases. Fisher and Peters (2010) created an alternative series of news based on the excess returns of defense contractor stocks for the period starting in 1958. Recent work by Ben Zeev and Pappa (forthcoming) uses the medium-horizon identification methods of Barsky and Sims (2011) to identify news shocks to defense spending from a time series model. In particular, Ben Zeev and Pappa identify defense spending news as a shock that (i) is orthogonal to current defense spending; and (ii) best explains future movements in defense spending over a horizon of 5 years.

All of these measures of anticipations have weaknesses, though. First, because they are associated with military events, there are likely confounding effects (eg, rationing, price controls, conscription, patriotic increases in labor supply). Second, as I show below, some of the shocks suffer from low first-stage F-statistics in some samples, indicating that they might not be relevant instruments for estimating multipliers.

Thus, there are two main differences in the shocks identified across these two classes of models. First, the SVAR shocks are more likely to be plagued by foresight problems. As I discussed in Section 2, this problem of foresight can be a serious flaw in SVARs. Second, the news alternatives are not rich enough in some subsamples and there may be confounding influences.

A more structural way to identify shocks to government purchases is through an estimated DSGE model. For example, one of the shocks identified by Smets and Wouters (2007) is a government purchases shock. Cogan et al. (2010) also estimate government spending multipliers in the context of an estimated DSGE model.

4.1.2 Summary of the Main Results from the Literature

Typically, the literature on government spending has sought to answer one or both of two main questions: (1) Are the empirical results consistent with theoretical DSGE models? (2) What are the government spending multipliers?

Let us begin by considering results that shed light on the first question. Most versions of standard neoclassical theory and standard new Keynesian theory predict that a rise in government purchases (financed with deficits or lump-sum taxes and not spent on public infrastructure, etc.) should raise GDP and hours, but should decrease consumption and real wages. Whether investment initially rises or falls depends on the persistence of the

increase in government spending. It is only when one adds extra elements, such as rule-of-thumb consumers and off-the-labor supply behavior of workers that one can produce rises in consumption and real wages in a model (eg, Galí et al., 2007).

Both SVARs and EVARs that use a news variable produce qualitative similar results for some variables. For example, both typically estimate an increase in GDP and hours and a fall in investment (at least after the first year) in response to a positive government spending shock. In contrast, the SVAR typically implies a rise in consumption and real wages, whereas the EVAR predicts a fall in consumption and real wages.

The second question the literature seeks to answer is the size of "the" government purchases multiplier. Unfortunately, most estimates are not for pure deficit-financed multipliers since most rises in government spending are accompanied by a rise in distortionary taxes, typically with a lag. This caveat should be kept in mind in the subsequent discussion of multiplier estimates.

One might assume that SVARs produce bigger multipliers since they predict increases in consumption. They do not. In Ramey (2013), I compared the effects of government spending on private spending, ie, GDP minus government spending, of several shocks based on the various identification methods. If the government spending multiplier is greater than unity, then private spending must increase. I found that all of the shocks lowered private spending, but that the Blanchard–Perotti (2002) shocks lowered it more, implying a smaller multiplier.

The estimated DSGE models of Smets and Wouters (2007) and Cogan et al. (2010) produce results that are close to the neoclassical model. In both cases, a shock to government spending lowers consumption and results in multipliers below unity.

In my survey of the literature on multipliers in Ramey (2011b), I found that most estimates of the government spending multiplier in aggregate data were between 0.8 and 1.2. The only multipliers that were larger were (1) those estimated on states or regions; and (2) some of those estimated allowing state dependence. As suggested in my survey, and as shown formally by Nakamura and Steinsson (2014) and Farhi and Werning (2012), the link between estimates of multipliers in a fiscal union (eg, across US states or regions) for aggregate multipliers is not entirely clear. Most of the cross-state or cross-region studies look at the effect of federal spending on a locality. Unfortunately, because constant terms or time fixed effects are included, these regressions difference out the effects of the financing, since taxes that finance federal spending are levied at the national level.[y] This explains why multipliers on federal spending at the state level will be higher than the aggregate multipliers. I will discuss the issue of state dependence in more detail momentarily.

---

[y] A notable exception is the paper by Clemens and Miran (2012), which identifies exogenous variation in state-level spending. Interestingly, they find multipliers around 0.5, which is closer to those found at the national level.

Since writing that survey, I realized that there were two potential biases in the way that many researchers calculated their multiplier, and as a result, many reported estimates are not comparable. First, many researchers followed Blanchard and Perotti's (2002) lead and calculated multipliers by comparing the *peak* output response to the *initial* government spending impact effect. While comparing values of impulse responses at peaks or troughs is a useful way to compare impulse responses, it is not a good way to calculate a multiplier. As argued by Mountford and Uhlig (2009), Uhlig (2010), and Fisher and Peters (2010), multipliers should instead be calculated as the *integral* (or present value) of the output response divided by the *integral* government spending response. The integral multipliers address the relevant policy question because they measure the cumulative GDP gain relative to the cumulative government spending during a given period. In many cases, Blanchard and Perotti's ratio of peak-to-impact method gives a higher number for the multiplier than the integral method. Second, most researchers estimating VARs use logarithms of variables. To convert the estimates to multipliers, they usually multiply the estimates by the sample mean of GDP to government spending ratio. As Owyang et al. (2013) point out, this can lead to serious biases in samples with significant trends in the GDP to government spending ratio. In the few cases where I have been able to adjust the estimates of multipliers to be integral multipliers, I have found that the multipliers are often below one.

With this additional caveat in mind, Table 3 shows a summary of a few of the estimates of multipliers for government purchases. Even with the variety of ways of calculating multipliers from the estimated impulse response functions, most of the estimates are from 0.6 to 1.5.

A number of researchers and policy makers have suggested that multipliers may be state dependent. Auerbach and Gorodnichenko (2012) use a smooth transition vector autoregression model and find evidence of larger multipliers in recessions. Ramey and Zubairy (2014) use the Jordà (2005) method (also used by Auerbach and Gorodnichenko, 2013 in a panel of countries) and find little evidence of state dependence, based on recessions, elevated unemployment rates, or the zero lower bound. They argue that their different finding is not so much due to the underlying parameter estimates but rather due to the additional assumptions that Auerbach and Gorodnichenko (2012) made when transforming those estimates into multipliers.

Most of the studies I have summarized focus on government purchases that do not involve infrastructure spending. The reason for the paucity of research on infrastructure spending is the difficulty of identifying shocks to infrastructure spending, particularly in the United States. The US highway system was an important part of government purchases starting in the late 1950s through the early 1970s. The problem with identifying the aggregate effects is that most of the spending was anticipated once the highway bill was passed in 1956. Most of the credible analyses have used clever indirect methods or used variation in cross-state expenditures. Fernald (1999) provides very strong evidence for a causal effect of the highway system on productivity by showing its greater effect on

Table 3 Summary of government spending multiplier estimates for the aggregate United States

Study	Sample	Identification	Implied spending multiplier
Barro (1981), Hall (1986), Hall (2009), Barro–Redlick (2011)	Annual historical samples	Use military spending as instrument for government spending	0.6–1
Rotemberg–Woodford (1992)	Quarterly, 1947–1989	Residuals from regression of military spending on own lags and lags of military employment	1.25
Ramey–Shapiro (1998), Edelberg et al. (1999), Burnside et al. (2004)	Quarterly, 1947 to the late 1990s or 2000s	Ramey–Shapiro dates, which are based on narrative evidence of anticipated military buildups	0.6–1.2, depending on sample and whether calculated as cumulative or peak
Blanchard–Perotti (2002)	Quarterly, 1960–1997	SVARS, Cholesky decomposition with G ordered first	0.9–1.29, calculated as peak multipliers
Mountford–Uhlig (2009)	Quarterly, 1955–2000	Sign restrictions on an SVAR	0.65 for a deficit-financed increase in spending
Bernstein and Romer (2009)	Quarterly	Average multipliers from FRB/US model and a private forecasting firm model	Rising to 1.57 by the 8th quarter
Cogan et al. (2010)	Quarterly, 1966–2004	Estimated Smets–Wouters model	0.64 at peak
Ramey (2011a,b)	Quarterly, 1939–2008 and subsamples	VAR using shocks to the expected present discounted value of government spending caused by military events, based on narrative evidence	0.6 –1.2, depending on sample
Fisher–Peters (2010)	Quarterly, 1960–2007	VAR using shocks to the excess stock returns of military contractors	1.5 based on cumulative effects
Auerbach and Gorodnichenko (2012)	Quarterly, 1947–2008	SVAR that controls for professional forecasts, Ramey news. Key innovation is regime-switching model	Expansion: −0.3 to 0.8 Recession: 1–3.6 (uses a variety of ways to calculate multipliers)
Ben Zeev and Pappa (forthcoming)	Quarterly, 1947–2007	Shock that (i) is orthogonal to current defense spending; and (ii) best explains future movements in defense spending over a horizon of 5 years	2.1 based on integral multiplier at 6 quarters

vehicle-intensive industries. These estimates do not directly inform us about aggregate effects, though. Leduc and Wilson (2013) identify news shocks about highway spending in a panel of US states using the arrival of new information about institutional formula factors. However, as discussed earlier, the multipliers they find cannot be converted to aggregate multipliers.

Gechert (2015) conducts a meta-analysis of 104 studies of multiplier effects across a variety of countries, including many different types of analyses from reduced form empirical to estimated DSGE models. With the caveat that the context and experiment varies across studies, Gechert finds that public spending multipliers are close to one, while public investment multipliers are around 1.5.

4.1.3 Explorations with Several Identified Shocks

I now study the effects of several of the leading identified government spending shocks in the Jordà local projection framework. This exploration is useful not only for gauging the robustness of the results to local projection methods but also for comparing the effects of the identified shocks using the same data, same specification, and the same way to calculate multipliers. Thus, any differences in results will be due to the identification methods rather than differences in data or implementation.

The three main shocks I study are (i) the shock identified using Blanchard and Perotti's (2002) method (which simply orders government spending first in a Cholesky decomposition); (ii) my narrative military news shock, updated in Ramey and Zubairy (2014); and (iii) Ben Zeev and Pappa's (forthcoming) defense news shock identified using Barsky and Sims' (2012) medium-run horizon method.[z] I also comment briefly on results using Fisher and Peters' (2010) military contractor excess returns.

In all cases, I use the following data transformations and functional forms. The first set of transformations is intended to facilitate the direct estimation of multipliers in order to avoid ad hoc transformation of estimates based on logs, as discussed by Owyang et al. (2013). One can use either the Hall (2009) and Barro–Redlick (2011) transformation or a Gordon–Krenn (2010) transformation. The Hall–Barro–Redlick transformation constructs variables as $(X_{t+h} - X_{t-1})/Y_{t-1}$, where X is the NIPA variable deflated by the GDP deflator and Y_{t-1} is real GDP before the shock hits in period t. The Gordon–Krenn transformation divides all NIPA variables by "potential GDP," estimated as an exponential trend. Both methods give similar results. I follow the Gordon–Krenn procedure, fitting log real GDP to a quadratic trend.[aa] Thus, the NIPA variables are

---

[z]   I estimated the Blanchard–Perotti shock using logarithms of real government spending, GDP, and taxes and four lags. One could instead estimate it directly in the regression using the Gordon-Krenn transformed variables. Ben Zeev and Pappa kindly provided me with estimates of their shock.

[aa] One could use the CBO estimate of real potential GDP instead. I found, however, that when I used the CBO estimate, the implied multipliers were noticeably smaller than when I used Hall–Barro–Redlick or Gordon–Krenn with either a quadratic or a quartic log trend.

transformed to be $z_t = X_t/Y_t^*$, where Y_t^* is the estimated trend in real GDP. The impulse responses using this transformation look qualitatively similar to those using log levels, but often have more narrow confidence bands.

The non–NIPA variables are transformed as follows. The average tax rate is federal current receipts divided by nominal GDP. The hours variable is the log of total hours per capita, where the total population is used in the denominator. Wages are given by the log of nominal compensation in the business sector, deflated by the price deflator for private business. The real interest rate is the 3-month Treasury bill rate minus the rate of inflation calculated using the GDP deflator.

The equation used to estimate the impulse responses for each variable z at each horizon h is given by

$$z_{t+h} = \alpha_h + \theta_h \cdot \text{shock}_t + \varphi_h(L)\gamma_{t-1} + \text{quadratic trend} + \varepsilon_{t+h} \qquad (18)$$

where z is the variable of interest, *shock* is the identified shock, γ is a vector of control variables, and $\varphi_h(L)$ is a polynomial in the lag operator. All regressions include two lags of the shock (to mop up any serial correlation), transformed real GDP, transformed real government purchases, and the tax rate. Regressions for variables other than real GDP, government purchases, and tax rates also include two lags of the left-hand side variable. The coefficient θ_h gives the response of z at time $t+h$ to the shock at time t.

As discussed earlier, the correct way to calculate a multiplier is as the integral under the impulse response of GDP divided by the integral of the impulse response of government spending. We could compute the multiplier using the following multistep method: (1) estimate Eq. (18) for GDP for each horizon and sum the coefficients θ_h up to some horizon H; (2) estimate Eq. (18) for government purchases for each horizon and sum the coefficients θ_h up to some horizon H; and (3) construct the multiplier as the answer from step (1) divided by the answer from step (2). Estimating the standard error, however, requires some ingenuity, such as estimating all of the regressions jointly in a panel estimation.

Alternatively, we can easily estimate the multiplier estimate and its standard error in one step if we cumulate the variables and reformulate the estimation problem as an instrumental variables (IV) estimation. In particular, we can estimate the following equation:

$$\sum_{i=0}^{h} z_{t+i} = \beta_h + m_h \cdot \sum_{i=0}^{h} g_{t+i} + \chi_h(L)\gamma_{t-1} + \text{quadratic trend} + \nu_{t+h} \qquad (19)$$

where the dependent variable is the *sum* of real GDP (or other NIPA variable) from t to $t+h$ and the government spending variable is the *sum* of the government purchases variable. We use the identified shock as the *instrument* for the sum of government purchases. The estimated coefficient, m_h, is the multiplier for horizon h. There are several advantages to this one-step IV method. First, the standard error of the multiplier is just the standard
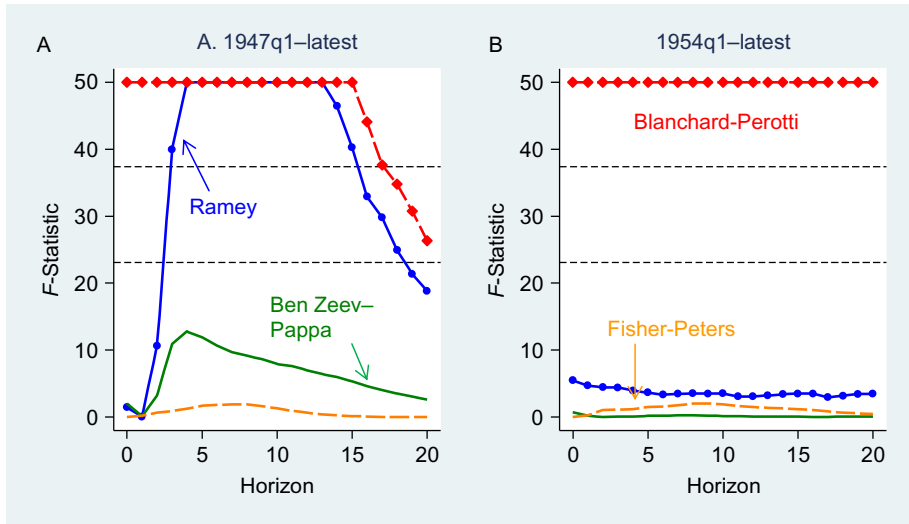
Fig. 4 First-stage *F*-statistics for government spending shocks. Note: The *F*-statistics are based on the regression of the sum of government spending from t to $t+h$ on the shock at t, plus the lagged control variables. Values above 50 have been capped at 50. The *horizontal dashed lines* are the Montiel Olea and Pflueger (2013) 5% worst case bias (*upper line*) and 10% worst case bias (*lower line*) thresholds.

error of the coefficient m_h.[bb] Second, the shock can have measurement error as long as the measurement error is not correlated with any measurement error in government spending. Third, formulating the estimation as an IV problem highlights the importance of instrument relevance.

Thus, I first consider how relevant each of the identified shocks is as an instrument for the integral of government spending. Stock et al. (2002) argue that the first-stage *F*-statistic should be above 10 for the IV estimates to be reliable, but their threshold applies only to first-stage regressions with serially uncorrelated error terms. Fortunately, follow-up work by Montiel Olea and Pflueger (2013) constructs thresholds for cases with serial correlation. For the first stage of Eq. (19), the thresholds are 23 for the 10% level and 37 for the 5% level.[cc] *F*-Statistics below those thresholds indicate a possible problem with instrument relevance.

Fig. 4 shows the first-stage *F*-statistics for the sum of government purchases on each identified shock for each horizon up to 20 quarters. Values above 50 have been capped at 50 for ease of viewing. The graph at the left shows the results for the sample starting in 1947 and the graph on the right shows the results for the sample starting in 1954, after the

---

[bb] Because of the serial correlation in the errors, any procedures for estimating standard errors should use methods that account for serial correlation.

[cc] These F-statistics and thresholds were derived using Pflueger and Wang's (2015) "weakivtest" Stata module.

