Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications

Jonas E. Arias
Federal Reserve Board

Juan F. Rubio-Ramírez*
Duke University, BBVA Research,
and Federal Reserve Bank of Atlanta

Daniel F. Waggoner
Federal Reserve Bank of Atlanta

August 29, 2014

Abstract

Are optimism shocks an important source of business cycle fluctuations? Are deficit-financed tax cuts better than deficit-financed spending to increase output? These questions have been previously studied using SVARs identified with sign and zero restrictions and the answers have been positive and definite in both cases. We show that these answers are wrong. While sign and zero restrictions are attractive because it allow the researcher to remain agnostic with respect to the responses of the key variables of interest, current implementations do not respect the agnosticism of the theory because they impose additional sign restrictions on variables that are seemingly unrestricted that bias the results and produce misleading confidence intervals. We provide an alternative algorithm that does not introduce any additional sign restriction, hence preserving the agnosticism of the theory. Without the additional restrictions, it is hard to claim that either optimism shocks are an important source of business cycle fluctuations or deficit-financed tax cuts work best at improving output.

*Corresponding author: Juan F. Rubio-Ramírez <juan.rubio-ramirez@duke.edu>, Economics Department, Duke University, Durham, NC 27708; 1-919-660-1865. We thank Paul Beaudry, Andrew Mountford, Deokwoo Nam, and Jian Wang for sharing supplementary material with us, and for helpful comments. We also thank Grátula Bedátula for her support and help. Without her this paper would have been impossible. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Board of Governors of the Federal Reserve System. This paper has circulated under the title "Algorithm for Inference with Sign and Zero Restrictions." Juan F. Rubio-Ramírez also thanks the NSF, the Institute for Economic Analysis (IAE) and the "Programa de Excelencia en Educación e Investigación" of the Bank of Spain, and the spanish ministry of science and technology Ref. ECO2011-30323-c03-01 for support.

1 Introduction

Are optimism shocks an important source of business cycle fluctuations? Are deficit-financed tax cuts better than deficit-financed spending to increase output? Several questions such as these have been previously studied in the literature using SVARs identified by imposing sign and zero restrictions on impulse response functions and frequently the answers have been definite. For example, Beaudry, Nam and Wang (2011) conclude that optimism shocks play a pivotal role in economic fluctuations and Mountford and Uhlig (2009) conclude that deficit-financed tax cuts are better for stimulating economic activity. Researchers combine SVARs with sign and zero restrictions because they allow the identification to remain agnostic with respect to the responses of key variables of interest. But this is just in theory. In practice this has not been the case.

We show that the current implementation of these techniques does, in fact, introduce sign restrictions in addition to the ones specified in the identification – violating the proclaimed agnosticism. The additional sign restrictions generate biased impulse response functions and artificially narrow confidence intervals around them. Hence, the researcher is going to be confident about the wrong thing. The consequence is that Beaudry, Nam and Wang (2011) and Mountford and Uhlig (2009) are not as agnostic as they pretend to be and that their positive and sharp conclusions are misleading and due to these additional sign restrictions. The heart of the problem is that none of the existing algorithms correctly draws from the posterior distribution of structural parameters conditional on the sign and zero restrictions. In this paper we solve this problem by providing an algorithm that draws from the correct posterior, hence not introducing any additional sign restrictions. Absent the additional sign restrictions, it is hard to support the claim that either optimism shocks are an important source of business cycle fluctuations or deficit-financed tax cuts work best at improving output. Once you are truly agnostic, Beaudry, Nam and Wang's (2011) and Mountford and Uhlig's (2009) main findings disappear.

In particular, we present an efficient algorithm for inference in SVARs identified with sign and zero restrictions that properly draws from the posterior distribution of structural parameters conditional on the sign and zero restrictions. We extend the sign restrictions methodology developed by Rubio-Ramírez, Waggoner and Zha (2010) to allow for zero restrictions. As was the case in Rubio-Ramírez, Waggoner and Zha (2010), we obtain most of our results by imposing sign and zero restrictions on the impulse response functions, but our algorithm allows for a larger class of restrictions. Two properties of the problem are relevant: (1) the set of structural parameters conditional on the sign and zero restrictions is of positive measure in the set of structural parameters conditional on the zero restrictions and (2) the posterior distribution of structural parameters conditional on the zero restrictions can be obtained from the product of the posterior distribution of the reduced-form parameters and the uniform distribution, with respect to the Haar measure, on the set of orthogonal matrices conditional on the zero restrictions. Drawing from the posterior of the reduced-form parameters is a well-understood problem. Our key theoretical contribution is to show how to efficiently draw from the uniform distribution with respect to the Haar measure on the set of orthogonal matrices conditional on the zero restrictions. This is the crucial step that allows us to draw from the posterior distribution of structural parameters conditional on the sign and zero restrictions and that differentiates our paper from existing algorithms.

Currently, the most widely used algorithm is Mountford and Uhlig's (2009) penalty function approach (PFA). Instead of drawing from the posterior distribution of structural parameters conditional on the sign and zero restrictions, the PFA selects a single value of the structural parameters by minimizing a loss function. We show that this approach has several drawbacks that crucially affect inference. First, the PFA imposes additional sign restrictions on variables that are seemingly unrestricted − violating the proclaimed agnosticism of the identification. The additional sign restrictions bias the impulse response functions. Indeed, for a class of sign and zero restrictions we can even formally recover the additional sign restrictions. Second, because it chooses a single value of structural parameters, the PFA creates artificially narrow confidence intervals around the impulse response functions that severely affect inference and the economic interpretation of the results.

We show the capabilities of our algorithm and the problems of the PFA by means of two applications previously analyzed in the literature using the PFA. The first application is related to optimism shocks. The aim of Beaudry, Nam and Wang (2011) is to provide new evidence on the relevance of optimism shocks as an important driver of macroeconomic fluctuations. In their most basic identification scheme, the authors claim to be agnostic about the response of consumption and hours to optimism shocks. Beaudry, Nam and Wang (2011) conclude that optimism shocks are clearly important for explaining standard business cycle type phenomena because they increase consumption and hours. Unfortunately, we show that the positive and sharp responses of consumption and hours reported in Beaudry, Nam and Wang (2011) are due to the additional sign restrictions on these variables introduced by the PFA that bias impulse response functions and create artificially narrow confidence intervals around them. Since these restrictions on consumption and hours were not part of the identification scheme, the PFA contravenes the proclaimed agnosticism of the identification. Once you are truly agnostic using our methodology, Beaudry, Nam and Wang's (2011) conclusion is very hard to support.

The second application identifies fiscal policy shocks, as in Mountford and Uhlig (2009), in order to analyze the effects of these shocks on economic activity. Government revenue and government spending shocks are identified by imposing sign restrictions on the fiscal variables themselves as well as imposing orthogonality to a business cycle shock and a monetary policy shock. The identification pretends to remain agnostic with respect to the responses of output and other variables of interest to the fiscal policy shocks. Mountford and Uhlig's (2009) main finding is that deficit-financed tax cuts work best among the different fiscal policies aimed at improving output. Analogously to Beaudry, Nam and Wang (2011), the PFA introduces additional sign restrictions on the response of output and other variables to fiscal policy shocks, again conflicting with the proclaimed agnosticism of the identification strategy. As before, the results obtained in Mountford and Uhlig (2009) are due to biased impulse response functions and the artificially narrow confidence intervals around them created by the additional sign restrictions. Using our truly agnostic methodology, we show that it is very difficult to endorse Mountford and Uhlig's (2009) results.

There is some existing literature that criticizes Mountford and Uhlig's (2009) PFA using arguments similar to the ones listed here. Baumeister and Benati (2013), Benati (2013), and Binning (2013) propose alternative algorithms, failing to provide any theoretical justification that their algorithms, in fact, draw from the posterior distribution of structural parameters conditional on the sign and zero restrictions. Caldara and Kamps (2012) also share our concerns about the PFA, but providing an alternative algorithm is out of the scope of their paper. In an environment with only sign restrictions, Baumeister and Hamilton (2013) highlight some related problems and advocate for using priors that reflect not just sign restrictions but also the relative plausibility of different parameter values within the allowable set.