Korean War. Several results emerge. First, the Blanchard–Perotti (BP) identified shock always has very high F-statistics. This is not surprising because the shock is equal to the portion of government spending not predicted by four lags of government spending, GDP, and taxes. Second, the Fisher–Peters defense contractor excess returns shock has very low F-statistics for all horizons and both samples, indicating that the return variable is not a good instrument for government spending. Third, the Ramey and Ben Zeev–Pappa (BZP) news shocks have low relevance for short horizons, but this is to be expected since those shocks capture news about *future* government spending. Fourth, in the full sample at horizons beyond three quarters, the Ramey news shock has F-statistics above the Montiel–Pflueger thresholds, whereas the BZP news shock F-statistics lie below them and range between 8 and 13. Fifth, in the sample that excludes the Korean War, all of the F-statistics are very low except for the Blanchard–Perotti shock. Thus, the BP shock surpasses the relevance safety threshold for all horizons in both samples, the Ramey news shock does so for the full sample for horizons at four to 20 quarters, while the BZP shock may have relevance problems at most horizons and the Fisher–Peters shock always has very low relevance. I thus exclude the Fisher–Peters shock from the rest of the analysis.

Fig. 5 shows the impulse responses estimated using Eq. (18), with estimates normalized across specifications to have the same peak in government purchases. The scales in the graphs of the NIPA variables should be interpreted as dollars; ie, a rise in government purchases that peaks at $1 leads to rise in GDP that peaks at 75 cents. The scales of the other graphs are in percentage points. The confidence bands are 90% bands based on Newey–West corrections of standard errors. They do not, however, take into account that two of the shocks are generated regressors.

Consider first the upper left graph in Fig. 5. Both the Ramey and BZP news variables imply similar paths of government purchases, with little effect for the first few quarters rising to a peak around five quarters after the shock. In contrast, the BP shock leads to an immediate rise in government spending. The graph in the top right shows that in response to all three shocks, GDP jumps immediately. The response of GDP is greatest for the BZP shock, but GDP begins to fall back to normal even before government purchases have hit their peak.

The tax rate series is simply federal receipts divided by GDP. This variable can rise either because of tax legislation or because higher GDP pushes more households into higher tax brackets. According to the estimates, tax rates begin to rise immediately for the BZP shock but only gradually for the Ramey news shock. Tax rates gradually fall after BP shock. Real interest rates (measured as the 3-month T-bill rate minus inflation) fall after a news shock, but rise slightly after a BP shock. Examination of the responses of the components of the real interest rate (not shown) reveals that the fall is due to both a drop in the nominal interest rate and a rise in inflation. As explained by Ramey (2011a), the rise in inflation is in large part driven by the spike up in prices at the beginning of the
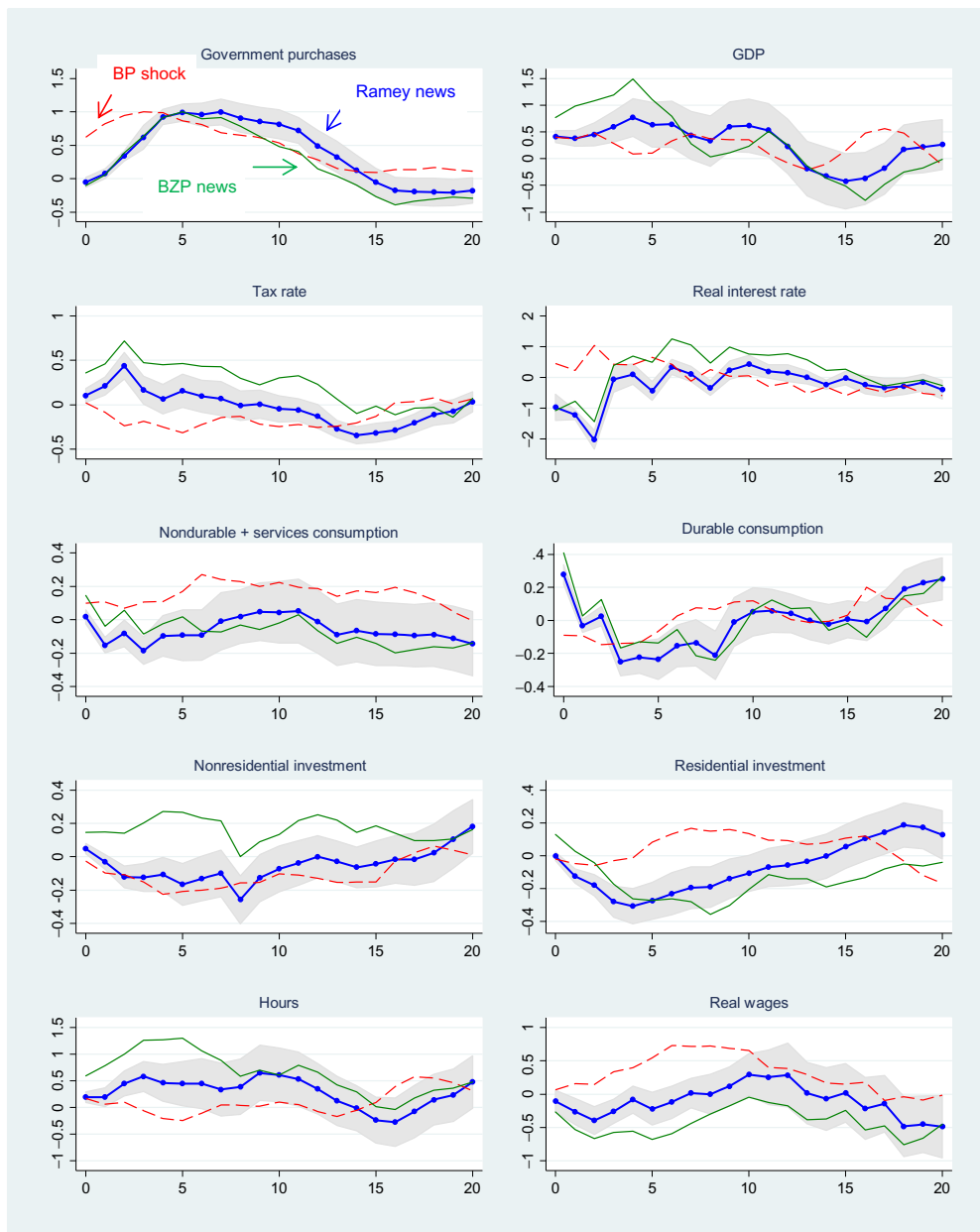
Fig. 5 Comparison of the effects of government spending shocks (*BP*: Blanchard–Perotti; *BZP*: Ben Zeev–Pappa). Light gray bands are 90% confidence bands.

Table 4 Multiplier estimates (HAC standard errors in parenthesis)

Horizon (in quarters)	Blanchard–Perotti	Ramey news	Ben Zeev–Pappa news
0	0.65 (0.24)	−7.53 (7.26)	−7.37 (5.85)
4	0.37 (0.23)	1.37 (0.33)	2.91 (1.13)
8	0.39 (0.32)	0.80 (0.25)	1.41 (0.61)
12	0.39 (0.44)	0.77 (0.27)	1.24 (0.71)
16	0.40 (0.58)	0.60 (0.36)	1.10 (1.01)
20	0.44 (0.63)	0.69 (0.48)	1.17 (1.46)

Notes: Multipliers estimated using Eq. (19). All regressions also include two lags of the shock (to mop up any serial correlation), real GDP (divided by potential GDP), real government purchases (divided by potential GDP), the tax rate, and a quadratic trend.

Korean War: with recent memories of WWII, firms thought that price controls were coming and raised their prices in advance.

Consider now four components of the national income accounts, shown in the middle graphs in Fig. 5. Nondurable plus services consumption falls after a Ramey news shock, responds little after a BZP shock, but rises after a BP shock. In Ramey (2009), I show using simulations of a DSGE model that one can estimate a rise in consumption if one treats an anticipated shock as an unanticipated shock. I argue that the rise in consumption after a BP shock can be explained by this type of identification problem.

Durable consumption spikes up initially and then falls after the two news shocks. As in the cast of prices, this initial spike on the arrival of news is driven mostly by the response of consumers to the beginning of the Korean War in 1950: with recent memories of WWII, consumers worried that rationing of durable goods was imminent so they hurried out to buy durable goods. Nonresidential investment rises in response to the BZP shock, but falls in response to both the Ramey news and BP shock. Residential investment falls in response to the two news shocks, but rises after a year in response to the BP shock.

Finally, both news shocks imply a rise in hours and a fall in the real wage, while the BP shock predicts very little response of hours, but a rise in the real wage.

Table 4 shows the estimated multipliers for various horizons.[dd] The impact multipliers for the two news shocks are negative because output jumps up, while government spending falls slightly. For the next few quarters, the multipliers for the two news shocks are large because output responds immediately to news of future government spending which has not yet fully transpired. Once government spending has risen to its peak, the implied multipliers using the Ramey news shock are just below unity, whereas those using the BZP news shock are above unity. For example, the BZP news shock multiplier is 1.4 at 8 quarters and 1.1 at 16 quarters. The responses in Fig. 5 reveal that the reason for the larger multiplier after a BZP shock is the large rise in nonresidential investment.

[dd] These estimates are based on the one-step method shown in Eq. (19).

It should be noted, though, that the BZP multipliers are estimated rather imprecisely as evidenced by the standard errors. This is one manifestation of the possible low instrument relevance of the BZP news shock. On the other hand, the multipliers implied by the Blanchard and Perotti shock are all low, most below 0.5. However, the estimates are not precise enough to reject a multiplier of either 0 or 1 at standard significance levels.

I now consider what each shock implies about the contribution to the forecast error variance of output. Although one can calculate forecast error variances using the estimated local projection coefficients, I found that the shares sometimes exceeded 100%. Thus, for present purposes I calculate the forecast error variance in a standard VAR with the shock, log government spending, log real GDP, and log taxes. The shock is ordered first and four lags are included, along with a quadratic trend.

Table 5 shows the forecast error variance decompositions of each of the three identified government spending shocks for government spending and output. Despite having the lowest contribution for government spending, the BZP shock has the highest contribution for output, but it is still 13% or below. The BP and Ramey news contributions tend to be 5% of below. Thus, none of the three shocks is an important contributor to the variance of output.

To summarize, most of the aggregate analyses find government spending multipliers between 0.6 and 1.5. The BP shock leads to smaller multipliers, but does imply that government spending leads to rises in consumption and real wages along with GDP and hours. In contrast, both the Ramey news and BZP news shocks lead to falls in real wages. Both news shock lead to an initial spike in durable consumption (due to the consumer fears of rationing), followed by a decline. The BZP shock produces a temporary blip in nondurable consumption, but then it falls to 0. The Ramey news shock implies a decline nondurable consumption. None of the methods suggests that government spending shocks explain an important part of GDP fluctuations.

Table 5 Shock contribution to the forecast error variance of government spending and output

| Horizon (in quarters) | Blanchard–Perotti | | Ramey news | | Ben Zeev–Pappa news | |
	Government spending	Output	Government spending	Output	Government spending	Output
0	100.0	5.5	1.0	2.2	1.4	5.6
4	96.2	3.3	31.8	2.6	14.0	10.1
8	90.5	2.9	50.2	2.9	27.0	12.6
12	86.5	2.5	50.5	2.5	29.8	12.1
16	83.1	2.4	46.7	2.4	29.4	11.8
20	80.2	2.3	43.0	2.2	28.7	11.7

Notes: Based on standard VAR with the shock, log output, log government spending, log taxes, and a quadratic trend. The shock is ordered first and four lags of the variables are included.

4.2 Tax Shocks

I now turn to the literature on tax shocks. Taxes were often an important component of the big Keynesian econometric models of the 1960s. The public finance literature has analyzed many details of the effects of taxes. In this section, I will focus on estimates of the effects of taxes in the macroeconomic literature since the 1990s.

4.2.1 Unanticipated Tax Shocks
4.2.1.1 Summary of the Literature
Macroeconomists have used both estimated DSGE models and time series models to estimate the effects of taxes. One of the first systematic analyses of macroeconomic tax effects in an estimated DSGE model was by McGrattan (1994). She extended the Kydland and Prescott (1982) model to include government consumption, labor income taxes, and capital income taxes and estimated the parameters using maximum likelihood. She found that the role of technology in business cycle fluctuations was much reduced, 41% rather than Kydland and Prescott's 75% estimate. She found that shocks to government consumption accounted for 28% of the forecast error variance of output, labor income tax shocks for 27%, and capital income tax shocks for 4%.

Among time series approaches, Blanchard and Perotti (2002) used an SVAR approach in which they identified tax shocks by imposing the elasticity of net taxes to GDP estimated from other studies. Returning to the discussion of the simple trivariate model from Section 2, consider the following system:

$$\begin{aligned}
\eta_{1t} &= b_{12}\eta_{2t} + b_{13}\eta_{3t} + \varepsilon_{1t} \\
\eta_{2t} &= b_{21}\eta_{1t} + b_{23}\eta_{3t} + \varepsilon_{2t} \\
\eta_{3t} &= b_{31}\eta_{1t} + b_{32}\eta_{2t} + \varepsilon_{3t}
\end{aligned} \tag{20}$$

where η_{1t} is the reduced form residual of net taxes, η_{2t} is the reduced form residual of government spending, and η_{3t} is the reduced form residual of GDP. Blanchard and Perotti (2002) identify the shock to government spending using a Cholesky decomposition in which government spending is ordered first (ie, $b_{21} = b_{23} = 0$). They identify exogenous shocks to net taxes by setting $b_{13} = 2.08$, an outside estimate of the cyclical sensitivity of net taxes. These three restrictions are sufficient to identify all of the remaining parameters and hence all three shocks. Blanchard and Perotti's estimated "impact multiplier" was -0.78. Their impact multiplier was calculated as the *trough* of GDP relative to the *initial* shock to taxes.

Mountford and Uhlig (2009) use sign restrictions to identify tax and spending shocks. Their results imply a multiplier of -5 at 12 quarters for a deficit-financed tax cut, when the multiplier is calculated as the ratio of the present value of the impulse response functions. In order to compare their results to Blanchard and Perotti, they also calculate "impact multipliers," meaning the value of the GDP response at a certain quarter to the initial shock impact on the fiscal variable. They find that whereas the Blanchard

and Perotti method implies a peak-to-impact multiplier of −1.3 at quarter 7, Mountford and Uhlig's results imply a peak-to-impact multiplier of −3.6.

In the context of the Blanchard and Perotti (2002) model, Caldara and Kamps (2012) demonstrate how the estimated multiplier depends crucially on their assumption about the elasticity of net tax revenue to GDP. Particularly important is their demonstration of how a small change in the assumed cyclical elasticity parameter can result in large changes in the estimated tax multiplier; to wit, this seems to be a case of a "multiplier multiplier" on the assumed elasticity. Caldara and Kamps (2012) propose a new method, which involves imposing probability restrictions on the output elasticities of taxes and spending. When they implement this method, they find peak-to-impact multipliers of −0.65 for tax shocks and peak-to-impact multipliers greater than unity for government spending shocks.

Barro and Redlick (2011) construct a new series of average marginal tax rates using IRS data and analyze its effects in a system that also considers government spending in annual data extending back to 1917. In their baseline specification, they find that an increase in the average marginal tax rate of 1 percentage point lowers GDP by 0.5%. Their calculations indicate a tax multiplier of −1.1.

Romer and Romer (2010) (R&R) use narrative methods to identify tax shocks. Based on presidential speeches and congressional reports, they construct several series of legislated tax changes and distinguish those series based on the motivation for enacting them. They argue that tax changes motivated by a desire to pay down the deficit or long-run growth considerations can be used to establish the causal effect of tax changes on output. When they estimate their standard dynamic single equation regression of output growth on its lags and on current and lagged values of the "exogenous" tax changes, they obtain estimates implying tax multipliers of −2.5 to −3 at 3 years. Leigh et al. (2010) use a similar narrative method to study fiscal consolidations across countries.[ee] Cloyne (2013) uses this method to identify exogenous tax shocks in the United Kingdom.

Favero and Giavazzi (2012) embed the R&R series in a somewhat restricted VAR and find smaller multipliers. In a series of papers, Mertens and Ravn (2011b, 2012, 2013, 2014) exploit the R&R narrative tax information in a way that significantly expands our understanding of the effects of tax shocks on the economy. I will focus on several of their contributions in this section and discuss the others in the next section. First, Mertens and Ravn (2011b, 2012) split the Romer and Romer series into anticipated vs unanticipated shocks based on the delay between the passing of the legislation and the implementation of the legislation. R&R had timed all of their shocks to coincide with the implementation rather the legislation. I will discuss the findings using unanticipated shocks here and discuss the findings using anticipated shocks below. Second, in their 2013

[ee] The Leigh et al. attempts to address measurement concerns in the very large literature on the effects of fiscal consolidations across countries, perhaps best exemplified by Giavazzi and Pagano (1990, 1996), Alesina and Perotti (1995, 1997), and Alesina and Ardagna (1998, 2010).

paper, Mertens and Ravn (2013) decomposed the unanticipated parts of the R&R series into personal income tax changes and corporate income tax changes and showed the differences in the two types of cuts on the economy. In their 2014 paper, Mertens and Ravn (2014) reconciled the Blanchard and Perotti SVAR estimates with the narrative estimates by introducing the proxy SVAR method discussed in detail in previous sections.

As discussed in Section 2.3.5, Mertens and Ravn's (2014) proxy SVAR provides a new method for identifying shocks using external instruments. In particular, they regress the reduced form residual of GDP, η_{3t}, from Eq. (20) on the reduced form residual of taxes, η_{1t}, using the R&R shock as an instrument. This leads to an unbiased estimate of b_{31} (since it is assumed that η_{2t} is the structural shock to government spending, which is uncorrelated with the other structural shocks). We can then use the estimated residual from that regression as one of the instruments in the regression of η_{1t} on η_{2t} and η_{3t}. This regression identifies b_{12} and b_{13}. When they implement their method, they estimate $b_{13} = 3.13$ with a 95% confidence band of (2.73, 3.55). Thus, their results suggest that Blanchard and Perotti's preset estimate of $b_{13} = 2.08$ is too low. Setting the output elasticity of tax revenue too low results in estimated tax shocks that include a reverse causality component (ie, there is a positive correlation between the cyclical components of taxes and output because of the positive causal effect of output on tax revenues). This is also an excellent illustration of Caldara and Kamps' (2012) insight on the link between the assumed structural tax elasticity and the estimated multipliers. Table 6 shows various tax multiplier estimates from the literature.