Not only does our method correctly draw from the posterior distribution of structural parameters conditional on the sign and zero restrictions but, at least for the two applications studied in this paper, it is also much faster than the PFA. Our methodology is between three and ten times faster than the PFA, depending on the number of sign and zero restrictions. It is also important to note that our approach can be embedded in a classical or Bayesian framework, although we follow only the latter. In addition, we wish to state that the aim of this paper is neither to dispute nor to challenge SVARs identified with sign and zero restrictions. In fact, our methodology preserves the virtues of the pure sign restriction approach developed in the work of Faust (1998), Canova and Nicoló (2002), Uhlig (2005), and Rubio-Ramírez, Waggoner and Zha (2010). Instead, our findings related to optimism and fiscal policy shocks just indicate that the respective sign and zero restrictions used by Beaudry, Nam and

Wang (2011) and Mountford and Uhlig (2009) are not enough to accurately identify these particular structural shocks. It seems that more restrictions are needed in order to identify such shocks, possibly zero restrictions. Finally, by characterizing the set of structural parameters conditional on sign and zero restrictions, our key theoretical contribution will also be of interest to the existing literature such as Faust (1998) and Barsky and Sims (2011), who identify shocks by maximizing the forecast error variance of certain variables subject to either sign or zero restrictions.

The paper is organized as follows. Section 2 shows some relevant results in the literature that we will later demonstrate to be wrong. Section 3 presents the methodology. It is here where we describe our theoretical contributions and algorithms. Section 4 offers some examples. Section 5 describes the PFA and highlights its shortcomings. Section 6 presents the first of our applications. Section 7 presents the second application. Section 8 concludes.

2 Being Confident About the Wrong Thing

Beaudry, Nam and Wang (2011) claim to provide evidence on the relevance of optimism shocks as an important driver of macroeconomic fluctuations by exploiting sign and zero restrictions using the PFA. More details about their work will be given in Section 6. At this point it suffices to say that in their most basic model, Beaudry, Nam and Wang (2011) use data on total factor productivity (TFP), stock price, consumption, the real federal funds rate, and hours worked. In a first attempt, they identify optimism shocks as positively affecting stock prices but being orthogonal to TFP at horizon zero. Hence, the identification scheme is agnostic about the response of both consumption and hours to optimism shocks. Figure 1 replicates the first block of Figure 1 in Beaudry, Nam and Wang (2011). As can be seen, both consumption and hours worked respond positively and strongly to optimism shocks. The results are also quite definite because of the narrow confidence intervals.

If right, this figure will clearly endorse the work of those who think that optimism shocks are relevant for business cycle fluctuations. But this is not the case. In Section 6 we will show that Figure 1 is wrong. The PFA introduces additional sign restrictions on consumption and hours; hence, Figure 1 does not correctly reflect the impulse response functions (IRFs) associated with the agnostic identification scheme described above. When compared with the correctly computed IRFs (as we will do in Section 6), Figure 1 reports upward-biased responses of consumption and hours worked with artificially narrow confidence intervals. In that sense, Beaudry, Nam and Wang (2011) are confident about the wrong thing.