Mertens and Ravn (2013) split the unanticipated Romer shocks into changes in personal income tax rates vs corporate income tax rates. They find that cuts in either tax rate have positive effects on output, employment, hours, and the tax base. Interestingly, a cut in the corporate tax rate does not decrease corporate tax revenues because the corporate income tax base responds so robustly. Personal income tax cuts tend to raise consumption and investment more than corporate income tax cuts do. See figures 2, 9, and 10 of Mertens and Ravn (2013) for more detail.

Oh and Reiss (2012) highlight the importance of transfers in the stimulus packages adopted in response to the Great Recession. They formulate a heterogeneous agent model and explore the predicted multipliers on transfers. There has been, however, very little empirical work on the multipliers associated with government transfers.[ff] A major challenge has been identifying exogenous movements in transfers. Hausman (2016) studied the large veteran's bonus of 1936, equaling 2% of GDP, and found that it led to immediate effects on consumption spending. His calculations suggest that it led to faster GDP growth in 1936, but followed by a quick reversal in growth in 1937. Romer and Romer (2016) study the effects of changes in Social Security benefit payments in aggregate US

[ff] There is a large literature on the effects of various transfers on individual household consumption and saving. However, these estimates do not translate directly to aggregate multipliers.

Table 6 Summary of some tax multiplier estimates for the aggregate Unites States

Study	Main sample	Identification	Implied tax multiplier
Evans (1969)	Quarterly, 1966–1974	Based on estimates of equations of Wharton, Klein–Goldberger, and Brookings models	−0.5 to −1.7, depending on horizon, type of tax, and model
Blanchard–Perotti (2002)	Quarterly, 1960–1997	Assumed output elasticities in an SVAR. "Taxes" are actually taxes less transfers	−0.78 to −1.33 (peak to impact)
Mountford–Uhlig (2009)	Quarterly, 1955–2000	Sign restrictions on a VAR. Use same variables as BP	−5 for a tax increase that reduces the deficit
Romer–Romer (2010)	Quarterly, 1947–2007	Legislated tax changes driven by an inherited government budget deficit or to promote future growth, based on narrative evidence	−3, based on peak effect. Romer–Romer (2009) show that these tax shocks do not raise government spending significantly, so these are close to pure tax shocks
Barro–Redlick (2011)	Annual, 1917–2006 and subsamples	Average marginal income tax rate	−1.1
Favero and Giavazzi (2012)	Quarterly, 1950–2006	Romer–Romer shocks embedded in an SVAR	−0.5
Caldara and Kamps (2012)	Quarterly, 1947–2006	SVAR using outside elasticities	−0.65 (peak to impact)
Mertens–Ravn (2014)	Quarterly, 1950–2006	Proxy SVAR using Romer–Romer unanticipated shocks	−3 at 6 quarters

data. They find very rapid responses of consumption to permanent changes in benefit, but the results dissipate within a few months. Moreover, there is no clear evidence of effects on aggregate output or employment.

Gechert (2015) conducts a meta-analysis of various types of multipliers. He finds that tax and transfer multiplier tend to be around 0.6–0.7.

4.2.1.2 Further Explorations

I now investigate the Mertens and Ravn (2014) reconciliation of the tax results in more detail. To do this, I first use Mertens and Ravn's (2014) specification, data, and sample. The specification is a trivariate SVAR with federal government spending, output, and federal tax revenue, all in real per capita logarithms.[gg] The SVAR includes four lags of the variables in addition to a quadratic trend and a dummy variable for the second quarter of 1975 (following Blanchard and Perotti, 2002). The tax shock is Mertens and Ravn's unanticipated shocks extracted from the R&R narrative, demeaned as they describe.

Fig. 6A shows the impulse responses for tax revenue and output from their proxy SVAR using their programs.[hh] The results show that a positive R&R tax shock that has an impact effect on tax revenues equal to 1% of GDP raises tax revenue for several quarters and then lowers it below zero (though not statistically different). Output falls significantly on impact and troughs around −3 after a year. The magnitude of the results is similar to those found by R&R (2010) with their entire exogenous series.

My further investigation reveals some potentially problems with instrument relevance, though. The first-stage regression of tax revenue on the unanticipated tax shock (controlling for the lags of the other variables in the VAR) has an F-statistic of 1.6 (based on robust standard errors), which suggests a possible problem with instrument relevance.[ii] Stock and Watson (2012) also noticed problems with first-stage F-statistics of some of these external instruments in their dynamic factor model.[jj] Of course, the feedback from GDP to tax revenues is a potential complication. An exogenous tax increase should raise revenue, holding GDP constant; however, the decline in GDP exerts a downward effect on tax revenues. Thus, perhaps it is better to think of the R&R tax shock as an instrument for tax *rates*. Ideally, one would use statutory rates, since the actual rate paid is partly endogenous (since a change in income can push an entity into a different tax bracket).

[gg] Blanchard and Perotti actually used *net* taxes, meaning taxes less transfers. I follow Mertens and Ravn and use taxes. One could augment the system to include transfers as a fourth variable and use Romer and Romer's (2014) narrative-based transfer shock series as an external instrument.

[hh] This is the same as Mertens and Ravn's (2014) figure 4 with the signs reversed to examine the effect of a tax increase.

[ii] These additional results are based on the same data definitions and specification as Mertens and Ravn (2014), but on revised data. The results are similar if I use their original data.

[jj] See Lundsford (2015) and Montiel Olea et al. (2015) for discussions of instrument relevance in the external instruments framework.
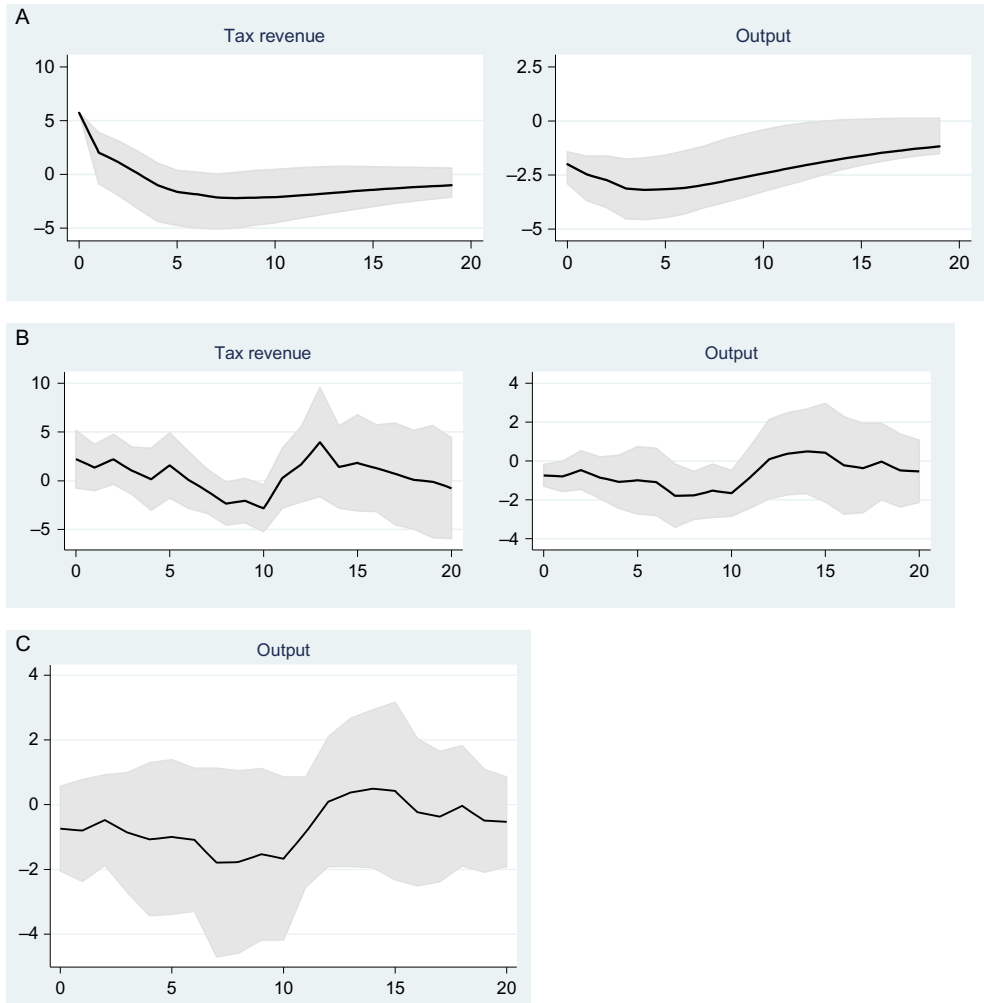
Fig. 6 Effects of unanticipated Romer tax shock, trivariate VAR, 1950q1–2006q4. (A) Mertens–Ravn (2014) proxy SVAR. (B) Jordà local projection, reduced form. (C) Jordà local projection, IV regression of output on tax revenue. Light gray bands are 90% confidence bands.

I do not have those data, so I simply construct an average tax rate as federal tax revenues divided by nominal GDP. I then estimate the first-stage regression described earlier with the average tax rate substituted for the log of taxes. The F-statistic on the R&R shock in this regression is 3.2, twice as high as the previous case but still well below the threshold for instrument relevance.

With the caveats about instrument relevance in mind, I further explore the robustness of Mertens and Ravn's (2014) results by estimating impulse responses using the Jordà local projection method and the Romer tax shock. I first estimate the reduced forms.

Table 7 Tax shock contribution to the forecast error variance of output

Horizon (in quarters)	Romer-Romer unanticipated	Leeper et al. (2012) anticipated tax series
0	1.6	0.4
4	0.4	5.7
8	0.5	4.8
12	1.1	4.4
16	1.8	4.3
20	2.1	4.3

Notes: Based on standard VAR with the shock, log output, log government spending, log taxes, and a quadratic trend. The shock is ordered first and four lags of the variables are included.

As discussed earlier, this involves regressing the dependent variable at $t + h$ on the shock at t, controlling for lags of other variables. To be consistent with Mertens and Ravn's specification, I use the same lags and variables in their proxy SVAR. Fig. 6B shows the impulse responses from the reduced form. Tax revenue increases in response to the shock initially and then falls below normal. In response to the tax shock, output falls on impact and then declines further to about −2 at 2 years, before beginning to recover. The confidence bands are wider, both because the Jordà method imposes fewer restrictions on the dynamics and because they incorporate the uncertainty about the impact of the tax shock on tax revenue. However, the point estimates for output for the first few years are broadly consistent with both Romer and Romer's (2010) original results and Mertens and Ravn's (2014) proxy SVAR.[kk] Table 7 shows the forecast error variance decomposition based on a standard VAR.[ll] Unanticipated tax shocks appear to account for very little of the forecast error variance of output.

As Mertens and Ravn (2014) note, however, external instruments tend to have measurement error, so they should not be used directly in an SVAR. A natural way to adjust for this in the Jordà setup is to estimate things as an IV regression, as discussed in Section 2. Thus, in a second specification I regress output at $t + h$ on the change in tax revenue at t, instrumented with the unanticipated part of the Romer tax shock, also controlling for the same variables as in the proxy SVAR (four lags of output, tax revenue, and government spending, as well as the deterministic terms). Fig. 6C shows the estimated impulse response of output for this specification. The point estimates for these results look very similar to those for output in Fig. 6B. The difference is that the confidence intervals are

[kk] If I use the Jordà method on the Romer's original specification and tax shock, I obtain results that are very close to theirs. This is as one would expect since they do not calculate impulses from a VAR.

[ll] As discussed in Section 4.1.3, although one can calculate forecast error variances using the estimated local projection coefficients, I found that the shares sometimes exceeded 100%. Thus, for present purposes I calculate the forecast error variance in a standard VAR with the shock, log government spending, log real GDP, and log taxes. The shock is ordered first and four lags are included, along with a quadratic trend.

very wide, always encompassing zero. Moreover, when I test whether the integral of the response for the first 12 quarters is different from zero, I cannot reject that it is zero.[mm]

To summarize, the literature on the effects of tax shocks has employed numerous methods, such as SVARs with calibrated elasticities, narrative approaches, and sign restrictions. Mertens and Ravn's (2014) reconciliation of some of the various approaches tends to support Romer and Romer's (2010) large estimated elasticities. My robustness checks suggest that while there might be a problem with instrument relevance, less restrictive ways to estimate impulse responses also generally support Romer and Romer's (2010) estimates of tax multipliers of −2 to −3.

4.2.2 News About Future Tax Changes
4.2.2.1 Summary of the Literature
Theory predicts that anticipated tax changes should have very different effects from unanticipated tax shocks (eg, Yang, 2005). If agents know that tax rates will increase in the future, they should respond by intertemporally substituting taxable activity into the present. Moreover, as discussed in Section 2, foresight about future tax changes can lead to identification problems in a standard SVAR. I will now review some recent results on the effects of anticipated tax changes on aggregate outcomes and provide some new results.

Mertens and Ravn (2011b, 2012) explore the effects of anticipated tax changes by splitting the Romers' narrative tax shock series into anticipated vs unanticipated shocks based on the delay between the passing of the legislation and the implementation of the legislation. Romer and Romer had timed all of their shocks to coincide with the implementation rather the legislation. Mertens and Ravn argue that the response of macroeconomic variables should be very different for anticipated vs unanticipated shocks.

Mertens and Ravn separate out the tax changes that were legislated more than 90 days before they were implemented. Because there are not a large number of these kinds of tax changes and because the lags between legislation and implementation vary significantly, Mertens and Ravn preserve the degrees of freedom in their estimation by combining various anticipated tax shocks according to the number of quarters left before implementation. Thus, their study does not trace out the effect of "news" per se; rather, it is more similar to an event study of the behavior of variables before and after the tax changes are implemented. Mertens and Ravn (2011b, 2012) estimate that anticipated and unanticipated tax shocks together account for 20% of the historical variation in output at business cycle frequencies. Particularly interesting is their finding that the so-called Volcker recession was in fact mostly caused by the Reagan tax cuts. The legislation was passed in summer 1981, but the actual tax cuts were phased in between 1982 and 1984. Mertens and Ravn's estimates imply that most of the decline in output from the second half of 1981 through 1982 was due to the negative output effects of anticipated future tax cuts.

[mm] Reducing the number of lags or control variables changes the results little.

Leeper, Richter, and Walker (2012) (LRW) construct an alternative measure of expected tax changes based on the spread between federal bonds and municipal bonds. They use their new series to inform their theoretical model but do not estimate effects of shocks to their series directly from the data. In the unpublished supplement to their 2013 *Econometrica* paper, Leeper et al. (2013) investigate the effect of their measure on output and show that expectations of a future tax increase raise output when the news arrives.

4.2.2.2 Further Explorations

I now explore the effects of several of the leading identified tax news shocks. Fig. 7 reproduces Mertens and Ravn (2011a,b) estimates of the effects of Romer tax shocks that were anticipated. Quarter 0 is the date of the implementation, negative quarters are quarters
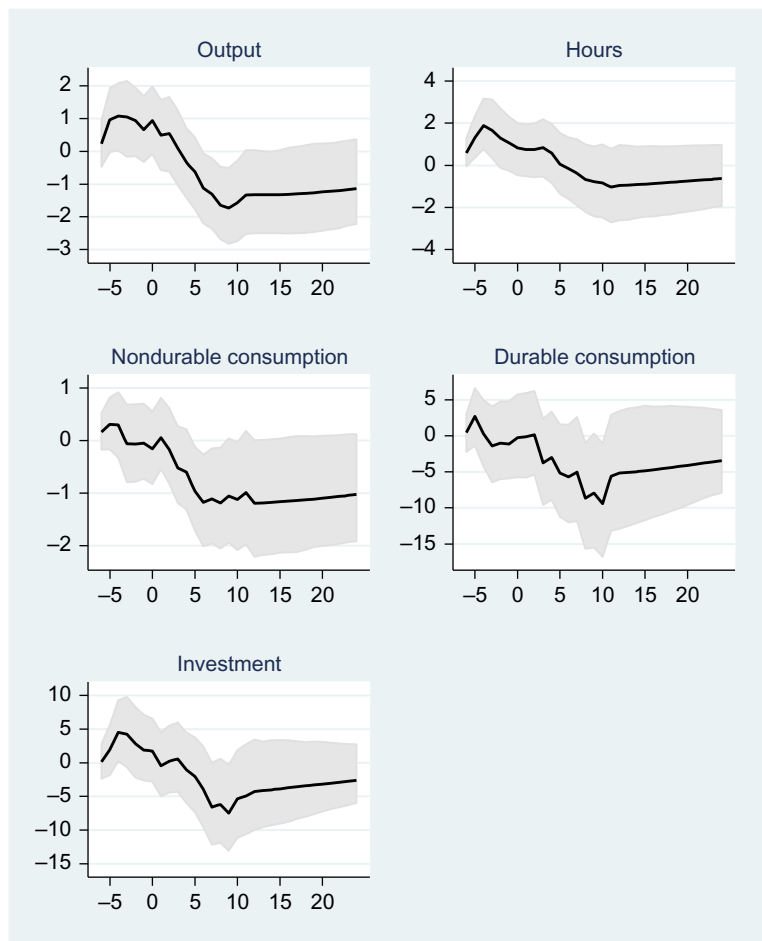

Fig. 7 Effects of news of future tax increases, Mertens–Ravn estimates based on Romer–Romer narrative, 1950q1–2006q4. Light gray bands are 90% confidence bands.

between the arrival of the news and before the implementation, and positive quarters are after the implementation. The graphs show clear evidence of anticipation effects and intertemporal substitution. Most variables, including output, hours, investment, and durable goods consumption expenditures, are higher than average in the interval between the announcement of a tax increase and its actual implementation. After implementation, all variables fall below normal, including nondurable consumption. Thus, the behavior of the data is very consistent with the theory.

To see how the results compare to Mertens and Ravn's results, I analyze the effects of Leeper et al.'s (2012) measure of average expected future tax rates from 1 to 5 years forward (AFTR15). Using a Jordà local projection, I estimate several sets of regressions at each horizon. In particular, I regress the endogenous variable of interest at $t+h$ on AFTR15 in period t, as well as on four lags of AFTR15, four lags of the endogenous variable and four lags of the average federal tax rate (total federal receipts divided by GDP). Because I do not orthogonalize the shock with respect to current values of any of the other variables, this identification scheme is similar to the one used by Leeper et al. (2013), where they order the tax news first in the Cholesky decomposition.

Fig. 8 shows the estimated responses to "news" that future tax rates will rise. The results are quite similar to those of Mertens and Ravn's results, even though the tax news variable is from a completely difference source and the model is estimated as responses to news rather than as an event study around the implementation. Output, hours, and investment start rising when news arrives at period 0 that tax rates will increase in the interval between 1 and 5 years. The variables remain high for a while and then fall below normal after a year or so.

Table 7 shows the forecast error variance decomposition for the LRW measure of expected tax changes. These shocks appear to account for more of the variance of output than the unanticipated tax changes, but still less than 6%.

In sum, some of the strongest and most robust findings in the fiscal literature are those associated with news about future tax changes. Expectations that future tax rates will increase lead to boom now followed by "busts." This is perhaps some of the strongest evidence that "news" can drive economic fluctuations.

4.3 Summary of Fiscal Results

In this section, I have summarized some of the main methods and findings concerning the effects of fiscal shocks. For both government spending and taxes, the methods that use external narrative series tend to find bigger effects on output than the more traditional SVAR method. For both government spending and taxes, anticipation effects are found to be very important.

Some of the literature has studied the effects of government spending and tax shocks jointly and made statements about "which" multiplier is larger. Some find larger
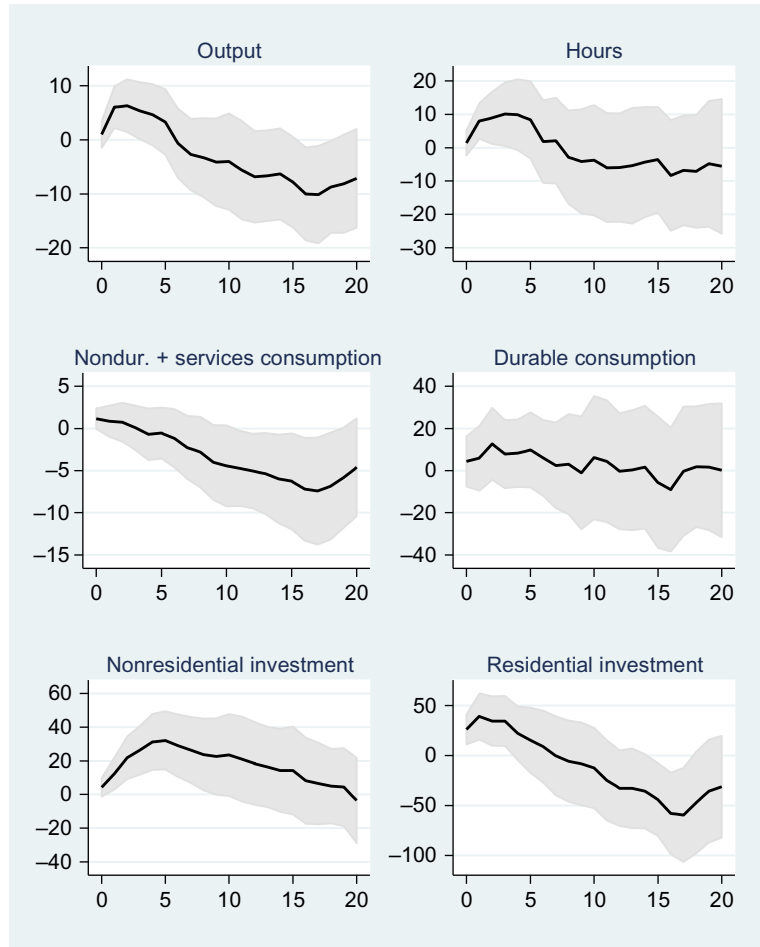
Fig. 8 Effect of news of future tax increase, Leeper et al. (2012) measure, Jordà local projection, 1954q1–2005q4. Light gray bands are 90% confidence bands.

government spending multipliers, and others find larger tax multipliers. My assessment is that the existing methods do not yield precise enough and robust enough estimates to be able to make this comparison.

5. TECHNOLOGY SHOCKS

Technology shocks are the most important type of nonpolicy shocks. In this section, I review the literature on technology shocks and present some new results comparing various shocks from the literature. I discuss both the classic unanticipated technology shocks and news about future changes in technology. I also distinguish between neutral and investment-specific technology (IST) shocks.

5.1 Neutral Technology Shocks

In 1982, Kydland and Prescott (1982) demonstrated the (then) surprising result that one could produce business cycle movements of key variables from a DSGE growth model beset by only one type of shock: variations in the growth rate of exogenous total factor productivity (TFP). To be concrete, consider the following aggregate production function:

$$Y_t = A_t F(L_t, K_t) \tag{21}$$

where Y_t is output, A_t is TFP, L_t is labor, and K_t is the capital stock at the beginning of period t. Neutral technology shocks, or TFP shocks, are shocks to the process driving A_t.

Several empirical regularities supported Kydland and Prescott's (1982) hypothesis. First, Solow (1957) showed that 87% of the growth in average labor productivity from 1909 to 1949 was due to TFP growth. If TFP growth was so important for growth, why not also for business cycles? Second, at the time that Kydland and Prescott published their article, a long-standing stylized fact was the procyclicality of labor productivity. In fact, this stylized fact was a problem for Keynesian "aggregate demand" explanations of business cycles, since diminishing returns would predict countercyclical labor productivity. Typically, the aggregate demand-driven business cycle literature had to resort to stories of labor hoarding or increasing returns to explain the procyclicality of labor productivity.

In follow-up work, Prescott (1986) used the Solow residual as his measure of exogenous TFP and used the standard deviation of that series along with his model to argue that the bulk of business cycle fluctuations could be explained by technology shocks. Beginning in the 1990s, though, several new results emerged that cast doubt on using the Solow residual as an exogenous technological progress for the purpose of business cycle analysis. First, Evans (1992) showed that variables such as money, interest rates, and government spending Granger-caused the Solow residual. Second, Hall (1988, 1990) developed a generalization of the Solow residual framework that relaxed the assumptions of competition and constant returns to scale. This framework demonstrated how endogenous components could enter the Solow Residual. Third, a number of papers, such as Shapiro (1993), Burnside et al. (1995), and Basu and Kimball (1997), used proxies such as the workweek of capital, electricity, or average hours to adjust the Solow residual for variations in the utilization of labor and/or capital. They found that much of the procyclicality of the Solow residual disappeared once it was adjusted.

Two approaches called into question whether technology shocks even led to business cycle-like movements. Galí (1999) and Basu et al. (2006) used different methods but both found results, suggesting that a positive technology shock led to a decline in labor inputs, such as hours. Both of these analyses assumed that all technology shocks were neutral technology shocks. I will discuss each of the approaches with the follow-up work in turn.

Galí (1999) used long-run restrictions to identify neutral technology shocks. He argued that a standard RBC model predicted that technology shocks were the only shocks that could have permanent effects on labor productivity. As discussed in

Section 2.3.6, Galí (1999) estimated a bivariate VAR with labor productivity and hours (or employment) and imposed the long-run restriction that technology shocks were the only shocks that could have a permanent effect on labor productivity. Francis and Ramey (2005) derived additional long-run restrictions from the theory and used them as an over-identification test and found that one could not reject the overidentifying restrictions. Francis and Ramey (2006) constructed new historical data for the United States and extended the analysis back to 1889. They found that a positive technology shock raised hours in the pre-WWII period but lowered them in the post-WWII period. They explained the switch with the difference in the serial correlation properties of productivity. In the early period, an identified technology shock raised productivity immediately, whereas in the later period an identified technology shock raised productivity more gradually. This gradual rise in the later period provides an incentive to reduce hours worked in the short run in anticipation of higher productivity in the long run.

Galí (1999) and Francis and Ramey (2005) both assumed that both (log) labor productivity and hours had a unit root and that their first differences were stationary. As Section 2.3.6 demonstrated, imposing long-run restrictions also requires the imposition of assumptions on stationarity. Christiano et al. (2003) argued that it makes no sense to model hours per capita as having a unit root since it is bounded above and below. They showed that if instead one assumes that hours are stationary and then impose the Galí long-run restriction, a positive technology shock leads to a rise in hours worked. Fernald (2007) noted the structural break in labor productivity growth, and when he allowed for that feature of the data, he found that hours fell after a positive technology shock. Francis and Ramey (2009) argued that the baby boom led to low-frequency movements in both labor productivity growth and hours worked per capita and that failure to correct for these led to the positive correlations found by Christiano et al. When they corrected for demographics, they found that a positive technology shock led to a decrease in hours. Gospodinov et al. (2013) discuss various econometric issues that arise in this setting with low-frequency movements.

Building on ideas of Uhlig (2003), Francis et al. (2014) introduced a new method of imposing long-run restrictions that overcame many of these problems. They identify the technology shock as the shock that maximizes the forecast error variance share of labor productivity at some finite horizon h. Using that scheme, they find that their identified technology leads to a fall in hours worked. They estimate that technology shocks contribute 15–40% of the forecast error variance of output at business cycle horizons. A variation by Barsky and Sims (2011) identifies the technology shock as the one that maximizes the *sum* of the forecast error variances up to some horizon h.

Several papers have questioned Galí's (1999) basic identifying assumption that technology shocks are the only shocks that have a long-run effect on labor productivity. Uhlig (2004) argues that capital taxation and shifts in preferences involving "leisure in the workplace" can also have long-run effects on labor productivity. He also introduces a

"medium-run" identification procedure that anticipates the procedures discussed earlier. He finds that the impact effect on hours is zero and that there is a small rise afterward. Mertens and Ravn (2011a) include the Romer and Romer (2010) exogenous tax shocks in a vector error correction model and find that once taxes are controlled and cointegration is allowed, a positive TFP shock raises hours in the short run. They also find that technology shocks account for 50–55% of the forecast error variance of output.

Basu et al. (2006) found that technology shocks were contractionary using a completely different method. Employing theoretical insights from Basu and Kimball (1997), they adjusted the annual Solow residual for utilization using hours per worker as a proxy. When they examined shocks to this purged Solow residual, they found that positive shocks to technology led to a decline in hours worked. Fernald (2014) has now constructed a quarterly version of this utilization-adjusted TFP series.[nn]

Alexopoulos (2011) identified technology shocks by creating an entirely new data series for measuring technology. Meticulously collecting and counting book publications for several types of technologies, she constructed several annual series on new technologies. She found that these series were not Granger-caused by standard macroeconomic variables. Using her new series in VARs, she found that a positive technology raises output and productivity. Contrary to the findings of Galí (1999) and Basu et al. (2006), she estimated that a positive shock to technology raises output, though the effect is weak.

Table 8 summarizes some of the estimates of the contribution of TFP shocks to output fluctuations at business cycle frequencies, based on approaches that identify technology shocks in time series models.

Numerous papers have identified technology shocks through estimated DSGE models. McGrattan (1994) estimated a neoclassical DSGE model with technology and fiscal shocks. Smets and Wouters (2007) estimated a New Keynesian DSGE model using Bayesian methods in order to explore the effects of various shocks. They incorporate a number of different shocks in the model, including neutral technology shocks, IST shocks (discussed in the next section), monetary shocks, government spending shocks, markup shocks, and risk premium shocks. Their estimates imply that a positive neutral technology shock lowers hours worked. Justiniano et al. (2010, 2011) also estimate a New Keynesian model, but incorporate also investment-specific shocks and MEI shocks. Schmitt-Grohe and Uribe (2012) estimate a DSGE model which allows all of their shocks to have an unanticipated component and a "news," or unanticipated, component. Miyamoto and Nguyen (2015) extend their estimation method by including series on survey expectations in the estimation. I will discuss these papers in more detail in the sections on IST shocks and news. Blanchard et al. (2013) estimate a DSGE model allowing for both "news" and "noise." Table 9 summarizes the estimates from DSGE models

---

[nn] The series is regularly updated and made available by John Fernald at http://www.frbsf.org/economic-research/economists/jfernald/quarterly_tfp.xls.

Table 8 Estimated importance of technology shocks in SVAR models

Study	Method	Type	News?	% of output explained
Galí (1999), Francis–Ramey (2005)	Long-run restrictions, hours in first differences	TFP	No	Very little
Christiano et al. (2004)	Long-run restrictions, hours in levels	TFP	No	31–45% for horizons up to 20 quarters
Christiano et al. (2004)	Long-run restrictions, hours in first differences	TFP	No	1–17% for horizons up to 20 quarters
Basu et al. (2006)	Utilization and effort adjusted TFP	TFP	No	17–40% from 1 to 3 years
Beaudry and Portier (2006)	Short-run or long-run restrictions	TFP	Yes	50%
Fisher (2006)	Long-run restrictions involving both labor productivity and investment goods prices	TFP	No	32% at 12 quarters (see papers for more details)
Fisher (2006)	Long-run restrictions involving both labor productivity and investment goods prices	IST	No	26% (49%) at 12 quarters in early (late) sample
Mertens and Ravn (2010)	Long-run restrictions, cointegration, include taxes	TFP	No	50–55% at business cycle frequencies
Barsky and Sims (2011)	Medium-horizon restrictions	TFP	Yes	9–43% for horizons up to 24 quarters.
Barsky and Sims (2011)	Medium-horizon restrictions	TFP	No	6–20% for horizons up to 24 quarters
Francis et al. (2014)	Medium-horizon restrictions	TFP	No	15–40% for horizons up to 32 quarters.
Francis et al. (2014)	Long-run restrictions	TFP	No	40–55% for horizons up to 32 quarters
Ben Zeev and Khan (2015)	Medium-horizon restrictions	IST	Yes	73% at 8 quarters
Ben Zeev and Khan (2015)	Medium-horizon restrictions	IST	No	Very little
Ben Zeev and Khan (2015)	Medium-horizon restrictions	TFP	No	10% at 8 quarters

Notes: TFP denotes neutral total factor productivity technology, IST denotes investment-specific technology, and MEI denotes marginal efficiency of investment.

Table 9 Estimated importance of technology shocks in DSGE models

Study	Model features	Type	News?	% of output explained at business cycle frequencies
Prescott (1986)	Calibrated neoclassical DSGE model	TFP	No	75%
McGrattan (1994)	Neoclassical model with distortionary taxes and government spending	TFP	No	41%
Greenwood et al. (2000)	Calibrated DSGE model, technology identified with relative price of investment	IST	No	30%
Smets and Wouters (2007)	New Keynesian model with many types of shocks	TFP	No	15–30% from horizon 1–10 quarters
Justiniano et al. (2011)	New Keynesian model with many types of shocks	TFP	No	25%
Justiniano et al. (2011)	New Keynesian model with many types of shocks	IST	No	0%
Justiniano et al. (2011)	New Keynesian model with many types of shocks	MEI	No	60%
Schmitt–Grohe and Uribe (2012)	Distinguishes unanticipated vs anticipated, TFP vs investment specific, no sticky prices	TFP	No	25%
Schmitt–Grohe and Uribe (2012)	Distinguishes unanticipated vs anticipated, TFP vs investment specific, no sticky prices	TFP	Yes	3%
Schmitt–Grohe and Uribe (2012)	Distinguishes unanticipated vs anticipated, TFP vs investment specific, no sticky prices	IST	No	21%
Schmitt–Grohe and Uribe (2012)	Distinguishes unanticipated vs anticipated, TFP vs investment specific, no sticky prices	IST	Yes	7%
Khan and Tsoukalas (2012)	New Keynesian model, distinguishes unanticipated vs anticipated	TFP	No	24%
Khan and Tsoukalas (2012)	New Keynesian model, distinguishes unanticipated vs anticipated	MEI	No	47%
Khan and Tsoukalas (2012)	New Keynesian model, distinguishes unanticipated vs anticipated	IST	No	1.2%
Khan and Tsoukalas (2012)	New Keynesian model, distinguishes unanticipated vs anticipated	TFP + MEI + IST	Yes	1.4%
Miyamoto and Nguyen (2013)	Extends Schmitt–Grohe–Uribe analysis by using data on expectations	TFP	No	19%
Miyamoto and Nguyen (2013)	Extends Schmitt–Grohe–Uribe analysis by using data on expectations	TFP	Yes	7%
Miyamoto and Nguyen (2013)	Extends Schmitt–Grohe–Uribe analysis by using data on expectations	IST	No	27%
Miyamoto and Nguyen (2013)	Extends Schmitt–Grohe–Uribe analysis by using data on expectations	IST	Yes	12%

Notes: TFP denotes neutral total factor productivity technology, IST denotes investment-specific technology, and MEI denotes marginal efficiency of investment.

of the contribution of various types of technology shocks to output fluctuations at busi-
ness cycle frequencies.