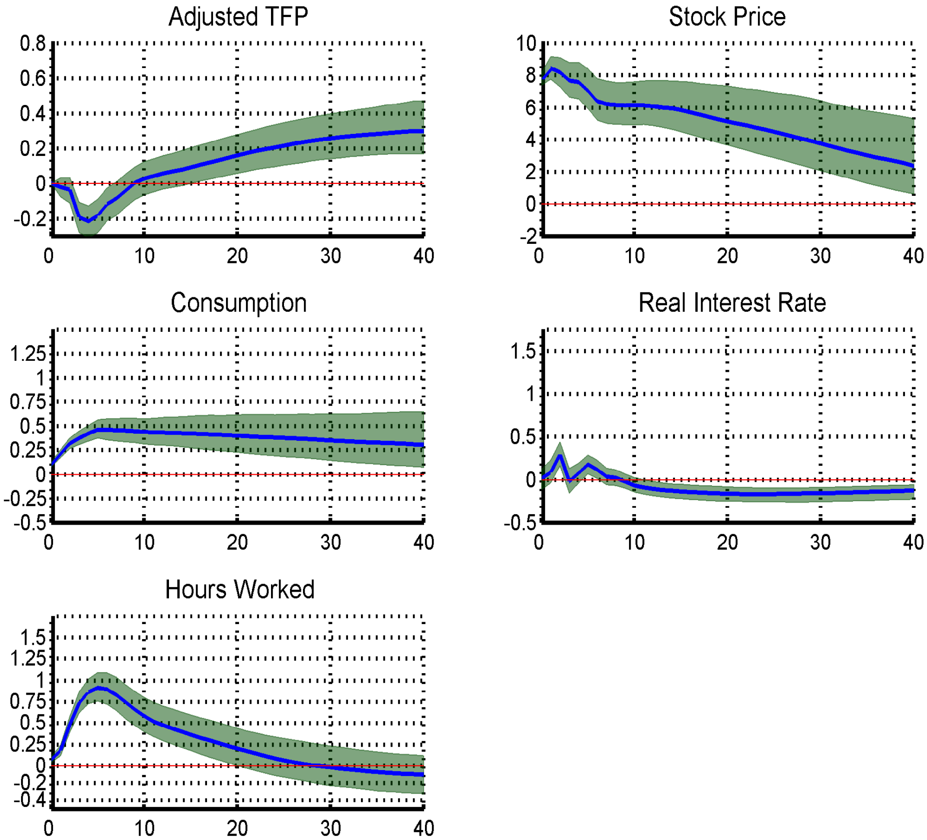


Figure 1: Beaudry, Nam and Wang (2011) Identification 1: Five-Variable SVAR

Mountford and Uhlig (2009) analyze the effects of fiscal policy shocks using SVARs identified with sign restrictions. Using data on output, consumption, total government spending, total government revenue, real wages, investment, the interest rate, adjusted reserves, prices of crude materials, and on output deflator, they identify a government revenue shock and a government spending shock by imposing sign restrictions on the fiscal variables themselves as well as imposing orthogonality to a generic business cycle shock and a monetary policy shock. No sign restrictions are imposed on the responses of output, consumption, and investment to fiscal policy shocks. Thus, the identification remains agnostic with respect to the responses of these key variables of interest to fiscal policy shocks. Using the identified fiscal policy shocks, they report many different results that will be analyzed in Section 7. At this stage we want to focus on their comparison of fiscal policy scenarios. They compare deficit-spending shocks, where total government spending rises by 1 percent and total government revenue remains unchanged during the four quarters following the initial shock, with deficit-financed tax cut shocks, where total government spending remains unchanged and total government revenue falls by 1 percent during the four quarters following the initial shock. More details about the fiscal policy scenarios will be provided in Section 7.

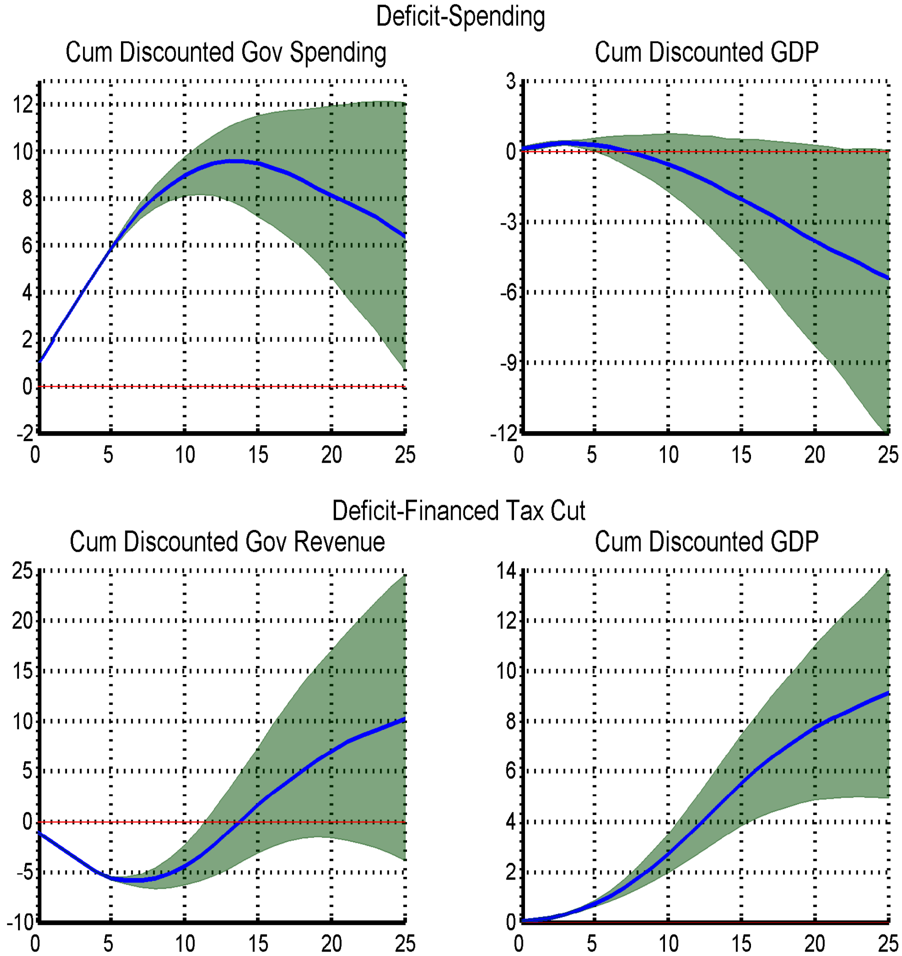


Figure 2: Mountford and Uhlig (2009) Cumulative Fiscal Multipliers

Figure 2 replicates Figure 13 in Mountford and Uhlig (2009). The figure shows that the median cumulative discounted IRF of output to a deficit-spending shock becomes negative after a few periods, while it is always positive in the case of a deficit-financed tax cut shock. It also shows narrow confidence intervals. If right, this figure will strongly support the work of those who think that deficit-financed tax cuts work best to improve output. But this is not the case. In Section 7 we will show that Figure 2 is, indeed, wrong. As is the case with optimism shocks, the PFA introduces additional sign restrictions on the response of output to the different fiscal policy shocks analyzed; hence, Figure 2 does not correctly reflect the IRFs associated with the agnostic identification scheme described above. When compared with the correctly computed IRFs (as we will do in Section 7), Figure 2 reports biased impulse response functions and artificially narrow confidence intervals. In that sense, Mountford and

Uhlig (2009) are also confident about the wrong thing.

3 Our Methodology

This section is organized into three parts. First, we describe the model. Second, we review the efficient algorithm for inference using sign restrictions on IRFs developed in Rubio-Ramírez, Waggoner and Zha (2010). Third, we extend this algorithm to also allow for zero restrictions. As mentioned, the algorithm proposed by Rubio-Ramírez, Waggoner and Zha (2010) and our extension can be embedded in a classical or Bayesian framework. In this paper we follow the latter.

3.1 The Model

Consider the structural vector autoregression (SVAR) with the general form, as in Rubio-Ramírez, Waggoner and Zha (2010)

$$\mathbf{y}_t'\mathbf{A}_0 = \sum_{\ell=1}^{p} \mathbf{y}_{t-\ell}'\mathbf{A}_\ell + \mathbf{c} + \varepsilon_t' \quad \text{for } 1 \leq t \leq T, \tag{1}$$

where $\mathbf{y}_t$ is an $n \times 1$ vector of endogenous variables, ε_t is an $n \times 1$ vector of exogenous structural shocks, $\mathbf{A}_\ell$ is an $n \times n$ matrix of parameters for $0 \leq \ell \leq p$ with $\mathbf{A}_0$ invertible, $\mathbf{c}$ is a $1 \times n$ vector of parameters, p is the lag length, and T is the sample size. The vector ε_t, conditional on past information and the initial conditions $\mathbf{y}_0, ..., \mathbf{y}_{1-p}$, is Gaussian with mean zero and covariance matrix $\mathbf{I}_n$, the $n \times n$ identity matrix. The model described in equation (1) can be written as