5.2 Investment-Related Technology Shocks

Greenwood et al. (1988) were the first to examine in a DSGE model Keynes' idea that
shocks to the marginal efficiency of investment (MEI) could be a source of business cycle
volatility. In follow-up work, Greenwood et al. (2000) used a calibrated DSGE model to
examine the importance of IST change in business cycles. They used the relative price of
new equipment to identify the process driving IST shocks and concluded that these
shocks could account for 30% of business cycle volatility.

Fisher (2006) extended Galí's (1999) analysis of neutral technology shocks by incor-
porating additional data and restrictions that separately identify neutral and IST shocks. In
particular, he assumed that only IST shocks affect the relative price of investment goods in
the long run and only neutral technology and IST technology shocks affect labor pro-
ductivity in the long run. Because of some sample instability, he estimated his model over
two periods: 1955q1–1979q2 and 1982q3–2000q4. He found that both technology
shocks together accounted for a substantial shared of the forecast error variance of output,
60% at 12 quarters in the early sample, 83% in the later sample.

Justiniano et al. (2010, 2011) estimate a New Keynesian DSGE model with a variety
of unanticipated shocks. Justiniano et al. (2011) distinguish between IST shocks and MEI
shocks. To be concrete, consider simplified versions of two equations from their DSGE
model:

$$I_t = \Psi_t Y_t^I \tag{22a}$$

$$K_{t+1} = (1-\delta)K_t + \mu_t I_t \tag{22b}$$

I_t is the production of investment goods and Ψ_t denotes the rate of transformation of final
goods, Y_t^I, into investment goods. Ψ_t is IST which, according to their model, should be
equal to the inverse of the relative price of investment goods to consumption goods. K_{t+1}
is the level of capital at the beginning of period $t+1$, δ is the depreciation rate, and μ_t is the
rate of transformation between investment goods and installed capital, or the *MEI*.
Previous research, such as by Greenwood et al. (2000) and Fisher (2006), had not distin-
guished IST from MEI and had assumed their product was equal to the inverse of the
relative price of investment goods. Justiniano et al. (2011) estimate that (unanticipated)
MEI shocks contribute 60% of the variance of output at business cycle frequencies.

Schmitt-Grohe and Uribe (2012), Khan and Tsoukalas (2012), and Miyamoto and
Nguyen (2015) estimate DSGE models that incorporate both TFP and IST shocks.
An important focus of their estimation is the distinction between unanticipated technol-
ogy changes and news about future changes, so I will discuss their work in the next
section on news.

Although there is a wide range of results, a general pattern that emerges is that when models include IST and/or MEI shocks, they tend to explain a significant portion of the variation in output at business cycle frequencies.

5.3 News About Future Technology Changes

Both Pigou (1927) and Keynes (1936) suggested that changes in expectations about the future may be an important driver of economic fluctuations. Beaudry and Portier (2006) reignited interest in the idea by providing time series evidence that news about future productivity could explain half of output fluctuations over the business cycle. Further-more, their estimates implied that hours and output rose when the news arrived, thus creating business cycle-type comovements. They identified news shocks using two methods; both involved identifying shocks that moved stock prices immediately, but affected productivity only with a lag. Beaudry and Lucke (2010) and Kurmann and Otrok (2013) used other techniques to reach similar conclusions. More recently, how-ever, Barsky and Sims (2011) and Barsky et al. (2014) have used medium-run restrictions and series on consumer confidence to identify news shocks and found that news shocks did not generate business cycle fluctuations. In particular, hours fell when news arrived. Fisher (2010), Kurmann and Mertens (2014), and Forni et al. (2014) have highlighted problems with Beaudry and Portier's identification method. For example, Kurmann and Mertens (2014) show that the larger VECM in Beaudry and Portier's (2006) paper is not identified. Forni et al. (2014) argue that the small-scale SVARs are affected by the "nonfundamentalness" problem discussed in Section 2.5. Thus, the empirical work based on time series identification is in flux. Beaudry and Portier (2014) present a comprehensive summary of the literature.

I would add that another potential problem with Beaudry and Portier's (2006) method for identifying TFP news shocks is the implicit assumption they make about stock prices. They assume that the future profits from the TFP shock will show up in current stock prices. It is not clear that this assumption holds for major innovations. Greenwood and Jovanovic (1999) and Hobijn and Jovanovic (2001) present theory and evidence that major technological innovations (such as information technology) actually lead to a temporary *decline* in stock market values because they lower the value of the existing technology. Revolutionary innovations usually arise in private companies and claims to future dividend streams only show up in stock prices after the initial public offerings. Thus, we should not necessarily see positive effects of news about future TFP on stock prices.

Ben Zeev and Khan (2015) identify both unanticipated IST shocks and IST news shocks. To do this, they extend Barsky and Sims (2011) medium-horizon restriction method for identifying news and employ Fisher's (2006) assumptions linking IST and the relative price of investment goods. They find that IST news shocks explain 73% of the forecast error variance of output at a horizon of eight quarters. They show that the IST shocks originally identified by Fisher (2006) were a combination of the

unanticipated IST shocks and news about IST. Ben Zeev and Khan's paper thus corroborates Fisher's finding that IST shocks are the major source of fluctuations, but goes on to show that it is the news part that is the most important.

Another strategy for identifying news is through estimation of a DSGE model, as pioneered by Schmitt-Grohe and Uribe (2012). This approach achieves part of its identification by assuming particular lags between the arrival of news and the realization of the change. Schmitt-Grohe and Uribe (2012) allow for a variety of unanticipated and news shocks for variables such as TFP, IST, and wage markups. They estimate that all news variables combined (including nontechnology shocks such as wage markup shocks) account for 50% of output fluctuations according to their estimates. An extension by Miyamoto and Nguyen (2015) uses actual survey forecasts for the expectations variables.

Khan and Tsoukalas (2012) estimate a New Keynesian DSGE model with both IST and MEI shocks and allow for both unanticipated changes and news shocks. They find that unanticipated MEI shocks contribute an important part of the variance of output (47%), but that technology news shocks are not important at all. Nontechnology news shocks do, however, contribute to the variance decomposition of hours. In particular, wage markup shocks account for over 40% of the variance of hours. Thus, their results on the importance of unanticipated technology shocks contrast with those of Schmitt-Grohe and Uribe (2012), but their results on the importance of news about wage markups are consistent with their findings. The estimates of the importance of technology news are summarized in Table 9.

5.4 Explorations with Estimated Technology Shocks

I now study the relationship between some of the leading shocks and explore the effects of a few of them in the Jordà local projection framework. I reestimate the Galí (1999), Christiano et al. (2003) (CEV), and Beaudry and Portier (2006) systems using updated data. In each case, I used a simple bivariate system. Both the Galí and CEV shocks use long-run restrictions, with the former assuming a unit root in hours per capita and the latter assuming a quadratic trend in hours. I use Beaudry and Portier's shock estimated with the short-run restriction; ie, it is the shock to stock prices that does not affect TFP on impact; the correlation of this shock with their shock estimated using long-run restrictions is 0.97. The Fernald shocks are simply the growth rate of Fernald's (2014) utilization-adjusted TFP for the aggregate economy or for the investment goods sector. The rest of the estimated shocks were kindly provided by Francis et al. (2014) (medium-horizon restrictions), Barsky and Sims (2011) (medium-run restrictions, consumer confidence), Justiniano et al. (2011) (estimated DSGE model), Ben Zeev and Khan (2015) (SVAR with medium-run restrictions), and Miyamoto and Nguyen (2015) (estimated DSGE model with forecast data). The joint sample is 1955q2–2006q4, except for the TFP news sample, which is limited to 1961q1–2006q4 by the Barsky and Sims shock availability. Correlations between subsets of shocks that are available over longer samples are similar to those reported for the joint sample. Table 10 shows the correlations, broken

Table 10 Correlation of various estimated technology shocks (sample is 1955q2–2006q4)
A. Unanticipated TFP shocks

	gali_tfp	cev_tfp	jf_tfp	ford_tfp	bzk_tfp	jpt_tfp	mn_tfp_p	mn_tfp_s
gali_tfp	1.00							
cev_tfp	0.62	1.00						
jf_tfp	0.68	0.42	1.00					
ford_tfp	0.75	0.62	0.62	1.00				
bzk_tfp	0.67	0.78	0.54	0.63	1.00			
jpt_tfp	0.68	0.69	0.53	0.54	0.63	1.00		
mn_tfp_p	0.17	0.16	0.20	0.28	0.08	0.16	1.00	
mn_tfp_s	0.52	0.59	0.47	0.52	0.58	0.62	0.10	1.00

B. TFP news shocks

	bp_news	bs_news	mn_p_n4	mn_p_n8	mn_s_n4	mn_p_n8
bp_news	1.00					
bs_news	0.25	1.00				
mn_p_n4	0.08	0.12	1.00			
mn_p_n8	0.05	0.00	0.29	1.00		
mn_s_n4	0.04	−0.04	0.53	−0.14	1.00	
mn_p_n8	0.05	0.00	0.29	1.00	−0.14	1.00

C. Unanticipated IST or MEI shocks

	jf_ist	bzk_ist	jpt_mei	jpt_ist	mn_ist_p	mn_ist_s
jf_ist	1.00					
bzk_ist	0.17	1.00				
jpt_mei	−0.27	0.05	1.00			
jpt_ist	0.19	0.49	−0.01	1.00		
mn_ist_p	0.03	0.31	0.17	0.20	1.00	
mn_ist_s	−0.13	0.11	0.27	0.14	−0.06	1.00

D. IST news shocks

	bzk_news	mn_p_n4	mn_p_n8	mn_s_n4	mn_s_n8
bzk_news	1.00				
mn_p_n4	0.15	1.00			
mn_p_n8	0.02	0.18	1.00		
mn_s_n4	0.12	0.07	0.12	1.00	
mn_s_n8	0.08	0.01	0.02	0.28	1.00

Abbreviations: *bp*, Beaudry-Portier; *bs*, Barsky–Sims; *bzk*, Ben Zeev and Khan; *cev*, Christiano, Eichenbaum, Vigfusson; *ford*, Francis, Owyang, Roush, DiCecio; *gali*, Gali; *ist*, investment-specific technology; *jf*, John Fernald; *jpt*, Justiniano, Primiceri, Tambolotti; *mei*, marginal efficiency of invest; *mn*, Miyamoto and Nguyen; *n4*, news with 4 quarter lead; *n8*, news with 8 quarter lead; *_p*, permanent; *_s*, stationary; *tfp*, total factor productivity.

down according to whether the shock is to TFP or IST (or MEI) and whether it is unanticipated or is news.

Table 10A shows the results for unanticipated TFP shocks, which have received the most attention. Most of the shocks have a correlation above 0.6 with the shock estimated using Galí's (1999) method. The exception is the Miyamoto and Nguyen (2015) permanent TFP shock. It is surprising that the Miyamoto and Nguyen stationary TFP shock has a higher correlation than the permanent TFP shock, since the Galí method only identifies permanent TFP shocks.

Table 10B shows news shocks about TFP. The correlation between Beaudry and Portier's (2006) shock estimated using short-run restrictions and Barsky and Sims' (2011) shock estimated using medium-horizon restrictions is only 0.25. The correlations of both of those SVAR-based shocks with the Miyamoto and Nguyen (2015) DSGE-based shocks are essentially 0.

Table 10C shows correlations of various estimates of unanticipated IST or MEI shocks. The correlations between the various estimates are quite low. For example, the correlation between Fernald's utilization-adjusted TFP for the investment goods sector and Justiniano et al.'s (2011) IST shock is only 0.19 and is −0.27 for Justiniano et al.'s MEI shock. The highest correlation of 0.49 is between Justiniano et al.'s IST shock and Ben Zeev and Khan's (2015) IST shock. The higher correlation is not surprising since both methods associate the IST shock with the inverse of the relative price of equipment.

Table 10D shows correlations of various estimates of IST news shocks. There is essentially no correlation between Ben Zeev and Khan's (2015) SVAR-based estimates and Miyamoto and Nguyen's (2015) DSGE-estimated shocks.

If we were simply trying to develop instruments for estimating structural parameters, it would not matter if various instruments had low correlation.[oo] If, however, we are trying to estimate shocks in order to determine their importance for macroeconomic fluctuations, a low correlation across various estimates is troubling. The large number of low correlations across methods and the widely varying results reported across papers suggest that we are still far from a consensus on the nature and importance of technology shocks. The problem is not one of lack of consensus of estimated DSGE vs SVAR methods. Even within each class of method, the results vary widely, as evidenced in Tables 8 and 9.

Moreover, many of the estimated shocks do not satisfy the property that they are unanticipated or news. Table 11 shows the p-values from two sets of tests. The first one tests for serially correlation of the shock by regressing the shock on its own two lags and testing their joint significance. The SVAR-estimated shocks do well on this test, but quite a few of the DSGE-estimated shocks fail this test. The second set of tests is for Granger causality (Granger, 1969). To conduct these tests, I augment this regression with two lags each of log real GDP, log real consumption, and log real stock prices. I chose

---

[oo] Sims (1998) made this argument in his discussion of Rudebusch's (1998) monetary shock critique.

Table 11 Tests for serial correlation and Granger causality

	_p_-Value test for significance of own lags	_p_-Value test for Granger causality
gali_tfp	0.986	**0.020**
cev_tfp	0.986	**0.000**
jf_tfp	0.718	**0.001**
jf_ist	0.644	**0.000**
ford_tfp	0.991	0.855
bp_tfp_news_sr	0.999	0.910
bs_tfp_news	0.834	0.935
bzk_ist_news	0.724	**0.049**
bzk_ist	0.981	0.740
bzk_tfp	0.949	0.992
jpt_tfp	0.101	**0.000**
jpt_mei	**0.006**	**0.000**
jpt_ist	0.941	0.854
mn_tfp_p	0.133	**0.000**
mn_ist_p	**0.000**	0.287
mn_tfp_s	**0.010**	**0.008**
mn_ist_s	**0.000**	**0.024**
mn_tfp_p_n4	**0.000**	**0.001**
mn_tfp_p_n8	**0.000**	**0.087**
mn_ist_p_n4	**0.000**	0.924
mn_ist_p_n8	**0.000**	**0.076**
mn_tfp_s_n4	**0.098**	0.134
mn_tfp_s_n8	0.353	0.783
mn_ist_s_n4	**0.000**	0.497
mn_ist_s_n8	**0.000**	**0.052**

Notes: The tests for serial correlation are conducted by regressing the shock on its own two lags and testing the joint significance. The tests for Granger causality are conducted by regressing the shock on its own two lags, as well as two lags of log real GDP, log real consumption, and log real stock prices. The test is on the joint significance of the lags of the three additional variables. P-values less than 0.1 are indicated in bold.
Abbreviations: _bp_, Beaudry–Portier; _bs_, Barsky–Sims; _bzk_, Ben Zeev and Khan; _cev_, Christiano, Eichenbaum, Vigfusson; _ford_, Francis, Owyang, Roush, DiCecio; _gali_, Gali; _ist_, investment-specific technology; _Jpt_, Justiniano, Primiceri, Tambolotti; _mei_, marginal efficiency of invest; _mn_, Miyamoto and Nguyen;_n4_, news with 4 quarter lead; _p_, permanent; _s_, stationary; _tfp_, total factor productivity.

consumption and stock prices because those variables have forward-looking components to them. Half of the shocks fail this test. Of course, the Galí and CEV shocks were estimated using a simple bivariate model. Had I augmented those systems with these variables, the shocks would have passed the tests by construction. The Francis et al., Beaudry and Portier, and Ben Zeev and Kahn shocks pass this test, as do about half of the DSGE-estimated shocks.

Next, I explore the effects of a few of the estimated shocks on several macroeconomic variables in a Jordà local projection framework. To do this, I estimate the following series of regressions:

$$z_{t+h} = \alpha_h + \theta_h \cdot \text{shock}_t + \varphi_h(L)\gamma_{t-1} + \text{quadratic trend} + \varepsilon_{t+h} \qquad (23)$$

The z is the variable of interest. The control variables include two lags each of the shock (to mop up any serial correlation in the shocks), log real GDP per capita, log real stock prices per capita, log labor productivity (equal to real GDP divided by total hours worked), and the dependent variable. The coefficient θ_h gives the response of z at time $t+h$ to a shock at time t. As discussed in Section 2, ε_{t+h} will be serially correlated, so the standard errors must incorporate a correction, such as Newey–West.

Fig. 9 shows the responses of real GDP, labor productivity hours, stock prices, consumption, and nonresidential investment to three different measures of unanticipated TFP shocks: the Francis, Owyang, Roush, and DiCecio (2014) (FORD) measure, which
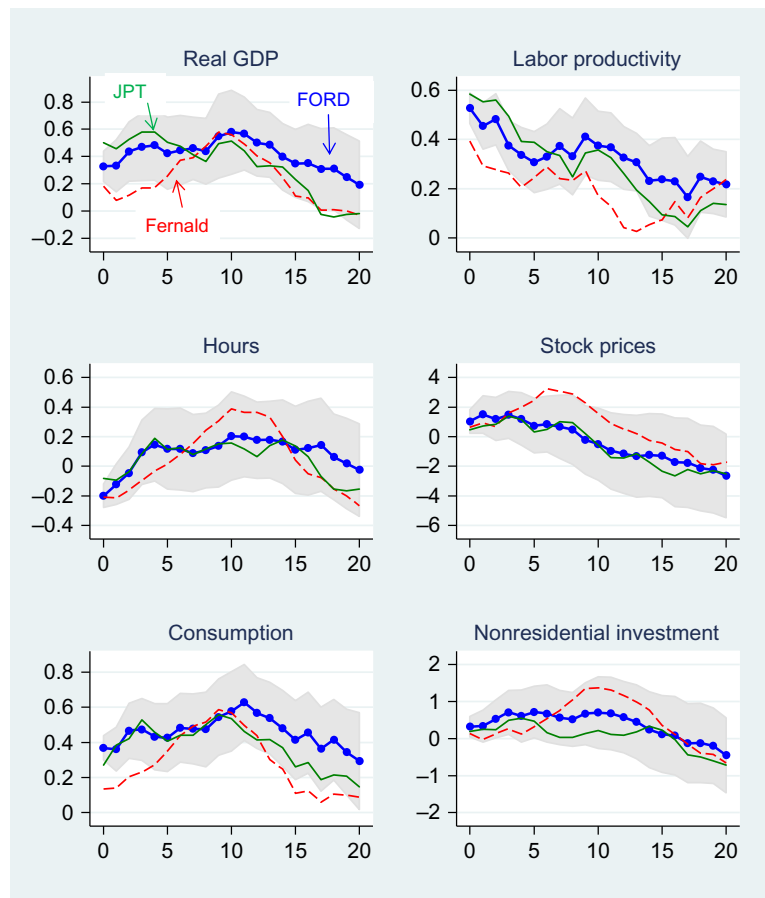

Fig. 9 Effects of TFP shock, Jordà local projection, various samples. Francis, Owyang, Roush, DiCecio (FORD): *blue lines with circles* (*black* in the print version); Fernald utilization-adj TFP: *dashed red* (*gray* in the print version) *lines*; Justiniano, Primiceri, Tambalotti (JPT) DSGE TFP: *solid green* (*gray* in the print version) *lines*. Light gray bands are 90% confidence bands.

uses medium-run restrictions; Fernald's (2014) utilization-adjusted TFP growth, and Justiniano, Primiceri, and Tambalotti's (2011) (JPT) estimate from their DSGE model. The responses to the FORD and JPT shocks are quite similar: GDP, labor productivity, stock prices, and consumption all jump immediately and significantly. Hours fall for a few quarters before rising. The Fernald shock implies a more hump-shaped response of GDP, hours, stock prices, consumption, and investment. Labor productivity rises immediately but then returns to normal at around 16 quarters. The Fernald shock also shows an initial decline in hours before they rise.

Fig. 10 shows the effects of the Ben Zeev and Khan (2015) IST news shock. Recall that this shock is an extension of Fisher's (2006) method, estimated using the Barsky–Sims (2011) medium-horizon method combined with information on relative prices of investment. This shock appears to generate a classic business cycle pattern. GDP, hours, stock
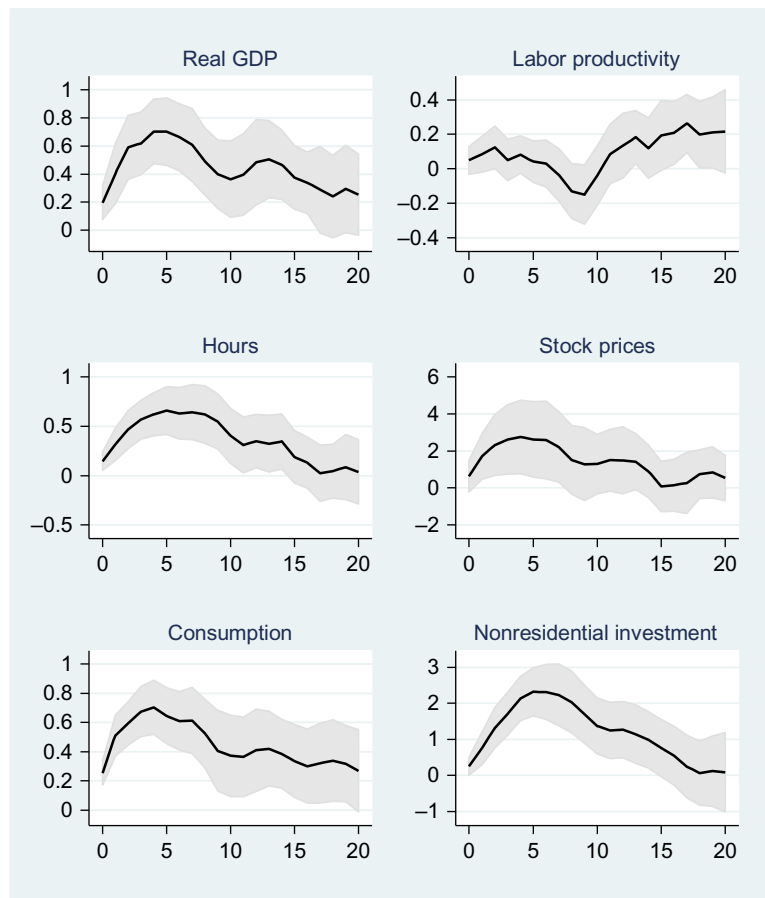

Fig. 10 Effects of news of investment-specific technology shocks, Ben Zeev–Khan (2015) measure, Jordà local projection, 1952q1–2012q1. Light gray bands are 90% confidence bands.

prices, consumption, and nonresidential investment increase with a prolonged hump shape. Labor productivity does nothing for about six quarters, falls around nine quarters, and then rises.

Fig. 11 shows the effects of JPT's MEI shock, estimated from their DSGE model. While this shock leads to temporary rises (for a year or less) in real GDP, labor productivity, consumption, and nonresidential investment, it leads to a fall in stock prices, which is puzzling.

Tables 12A and 12B show the forecast error variance decompositions for these five shocks for both output and hours. The decompositions are calculated from a standard VAR with the shock, and log per capita values of real GDP, hours, stock prices, consumption, and nonresidential investment. Although some of the unanticipated TFP shocks can account for up to 16% of output, none accounts for much of the variance
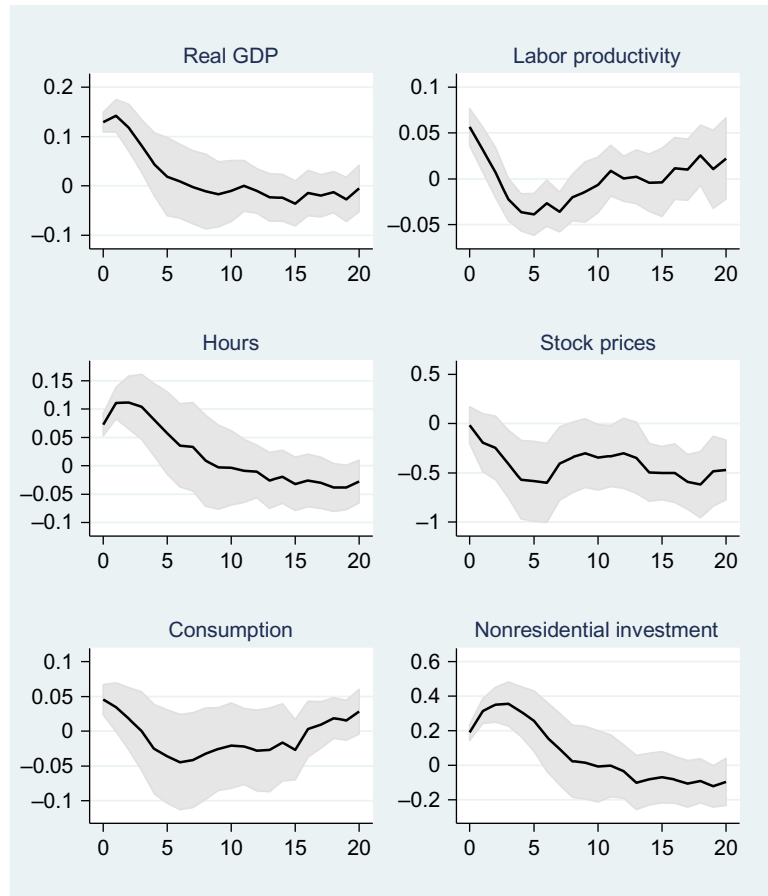

Fig. 11 Effects of marginal efficiency of investment shock, Justiniano et al. (2011) measure, Jordà local projection, 1954q3–2009q1. Light gray bands are 90% confidence bands.

Table 12A TFP shock contribution to the forecast error variance of output and hours

	FORD TFP		Fernald TFP		JPT TFP	
Horizon (in quarters)	Output	Hours	Output	Hours	Output	Hours
0	16.2	10.5	6.1	10.5	28.2	1.0
4	13.1	2.0	2.0	2.4	15.1	0.9
8	14.3	1.9	2.8	1.3	15.9	1.6
12	14.3	1.6	3.1	1.2	16.3	1.6
16	14.0	1.5	3.1	1.5	16.0	1.6
20	13.7	1.5	3.0	2.0	15.7	1.9

Notes: These results are based on a standard VAR with the shock, output, hours stock prices, consumption, and nonresidential investment. The shock is ordered first. Four lags are included, along with a quadratic trend.

Table 12B Investment-related technology shock contribution to the forecast error variance of output and hours

	Ben Zeev–Khan IST news		JPT MEI	
Horizon (in quarters)	Output	Hours	Output	Hours
0	7.8	6.9	49.6	26.4
4	33.2	31.3	19.8	20.9
8	36.8	38.5	11.9	12.1
12	36.8	38.8	11.4	10.5
16	36.4	37.9	11.3	10.1
20	35.9	36.8	11.1	9.8

Notes: These results are based on a standard VAR with the shock, output, hours stock prices, consumption, and nonresidential investment. The shock is ordered first. Four lags are included, along with a quadratic trend.

of hours. In contrast, the Ben Zeev and Khan IST news shock accounts for well over a third of the forecast error variance of both output and hours. JPT's MEI shock accounts for large fractions on impact, 50% for output and 26% for hours, but the effects dissipate fairly quickly.

5.5 Summary of Technology Shocks

The literature investigating the effects of technology shocks has moved far beyond the simple Solow residual. Various methods have been introduced to deal with changes in measured TFP that are due to variable utilization. Moreover, the literature has moved beyond neutral technology shocks to recognize the potential importance of IST shocks and MEI shocks. In addition, recent contributions have investigated the importance of news shocks.

My analysis shows, however, that some of the estimated shocks are not highly correlated with other versions. Moreover, many of the shocks are serially correlated or Granger-caused by other variables. This suggests that more research needs to be done

to refine these shock measures. One of the shocks that seem to be promising both for generating business-cycle comovement and for contributing significant amounts to the variance of output is the shock that captures news about IST change.

6. ADDITIONAL SHOCKS

So far, this chapter has focused on only three types of shocks—monetary, fiscal, and technology shocks. There are numerous other candidates for potentially important macroeconomic shocks. Here, I will briefly review a few.

One obvious additional candidate for a macroeconomic shock is oil shocks. Hamilton (1983) argued that oil supply shocks were a major driver of economic fluctuations. Since then, a large literature has examined the effects of oil supply shocks. One of the major themes of the literature is the changing estimated effects of oil price shocks, identified by ordering oil prices first in a linear VAR. In particular, after the 1970s oil price changes seemed to have smaller effects. One potential explanation is asymmetries. Several analyses, such as by Davis and Haltiwanger (2001) and Hamilton (2003), argued that oil price increases have larger effects than oil price decreases. Subsequent research, however, found that there was not strong evidence of asymmetry (eg, Kilian and Vigfusson, 2011). A second potential explanation for the changing effects of oil supply shocks is that the oil price increases in earlier periods were accompanied by price controls, which led to many distortions (Ramey and Vine, 2011). When they constructed an implicit cost of oil that incorporated a proxy for the distortion costs, they did not find much evidence of changing effects. A third explanation was by Kilian (2009) and was a critique of standard identification methods. He argued that many of the changes in oil prices are driven now driven by demand shocks, not supply shocks, so a standard Cholesky decomposition with oil prices ordered first does not properly identify oil supply shocks. Stock and Watson's (forthcoming) chapter in this *Handbook* uses oil shocks as a case study of their methods. They find that oil supply shocks, identified using Kilian's (2009) method, do not account for a significant fraction of the forecast error variance of output.

Credit shocks are another possible candidate for a macroeconomic shock. There is huge literature analyzing the importance of credit and credit imperfections in economic fluctuations and growth. Most of this literature focuses on credit as an important propagation and amplification mechanism (eg, "the credit channel" of monetary policy), rather than as an important independent source of shocks. Gilchrist and Zakrajšek's (2012) recent analysis of the effects of innovations to their new excess bond premium variable can be interpreted as an analysis of credit market shocks. They showed that innovations to the excess bond premium that were orthogonal to the current state of the economy had significant effects on macroeconomic variables. They interpret a negative "shock" to this variable as signaling a reduction in the effective risk-bearing capacity of the financial sector.

The role of uncertainty in the business cycle has received heightened attention recently. In addition to standard firm–level uncertainty and financial uncertainty, recent work has suggested a possible role for policy uncertainty. More research needs to be done to untangle uncertainty as an endogenous propagation mechanism vs uncertainty as an independent source of macroeconomic shocks.

Labor supply shocks are yet another possible source of macroeconomic shocks. It is well known that cyclical variations in the "labor wedge" are an important component of business cycles. Shapiro and Watson (1988) estimated an SVAR with long-run restrictions and found that labor supply shocks were the dominant driver of business cycles. In estimated DSGE models with many shocks, wage markup shocks are often found to play an important role. This is particularly the case for *news* about wage markups. For example, both Khan and Tsoukalas (2012) and Schmitt-Grohe and Uribe (2012) find that wage markup news shocks account for 60% of the variance share of hours. A key question is whether exogenous shocks to the labor market are an important part of fluctuations or whether we are accidentally identifying an internal propagation mechanism as an exogenous shock.

7. SUMMARY AND CONCLUSIONS

This chapter has summarized the new methods and new findings concerning macroeconomic shocks and their propagation. Identification is particularly challenging in macroeconomics because researchers ask questions for which *dynamics* are all important, *general equilibrium* effects are crucial, and *expectations* have powerful effects.

The literature has made substantial progress in thinking seriously about identification of shocks since the early days of Cholesky decompositions. It now exploits new data sources, such as narrative records, survey expectations, and high–frequency financial data, combines theory with extra data series (eg, the relative price of investment goods), and incorporates that information in estimated DSGE models and SVARs.

The introduction to this chapter posed the question: Are we destined to remain forever ignorant of the fundamental causes of economic fluctuations? I would argue "no." Although we still have far to go, substantial progress has been made since Cochrane (1994) asked that question.

In support of my answer, I offer the following forecast error variance decomposition that combines some of the leading shocks I have discussed in this chapter. I specify a VAR that contains the shocks as well as some macroeconomic variables. In particular, it contains (in this order) Ben Zeev and Pappa's (forthcoming) military news shock, Leeper et al. (2012) news about future taxes from bond prices, the Romer and Romer's (2010) unanticipated tax shocks (as constructed by Mertens and Ravn, 2012), Francis et al.'s (2014) medium-horizon restriction TFP shock, Ben Zeev and Khan's (2015)

IST news shock, and Justiniano et al.'s (2011) MEI shock. The macroeconomic variables include the logs real per capita values of GDP and total hours, as well as the log of commodity prices and the GDP deflator. Ordered last is the federal funds rate. Four lags are used and a quadratic trend is included.

Table 13 shows the forecast error variance decomposition of log real GDP per capita and log hours per capita. Because of data limitations on some of the shocks, the sample starts after the Korean War. It should not be surprising, then, that the government spending shocks are not very important. The tax news shocks contribute a small amount, less than 10%. The unanticipated tax shocks are unimportant.

Which shocks are important? The most important for both output and hours is Ben Zeev and Khan's (2015) news about IST change. This variable contributes an important part of the forecast error variance of both output and hours. For example, at 8 quarters the contribution to hours is 40%. The 90% confidence interval (not shown in the table) is (25, 54). Justiniano et al.'s (2011) MEI shock contributes 42% on impact, with a 90% confidence interval of (34, 50), but this falls to 24% by 1 year. If we associate the innovations to the federal funds rate with monetary policy shocks, then monetary policy shocks contribute up to 8% of the variance of output, but up to 18% of the variance in hours.

Table 13 Combined VAR: shock contribution to the forecast error variance of output and hours: 1954q3–2005q4

Horizon	bzp_gov	lrw	rrtaxu	ford_tfp	bzk_ist_news	jpt_mei	ffr
A. Output							
0	5.5	0.1	2.4	15.8	11.8	42.1	0.0
4	1.6	5.6	1.6	15.1	28.8	23.9	2.0
8	1.4	4.8	1.5	13.9	26.9	16.3	6.1
12	3.0	4.8	1.2	12.6	22.1	13.6	8.1
16	4.4	6.9	1.2	11.2	19.6	12.5	7.8
20	4.9	8.5	1.2	10.7	18.6	11.9	7.4
B. Hours							
0	2.3	0.8	0.3	17.6	13.2	20.5	0.0
4	0.5	6.6	0.8	3.7	38.3	22.1	3.2
8	0.9	6.3	0.9	2.4	39.5	14.2	10.9
12	4.1	5.2	0.7	1.8	33.4	11.5	16.8
16	7.3	6.0	0.7	1.7	28.6	10.6	18.3
20	8.9	7.0	0.8	2.0	26.7	10.2	18.1

Notes: These results are from a standard VAR with four lags and a quadratic trend estimated from 1954q3 to 2005q4. The variables are as follows, in this order: Bzp_gov, lrw, rrtaxu, ford_tfp, bzk_ist_news, jpt_mei, log real GDP per capita, log total hours per capita, log commodity prices, log GDP deflator, federal funds rate.
Abbreviations: *bzk*, Ben Zeev and Khan; *bzp*, Ben Zeev and Pappa; *ffr*, federal funds rate; *ford*, Francis, Owyang, Roush, DiCecio; *lrw*, Leeper, Richter, Walker anticipated future tax; *ist*, investment-specific technology; *Jpt*, Justiniano, Primiceri, Tambolotti; *mei*, marginal efficiency of invest; *rrtaxu*, Romer–Romer unanticipated tax; *tfp*, total factor productivity.

In sum, the three fiscal shocks, the three technology shocks, and the federal funds rate shock contribute 63–79% of the variances of output and hours at horizons of 4–20 quarters. While much more work should be done exploring the plausibility of the identifying assumptions, testing the robustness of these shock estimates, and making sure that they do satisfy the properties a shock should satisfy, these results suggest that we are indeed closer to understanding Slutsky's random shocks that drive macroeconomic fluctuations.

ACKNOWLEDGMENTS

REFERENCES

Alesina, A., Ardagna, S., 1998. Tales of fiscal adjustment. Econ. Policy 13 (27), 487–545.
Alesina, A., Ardagna, S., 2010. Large changes in fiscal policy: taxes versus spending. In: Brown, J.R. (Ed.), Tax Policy and the Economy, vol. 24. National Bureau of Economic Research, Cambridge, MA.
Alesina, A., Perotti, R., 1995. Fiscal expansions and fiscal adjustments in OECD countries. Econ. Policy 10 (21), 205–248.
Alesina, A., Perotti, R., 1997. Fiscal adjustments in OECD countries: composition and macroeconomic effects. IMF Staff. Pap. 44, 210–248.
Alexopoulos, M., 2011. Read all about it!! What happens following a technology shock? Am. Econ. Rev. 101 (4), 1144–1179.
Amir Ahmadi, P., Uhlig, H., 2015. Sign Restrictions in Bayesian FAVARs with an Application to Monetary Policy Shocks. NBER Working Paper 21738.
Angrist, J.D., Jordà, O., Kuersteiner, G., 2013. Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited. NBER Working Paper 19355.
Arias, J.E., Caldara, D., Rubio-Ramirez, J.F., 2015a. The Systematic Component of Monetary Policy in SVARS: An Agnostic Identification Procedure. January 2015 Working Paper.
Arias, J.E., Rubio-Ramirez, J.F., Waggoner, D.F., 2015b. Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications. November 2015 Working Paper.
Auerbach, A., Gorodnichenko, Y., 2012. Measuring the output responses to fiscal policy. Am. Econ. J. Econ. Pol. 4 (2), 1–27.
Auerbach, A., Gorodnichenko, Y., 2013. Fiscal multipliers in recession and expansion. In: Alesina, A., Giavazzi, F. (Eds.), Fiscal Policy After the Financial Crisis. University of Chicago Press, Chicago, IL.
Bagliano, F.C., Favero, C.A., 1999. Information from financial markets and VAR measures of monetary policy. Eur. Econ. Rev. 43 (4–6), 825–837.
Barakchian, S.M., Crowe, C., 2013. Monetary policy matters: evidence from new shocks. J. Monet. Econ. 60 (8), 950–966.
Barro, R.J., 1977. Unanticipated money growth and unemployment in the United States. Am. Econ. Rev. 67 (2), 101–115.

Barro, R.J., 1978. Unanticipated money, output, and the price level in the United States. J. Polit. Econ. 86 (4), 549–580.

Barro, R.J., 1981. Output effects of government purchases. J. Polit. Econ. 89 (6), 1086–1121.

Barro, R.J., Redlick, C.J., 2011. Macroeconomic effects from government purchases and taxes. Q. J. Econ. 126 (1), 51–102.

Barsky, R.B., Sims, E.R., 2011. News shocks and business cycles. J. Monet. Econ. 58 (3), 273–289.

Barsky, R.B., Sims, E., 2012. Information, animal spirits, and the meaning of innovations in consumer confidence. Am. Econ. Rev. 102 (4), 1343–1377.

Barsky, R.B., Basu, S., Lee, K., 2014. Whither News Shocks? NBER Working Paper 20666.

Barth III, M.J., Ramey, V.A., 2002. The cost channel of monetary transmission. In: Bernanke, B.S., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2001, vol. 16. MIT Press, Cambridge, MA, pp. 199–256.

Basu, S., Kimball, M.S., 1997. Cyclical productivity with unobserved input variation. No. w5915, National Bureau of Economic Research.

Basu, S., Fernald, J.G., Kimball, M.S., 2006. Are technology improvements contractionary? Am. Econ. Rev. 96, 1418–1448.

Baumeister, C., Hamilton, J.D., 2015. Sign restrictions, structural vector autoregressions, and useful prior information. Econometrica 83, 1963–1999.

Beaudry, P., Lucke, B., 2010. Letting different views about business cycles compete. In: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), NBER Macroeconomics Annual 2009, vol. 24. University of Chicago Press, Chicago, IL, pp. 413–455.

Beaudry, P., Portier, F., 2006. Stock prices, news, and economic fluctuations. Am. Econ. Rev. 96 (4), 1293–1307.

Beaudry, P., Portier, F., 2014. News driven business cycles: insights and challenges. J. Econ. Lit. 52 (4), 993–1074.

Beaudry, P., Fève, P., Guay, A., 2015. When is Nonfundamentalness in VARs a Real Problem? An Application to News Shocks. No. W21466, National Bureau of Economic Research.

Ben Zeev, N., Khan, H., 2015. Investment-specific news shocks and U.S. business cycles. J. Money Credit Bank. 47, 1443–1464.

Ben Zeev, N., Pappa, E., forthcoming. Chronicle of a War Foretold: The Macroeconomic Effects of Anticipated Defense Spending Shocks. Econ. J. http://doi.org/10.1111/ecoj.12349.

Bernanke, B.S., 1986. Alternative explanations of the money–income correlation. Carn.-Roch. Conf. Ser. Public Policy 25, 49–99.

Bernanke, B.S., Blinder, A.S., 1992. The Federal funds rate and the channels of monetary transmission. Am. Econ. Rev. 82 (4), 901–921.

Bernanke, B.S., Mihov, I., 1998. Measuring monetary policy. Q. J. Econ. 113 (3), 869–902.

Bernanke, B.S., Boivin, J., Eliasz, P., 2005. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. Q. J. Econ. 120 (1), 387–422.

Bernstein, J., Romer, C.D., 2009. The job impact of the American Recovery and Reinvestment Plan. Office of the Vice President-Elect. Working paper.

Blanchard, O., Perotti, R., 2002. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. Q. J. Econ. 117, 1329–1368.

Blanchard, O., Quah, D., 1989. The dynamic effects of aggregate demand and supply disturbances. Am. Econ. Rev. 79 (4), 655–673.

Blanchard, O., Watson, M.W., 1986. Are all business cycles alike? In: Gordon, R.J. (Ed.), The American Business Cycle: Continuity and Change. NBER. The University of Chicago Press, Chicago, IL.

Blanchard, O.J., L'Huillier, J.P., Lorenzoni, G., 2013. News, noise, and fluctuations: an empirical exploration. Am. Econ. Rev. 103 (7), 3045–3070.

Boivin, J., 2006. Has U.S. monetary policy changed? Evidence from drifting coefficients and real–time data. J. Money Credit Bank. 38 (5), 1149–1173.

Boivin, J., Giannoni, M.P., 2006. Has monetary policy become more effective? Rev. Econ. Stat. 88 (3), 445–462.

Boivin, J., Kiley, M.T., Mishkin, F.S., 2010. How has the monetary transmission mechanism evolved over time? In: Friedman, B.M., Woodford, M. (Eds.), Handbook of Monetary Economics. Elsevier.

Boschen, J.F., Mills, L.O., 1995. The relation between narrative and money market indicators of monetary policy. Econ. Inq. 33 (1), 24–44.

Burnside, C., Eichenbaum, M., Rebelo, S., 1995. Capital utilization and returns to scale. In: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual 1995, vol. 10. MIT Press, Cambridge, MA, pp. 67–124.

Burnside, C., Eichenbaum, M., Fisher, J., 2004. Fiscal shocks and their consequences. J. Econ. Theory 115, 89–117.

Caldara, D., Kamps, C., 2012. The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, DC.

Campbell, J.R., Evans, C.L., Fisher, J.D.M., Justiniano, A., 2012. Macroeconomic effects of federal reserve forward guidance. Brook. Pap. Econ. Act. (Spring), 1–80.

Canova, F., De Nicolo, G., 2002. Monetary disturbances matter for business fluctuations in the G-7. J. Monet. Econ. 49 (6), 1131–1159.

Canova, F., Pina, J.P., 2005. What VAR tell us about DSGE models? In: Diebolt, C., Kyrtsou, C. (Eds.), New Trends in Macroeconomics. Springer, Berlin Heidelberg, pp. 89–123.

Canova, F., Sala, L., 2009. Back to square one: identification issues in DSGE models. J. Monet. Econ. 56, 431–449.

Chang, P.-L., Sakata, S., 2007. Estimation of impulse response functions using long autoregression. Econ. J. 10 (2), 453–469.

Chari, V.V., Kehoe, P.J., McGrattan, E.R., 2008. Are structural VARs with long-run restrictions useful in developing business cycle theory? J. Monet. Econ. 55 (8), 1337–1352.

Christiano, L.J., Eichenbaum, M., 1992. Liquidity effects and the monetary transmission mechanism. Am. Econ. Rev. 82 (2), 346.

Christiano, L.J., Eichenbaum, M., Evans, C.L., 1999. What have we learned and to what end? In: Woodford, M., Taylor, J.D. (Eds.), Handbook of Macroeconomics. Elsevier, Amsterdam.

Christiano, L.J., Eichenbaum, M., Vigfusson, R., 2003. What happens after a technology shock? In: NBER Working Paper Series 9819. National Bureau of Economic Research, Cambridge, MA.

Christiano, L.J., Eichenbaum, M., Evans, C.L., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. J. Polit. Econ. 113 (1), 1–45.

Clemens, J., Miran, S., 2012. Fiscal policy multipliers on subnational government spending. Am. Econ. J.: Econ. Pol. 4 (2), 46–68.

Cloyne, J., 2013. Discretionary tax changes and the macroeconomy: new narrative evidence from the United Kingdom. Am. Econ. Rev. 103 (4), 1507–1528.

Cochrane, J., 1994. Shocks. Carn.-Roch. Conf. Ser. Public Policy 41, 295–364.

Cochrane, J., Comments on 'A New Measure of Monetary Shocks: Derivation and Implications' By Christina Romer and David Romer. July 17, 2004, presented at NBER EFG Meeting. http://faculty.chicagobooth.edu/john.cochrane/research/papers/talk_notes_new_measure_2.pdf.

Cochrane, J., Piazzesi, M., 2002. The fed and interest rates—a high-frequency identification. Am. Econ. Rev. 92 (2), 90–95.

Cogan, J.F., Cwik, T., Taylor, J.B., Wieland, V., 2010. New Keynesian versus old Keynesian government spending multipliers. J. Econ. Dyn. Control. 34 (3), 281–295.

Coibion, O., 2012. Are the effects of monetary policy shocks big or small? Am. Econ. J. Macroecon. 4 (2), 1–32.

Coibion, O., Gorodnichenko, Y., 2011. Monetary policy, trend inflation, and the great moderation: an alternative interpretation. Am. Econ. Rev. 101 (1), 341–370.

Cover, J.P., 1992. Asymmetric effects of positive and negative money-supply shocks. Q. J. Econ. 107 (4), 1261–1282.

D'Amico, S., King, T.B., 2015. What Does Anticipated Monetary Policy Do? Federal Reserve Bank of Chicago Working Paper 2015–10, November 2015.

Davis, S.J., Haltiwanger, J., 2001. Sectoral job creation and destruction responses to oil price changes. J. Monet. Econ. 48 (3), 465–512.

Edelberg, W., Eichenbaum, M.S., Fisher, J.D.M., 1999. Understanding the effects of a shock to government purchases. Rev. Econ. Dynamics. 2 (1), 166–206.

Eichenbaum, M.S., 1992. Comment on interpreting the macroeconomic time series facts: the effects of monetary policy. Eur. Econ. Rev. 36 (5), 1001–1011.

Elliott, G., 1998. On the robustness of cointegration methods when regressors almost have unit roots. Econometrica 66 (1), 149–158.

Erceg, C.J., Guerrieri, L., Gust, C., 2005. Can long-run restrictions identify technology shocks? J. Eur. Econ. Assoc. 3 (6), 1237–1278.

Evans, M.K., 1969. Reconstruction and estimation of the balanced budget multiplier. R. Econ. Stats. 51 (1), 14–25.

Evans, C.L., 1992. Productivity shocks and real business cycles. J. Monet. Econ. 29 (2), 191–208.

Farhi, E., Werning, I., 2012. Fiscal Multipliers: Liquidity Traps and Currency Unions. NBER Working Paper No. 18381, September 2012.

Fatás, A., Mihov, I., 2001. The effects of fiscal policy on consumption and employment: theory and evidence. CEPR Discussion Paper No. 2760.

Faust, J., 1998. The robustness of identified VAR conclusions about money. Carn.-Roch. Conf. Ser. Public Policy 49, 207–244.

Faust, J., Leeper, E.M., 1997. When do long-run identifying restrictions give reliable results? J. Bus. Econ. Stat. 15 (3), 345–353.

Faust, J., Swanson, E.T., Wright, J.H., 2004. Identifying VARS based on high frequency futures data. J. Monet. Econ. 51 (6), 1107–1113.

Favero, C., Giavazzi, F., 2012. Measuring tax multipliers: the narrative method in fiscal VARs. Am. Econ. J. Econ. Pol. 4 (2), 69–94.

Fernald, J.G., 1999. Roads to prosperity? Assessing the link between public capital and productivity. Am. Econ. Rev. 89 (3), 619–638.

Fernald, J.G., 2007. Trend breaks, long-run restrictions, and contractionary technology improvements. J. Monet. Econ. 54 (8), 2467–2485.

Fernald, J.G., 2014. A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. Federal Reserve Bank of San Francisco Working Paper 2012–19, April 2014.

Fernandez-Villaverde, J., Rubio-Ramirez, J., Sargent, T.J., Watson, M.W., 2007. A, B, C's (and D's) of understanding VARs. Am. Econ. Rev. 97 (3), 1021–1026.

Fisher, J.D.M., 2006. The dynamic effects of neutral and investment-specific technology shocks. J. Polit. Econ. 114 (3), 413–451.

Fisher, J.D.M., 2010. Comment on 'Letting Different Views of the Business Cycle Compete'. In: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), NBER Macroeconomics Annual 2009, vol. 24, University of Chicago Press, Chicago, IL, pp. 457–474.

Fisher, J.D.M., Peters, R., 2010. Using stock returns to identify government spending shocks. Econ. J. 120, 414–436.

Forni, M., Gambetti, L., Sala, L., 2014. No news in business cycles. Econ. J. 124, 1168–1191.

Francis, N., Ramey, V.A., 2006. The source of historical fluctuations: an analysis using long-run restrictions. In: Clarida, R., Frankel, J., Giavazzi, F., West, K. (Eds.), NBER International Seminar on Macroeconomics 2004. The MIT Press, Cambridge, MA, pp. 17–49.

Francis, N., Ramey, V.A., 2005. Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited. J. Monet. Econ. 52 (8), 1379–1399.

Francis, N., Ramey, V.A., 2009. Measures of per capita hours and their implications for the technology-hours debate. J. Money Credit Bank. 41 (6), 1071–1097.

Francis, N., Owyang, M.T., Rousch, J.E., DiCecio, R., 2014. A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks. Rev. Econ. Stat. 96 (4), 638–647.

Friedman, M., Schwartz, A., 1963. A Monetary History of the United States: 1867–1960. National Bureau of Economic Research, Princeton University Press, Princeton, NJ.

Frisch, R., 1933. Propagation problems and impulse problems in dynamic economics. In: Economic Essays in Honor of Gustav Cassel. Allen & Unwin, London, pp. 171–205.

Galí, J., 1999. Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations. Am. Econ. Rev. 89, 249–271.

Galí, J., David López-Salido, J., Vallés, J., 2007. Understanding the effects of government spending on consumption. J. Eur. Econ. Assoc. 5 (1), 227–270.

Gechert, S., 2015. What Fiscal Policy Is Most Effective? A Meta Regression Analysis. Oxford Economic Papers. 67 (3), 553–580.

Gertler, M., Karadi, P., 2015. Monetary policy surprises, credit costs, and economic activity. Am. Econ. J. Macroecon. 7 (1), 44–76.

Giavazzi, F., Pagano, M., 1990. Can severe fiscal consolidations be expansionary? Tales of two small European countries. In: Blanchard, O., Fischer, S. (Eds.), NBER Macroeconomics Annual, Vol. 5. National Bureau of Economic Research, Cambridge, Massachusetts.

Giavazzi, F., Pagano, M., 1996. Non-Keynesian effects of fiscal policy changes: international evidence and the Swedish experience. Swedish Econ. Pol. Rev. 3 (1), 67–103.

Gilchrist, S., Zakrajšek, E., 2012. Credit spreads and business cycle fluctuations. Am. Econ. Rev. 102 (4), 1692–1720.

Gordon, R.J., Krenn, R., 2010. The End of the Great Depression: VAR Insight on the Roles of Monetary and Fiscal Policy. NBER Working Paper 16380, September.

Gospodinov, N., María Herrera, A., Pesavento, E., 2013. Unit roots, cointegration, and pretesting in VAR models. In: Fomby, T.B., Kilian, L., Murphy, A. (Eds.), VAR Models in Macroeconomics—New Developments and Applications: Essays in Honor of Christopher A. Sims. Emerald Group Publishing Limited, pp. 81–115.

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37 (3), 424–438.

Greenwood, J., Jovanovic, B., 1999. The information-technology revolution and the stock market. Am. Econ. Rev. 89 (2), 116–122.

Greenwood, J., Hercowitz, Z., Huffman, G.W., 1988. Investment, capacity utilization, and the real business cycle. Am. Econ. Rev. 78 (3), 402–417.

Greenwood, J., Hercowitz, Z., Krusell, P., 2000. The role of investment-specific technological change in the business cycle. Eur. Econ. Rev. 44 (1), 91–115.

Gürkaynak, R.S., Sack, B., Swanson, E., 2005. The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models. Am. Econ. Rev. 95 (1), 425–436.

Hall, R.E., 1980. Labor supply and aggregate fluctuations. In: Carnegie-Rochester Conference Series on Public Policy, vol. 12, North-Holland.

Hall, R.E., 1986. The role of consumption in economic fluctuations. In: Gordon, R.J. (Ed.), The American Business Cycle: Continuity and Change. NBER, University of Chicago Press, Chicago, IL, pp. 237–266.

Hall, R.E., 1988. The relation between price and marginal cost in U.S. industry. J. Polit. Econ. 96 (5), 921–947.

Hall, R.E., 1990. Invariance properties of Solow's productivity residual. In: Diamond, P. (Ed.), Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow's Birthday. MIT Press, Cambridge, MA, pp. 71–112.

Hall, R.E., 2009. By how much does GDP rise if the government buys more output? Brook. Pap. Econ. Act. 2, 183–231.

Hamilton, J.D., 1983. Oil and the macroeconomy since World War II. J. Polit. Econ. 91 (2), 228–248.

Hamilton, J.D., 1985. Historical causes of postwar oil shocks and recessions. Energy J. 6 (1), 97–116.

Hamilton, J.D., 2003. What is an oil shock? J. Econ. 113 (2), 363–398.

Hamilton, J.D., 2010. Macroeconomics and ARCH. In: Bollerslev, T., Russell, J., Watson, M. (Eds.), Volatility and Time Series Econometrics: Essays in Honor of Robert Engle. Oxford University Press, Oxford, pp. 79–96.

Hansen, L.P., Sargent, T.J., 1991. Two difficulties in interpreting vector autoregressions. In: Hansen, L.P., Sargent, T.J. (Eds.), Rational Expectations Econometrics. Westview Press, Boulder, CO, pp. 77–119.

Hanson, S.G., Stein, J.C., 2015. Monetary policy and long-term real rates. J. Financ. Econ. 115 (3), 429–448.

Hausman, J.K., 2016. Fiscal policy and economic recovery: the case of the 1936 Veterans' bonus. Am. Econ. Rev. 106 (4), 1100–1143.

Hobijn, B., Jovanovic, B., 2001. The information-technology revolution and the stock market: evidence. Am. Econ. Rev. 91 (5), 1203–1220.

Hoover, K.D., Perez, S.J., 1994. Post hoc ergo propter once more an evaluation of 'does monetary policy matter?' In the spirit of James Tobin. J. Monet. Econ. 34 (1), 47–74.

Jordà, Ò., 2005. Estimation and inference of impulse responses by local projections. Am. Econ. Rev. 95 (1), 161–182.

Justiniano, A., Primiceri, G.E., Tambalotti, A., 2010. Investment shocks and business cycles. J. Monet. Econ. 57 (2), 132–145.

Justiniano, A., Primiceri, G.E., Tambalotti, A., 2011. Investment shocks and the relative price of investment. Rev. Econ. Dyn. 14 (1), 102–121.

Keynes, J.M., 1936. The General Theory of Employment, Interest and Money. Macmillan, London.

Khan, H., Tsoukalas, J., 2012. The quantitative importance of news shocks in estimated DSGE models. J. Money Credit Bank. 44 (8), 1535–1561.

Kilian, L., 2009. Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. Am. Econ. Rev. 99 (3), 1053–1069.

Kilian, L., Vigfusson, R.J., 2011. Are the responses of the US economy asymmetric in energy price increases and decreases? Quant. Econ. 2 (3), 419–453.

King, R., Plosser, C., Stock, J., Watson, M.W., 1991. Stochastic trends and economic fluctuations. Am. Econ. Rev. 81 (4), 819–840.

Kliem, M., Kriwoluzky, A., 2013. Reconciling narrative monetary policy disturbances with structural VAR model shocks? Econ. Lett. 121 (2), 247–251.

Komunjer, I., Ng, S., 2011. Dynamic identification of DSGE models. Econometrica 79 (6), 1995–2032.

Koop, G., Hashem Pesaran, M., Potter, S.M., 1996. Impulse response analysis in nonlinear multivariate models. J. Econ. 74 (1), 119–147.

Krishnamurthy, A., Vissing-Jorgensen, A., 2011. The effects of quantitative easing on interest rates. Brook. Pap. Econ. Acti. Fall, 215–287.

Kurmann, A., Otrok, C., 2013. News shocks and the slope of the term structure of interest rates. Am. Econ. Rev. 103 (6), 2612–2632.

Kurmann, A., Mertens, E., 2014. Stock prices, news, and economic fluctuations: comment. Am. Econ. Rev. 104 (4), 1439–1445.

Kuttner, K.N., 2001. Monetary policy surprises and interest rates: evidence from the Fed funds futures market. J. Monet. Econ. 47 (3), 523–544.

Kydland, F.E., Prescott, E.C., 1982. Time to build and aggregate fluctuations. Econometrica 50 (6), 1345–1370.

Lakdawala, A., 2015. Changes in Federal Reserve Preferences. Michigan State University Working Paper, April 2015.

Leduc, Sylvain, Wilson, Daniel, 2013. Roads to prosperity or bridges to nowhere? Theory and evidence on the impact of public infrastructure investment. In: Acemoglu, D., Parker, J., Woodford, M. (Eds.), NBER Macroecon. Annu. 2012. Vol. 27, pp. 89–142.

Leeper, E.M., 1997. Narrative and VAR approaches to monetary policy: common identification problems. J. Monet. Econ. 40, 641–657.

Leeper, E.M., Richter, A., Walker, T.B., 2012. Quantitative effects of fiscal foresight. Am. Econ. J. Econ. Pol. 4 (2), 1–27.

Leeper, E.M., Walker, T.B., Susan Yang, S.-C., 2013. Fiscal foresight and information flows. Econometrica 81 (3), 1115–1145. Also, unpublished supplement at: https://www.econometricsociety.org/sites/default/files/8337_extensions_0.pdf.

Leigh, D., Devries, P., Freedman, C., Guajardo, J., Laxton, D., Pescatori, A., 2010. Will it hurt? Macroeconomic effects of fiscal consolidation. World Economic Outlook, IMF, October 2010 (Chapter 3).

Litterman, R.B., Weiss, L., 1985. Money, real interest rates, and output: a reinterpretation of post-war data. Econometrica 53 (1), 129–156.

Lundsford, K.G., 2015. Identifying Structural VARs with a Proxy Variable and a Test for a Weak Proxy. Federal Reserve Bank of Cleveland Working Paper, December 15–28, 2015.

Marcellino, M., Stock, J.H., Watson, M.W., 2006. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. J. Econ. 135 (1), 499–526.

McCallum, B.T., 1983. A reconsideration of Sims' evidence concerning monetarism. Econ. Lett. 13 (2–3), 167–171.

McGrattan, E.R., 1994. The macroeconomic effects of distortionary taxation. J. Monet. Econ. 33 (3), 573–601.

Mertens, K., Ravn, M.O., 2011a. Technology-hours redux: tax changes and the measurement of technology shocks. In: NBER International Seminar on Macroeconomics 2010.

Mertens, K., Ravn, M.O., 2011b. Understanding the aggregate effects of anticipated and unanticipated tax policy shocks. Rev. Econ. Dyn. 14 (1), 27–54.

Mertens, K., Ravn, M.O., 2012. Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. Am. Econ. J. Econ. Pol. 4 (2), 145–181.

Mertens, K., Ravn, M.O., 2013. The dynamic effects of personal and corporate income tax changes in the United States. Am. Econ. Rev. 103 (4), 1212–1247.

Mertens, K., Ravn, M.O., 2014. A reconciliation of SVAR and narrative estimates of tax multipliers. J. Monet. Econ. 68, S1–S19.

Miyamoto, W., Nguyen, T.L., 2015. News Shocks and Business Cycles: Evidence from Forecast Data. Santa Clara University working paper.

Montiel Olea, J.L., Pflueger, J., 2013. A robust test for weak instruments. J. Bus. Econ. Stat. 31 (3), 358–369.

Montiel Olea, J.L., Stock, J.H., Watson, M.W., 2015. Uniform Inference in SVARs with External Instruments. December 2015 manuscript.

Mountford, A., Uhlig, H., 2009. What are the effects of fiscal policy shocks? J. Appl. Econ. 24, 960–992.

Nakamura, E., Steinsson, J., 2014. Fiscal stimulus in a monetary union: evidence from US regions. Am. Econ. Rev. 104 (3), 753–792.

Nakamura, E., Steinsson, J., 2015. High Frequency Identification of Monetary Non-Neutrality. October 2015 Working Paper.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.

Oh, H., Reis, R., 2012. Targeted transfers and the fiscal response to the great recession. J. Monet. Econ. 59, S50–S64.

Olivei, G., Tenreyro, S., 2007. The timing of monetary policy shocks. Am. Econ. Rev. 97, 636–663.

Owyang, M., Ramey, G., 2004. Regime switching and monetary policy measurement. J. Monet. Econ. 51 (8), 1577–1597.

Owyang, M.T., Ramey, V.A., Zubairy, S., 2013. Are government spending multipliers greater during periods of slack? Evidence from twentieth-century historical data. Am. Econ. Rev. 103 (3), 129–134.

Pappa, E., 2009. The effects of fiscal shocks on employment and the real wage. Int. Econ. Rev. 50 (1), 217–244.

Perotti, R., 2005. Estimating the effects of fiscal policy in OECD countries. CEPR Discussion Paper 4842, January.

Perotti, R., 2011. Expectations and Fiscal Policy: An Empirical Investigation. Bocconi Working Paper.

Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57 (6), 1361–1401.

Pflueger, C.E., Wang, S., 2015. A robust test for weak instruments in Stata. Stata J. 15 (1), 216–225.

Piazzesi, M., Swanson, E.T., 2008. Futures prices as risk-adjusted forecasts of monetary policy. J. Monet. Econ. 55 (4), 677–691.

Pigou, A.C., 1927. Industrial Fluctuations. Macmillan, London.

Poterba, J.M., 1986. Explaining the yield spread between taxable and tax-exempt bonds: the role of expected tax policy. In: Rosen, H.S. (Ed.), Studies in State and Local Public Finance. University of Chicago Press, pp. 5–52.

Prescott, E.C., 1986. Theory ahead of business-cycle measurement. Carn.-Roch. Conf. Ser. Public Policy 25, 11–44.

Primiceri, G., 2005. Time varying structural vector autoregressions and monetary policy. Rev. Econ. Stud. 72 (3), 821–852.

Ramey, V.A., 2009. Identifying Government Spending Shocks: It's All in the Timing. NBER Working Paper No. 15464, October 2009.

Ramey, V.A., 2011a. Identifying government spending shocks: it's all in the timing. Q. J. Econ. 126, 1–50.

Ramey, V.A., 2011b. Can government purchases stimulate the economy? J. Econ. Lit. 49 (3), 673–685.

Ramey, V.A., 2013. Government spending and private activity. In: Alesina, A., Giavazzi, F. (Eds.), Fiscal Policy After the Financial Crisis. University of Chicago Press, Chicago, IL.

Ramey, V.A., Shapiro, M., 1998. Costly capital reallocation and the effects of government spending. Carn.-Roch. Conf. Ser. Public Policy 48, 145–194.

Ramey, V.A., Vine, D.J., 2011. Oil, automobiles, and the US economy: how much have things really changed? In: Acemoglu, D., Woodford, M. (Eds.), NBER Macroeconomics Annual 2010, Vol. 25. University of Chicago Press, pp. 333–367.

Ramey, V.A., Zubairy, S., 2014. Government Spending Multipliers in Good Times and in Bad: Evidence from 20th Century Historical Data. November 2014 Working Paper.

Romer, C.D., Romer, D.H., 1989. Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. NBER Macroeconomic Annual 1989.

Romer, C.D., Romer, D.H., 1997. Identification and the narrative approach: a reply to Leeper. J. Monet. Econ. 40, 659–665.

Romer, C.D., Romer, D.H., 2000. Federal reserve information and the behavior of interest rates. Am. Econ. Rev. 90 (3), 429–457.

Romer, C.D., Romer, D.H., 2004. A new measure of monetary policy shocks: derivation and implications. Am. Econ. Rev. 94 (4), 1055–1084.

Romer, C.D., Romer, D.H., 2010. The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. Am. Econ. Rev. 100, 763–801.

Romer, C.D., Romer, D.H., 2016. Transfer payments and the macroeconomy: the effects of social security benefit increases, 1952-1991. Berkeley working paper, March.

Rotemberg, J., Woodford, M., 1992. Oligopolistic pricing and the effects of aggregate demand on economic activity. J. Polit. Econ. 100, 1153–1297.

Rudebusch, G., 1998. Do measures of monetary policy in a VAR make sense? In: Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance. International Economic Review, vol. 39(4), pp. 907–931.

Schmitt-Grohe, S., Uribe, M., 2012. What's news in business cycles? Econometrica 80 (6), 2733–2764.

Shapiro, M.D., 1993. Cyclical productivity and the workweek of capital. Am. Econ. Rev. 83 (2), 229–233.

Shapiro, M.D., 1994. Federal reserve policy: cause and effect. In: Gregory Mankiw, N. (Ed.), Monetary Policy. National Bureau of Economic Research, The University of Chicago Press, Chicago, IL.

Shapiro, M.D., Watson, M.W., 1988. The sources of business cycle fluctuations. In: Fischer, S. (Ed.), NBER Macroeconomics Annual, vol. 3. MIT Press, Cambridge, MA.

Sims, C.A., 1972. Money, income, and causality. Am. Econ. Rev. 62 (4), 540–552.

Sims, C.A., 1980a. Macroeconomics and reality. Econometrica 48, 1–48.

Sims, C.A., 1980b. Comparison of interwar and postwar business cycles: monetarism reconsidered. Am. Econ. Rev. 70 (2), 250–257.

Sims, C.A., 1992. Interpreting the macroeconomic time series facts: the effects of monetary policy. Eur. Econ. Rev. 36 (5), 975–1000.

Sims, C.A., 1998. Discussion of Glenn Rudebusch, "Do measures of monetary policy in a VAR make sense?" In: Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance (Nov. 1998). International Economic Review, vol. 39, pp. 907–931.

Sims, C.A., Zha, T., 2006a. Were there Regime Switches in U.S. monetary policy? Am. Econ. Rev. 96 (1), 54–81.

Sims, C.A., Zha, T., 2006b. Does monetary policy generate recessions? Macroecon. Dyn. 10 (2), 231–272.

Sims, C.A., Stock, J.H., Watson, M.W., 1990. Inference in linear time series models with some unit roots. Econometrica 58 (1), 113–144.

Slutsky, E., 1937. The summation of random causes as the source of cyclic processes. Econometrica 5 (2), 105–146.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium model of the Euro area. J. Eur. Econ. Assoc. 1 (5), 1123–1175.

Smets, F., Wouters, R., 2007. Shocks and frictions in U.S. business cycles: a Bayesian DSGE approach. Am. Econ. Rev. 97 (3), 586–606.

Solow, R.M., 1957. Technical change and the aggregate production function. Rev. Econ. Stat. 39 (3), 312–320.

Stock, J.H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. J. Am. Stat. Assoc. 97 (460), 1167–1179.

Stock, J.H., Watson, M.W., 2008. NBER Summer Institute Minicourse 2008: What's New in Econometrics—Time Series, Lecture 7: Structural VARs. National Institute for Economic Research, Cambridge, MA. www.nber.org/minicourse_2008.html.

Stock, J.H., Watson, M.W., 2012. Disentangling the channels of the 2007–09 recession. Brook. Pap. Econ. Act. 2012 (Spring), 81–135.

Stock, J.H., Watson, M.W., 2016. Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In: Taylor, J.B., Uhlig, H. (Eds.), Handbook of Macroeconomics, vol. 2A. Elsevier, Amsterdam, Netherlands, pp. 415–525.

Stock, J.H., Wright, J.H., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. J. Bus. Econ. Stat. 20 (4), 518–529.

Strongin, S., 1995. The identification of monetary policy disturbances explaining the liquidity puzzle. J. Monet. Econ. 35 (3), 463–497.

Tenreyro, S., Thwaites, G., forthcoming. Pushing on a String: US Monetary Policy is Less Powerful in Recessions. Am. Econ. J.: Macro.

Thoma, M.A., 1994. Subsample instability and asymmetries in money-income causality. J. Econ. 64 (1–2), 279–306.

Uhlig, H., 1997. What are the effects of monetary policy on output? Results from an agnostic identification procedure. Tilburg University Manuscript.

Uhlig, H., 2003. What drives GNP? Unpublished manuscript, Euro Area Business Cycle Network.

Uhlig, H., 2004. Do technology shocks lead to a fall in total hours worked? J. Eur. Econ. Assoc. 2 (2–3), 361–371.

Uhlig, H., 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. J. Monet. Econ. 52 (2), 381–419.

Uhlig, H., 2010. Some fiscal calculus. Am. Econ. Rev. 100 (2), 30–34.

Velde, F.R., 2009. Chronicle of a deflation unforetold. J. Polit. Econ. 117 (4), 591–634.

Weise, C.L., 1999. The asymmetric effects of monetary policy: a nonlinear vector autoregression approach. J. Money, Credit, Bank. 31 (1), 85–108.

Wieland, J., Yang, M.-J., 2015. Financial Dampening. September 2015 Working Paper.

Wu, J.C., Xia, F.D., 2016. Measuring the macroeconomic impact of monetary policy at the zero lower bound. J. Money, Credit, Bank. 48 (2–3), 253–291.

Yang, S.-C.S., 2005. Quantifying tax effects under policy foresight. J. Monet. Econ. 52 (8), 1557–1568.