$$\mathbf{y}_t'\mathbf{A}_0 = \mathbf{x}_t'\mathbf{A}_+ + \varepsilon_t' \quad \text{for } 1 \leq t \leq T, \tag{2}$$

where $\mathbf{A}_+' = \begin{bmatrix} \mathbf{A}_1' & \cdots & \mathbf{A}_p' & \mathbf{c}' \end{bmatrix}$ and $\mathbf{x}_t' = \begin{bmatrix} \mathbf{y}_{t-1}' & \cdots & \mathbf{y}_{t-p}' & 1 \end{bmatrix}$ for $1 \leq t \leq T$. The dimension of $\mathbf{A}_+$ is $m \times n$, where $m = np + 1$. The reduced-form representation implied by equation (2) is

$$\mathbf{y}_t' = \mathbf{x}_t'\mathbf{B} + \mathbf{u}_t' \quad \text{for } 1 \leq t \leq T, \tag{3}$$

where $\mathbf{B} = \mathbf{A}_+\mathbf{A}_0^{-1}$, $\mathbf{u}_t' = \varepsilon_t'\mathbf{A}_0^{-1}$, and $\mathbb{E}\left[\mathbf{u}_t\mathbf{u}_t'\right] = \mathbf{\Sigma} = (\mathbf{A}_0\mathbf{A}_0')^{-1}$. The matrices $\mathbf{B}$ and $\mathbf{\Sigma}$ are the reduced-form parameters, while $\mathbf{A}_0$ and $\mathbf{A}_+$ are the structural parameters.

Most of the literature imposes restrictions on the IRFs. As we will see by the end of this section, the theorems and algorithms described in this paper allow us to consider a more general class of restrictions. In any case, we will use IRFs to motivate our theory. Thus, we now characterize them. We begin by introducing IRFs at finite horizons and then do the same at the infinite horizon. Once the IRFs are defined, we will show how to impose sign restrictions. In the finite horizon case, we have the following definition.

Definition 1. *Let $(\mathbf{A}_0, \mathbf{A}_+)$ be any value of structural parameters. The IRF of the i-th variable to the j-th structural shock at finite horizon h corresponds to the element in row i and column j of the matrix*

$$\mathbf{L}_h(\mathbf{A}_0, \mathbf{A}_+) = \left(\mathbf{A}_0^{-1}\mathbf{J}'\mathbf{F}^h\mathbf{J}\right)', \text{ where}$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{A}_1\mathbf{A}_0^{-1} & \mathbf{I}_n & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{p-1}\mathbf{A}_0^{-1} & \mathbf{0} & \cdots & \mathbf{I}_n \\ \mathbf{A}_p\mathbf{A}_0^{-1} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \text{ and } \mathbf{J} = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}.$$

In the infinite horizon case, we assume the i-th variable is in first differences.

Definition 2. *Let $(\mathbf{A}_0, \mathbf{A}_+)$ be any value of structural parameters. The IRF of the i-th variable to the j-th structural shock at the infinite horizon (sometimes called the long-run IRF) corresponds to the element in row i and column j of the matrix*

$$\mathbf{L}_\infty(\mathbf{A}_0, \mathbf{A}_+) = \left(\mathbf{A}_0' - \sum_{\ell=1}^{p}\mathbf{A}_\ell'\right)^{-1}.$$

It is important to note that $\mathbf{L}_h(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q}) = \mathbf{L}_h(\mathbf{A}_0, \mathbf{A}_+)\mathbf{Q}$ for $0 \le h \le \infty$ and $\mathbf{Q} \in O(n)$, where $O(n)$ denotes the set of all orthogonal $n \times n$ matrices.

3.2 Algorithm for Sign Restrictions

Let us assume that we want to impose sign restrictions at several horizons, both finite and infinite. It is convenient to stack the IRFs for all the relevant horizons into a single matrix of dimension $k \times n$, which we denote by $f(\mathbf{A}_0, \mathbf{A}_+)$. For example, if the sign restrictions are imposed at horizon zero and infinity, then

$$f(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} \mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_\infty(\mathbf{A}_0, \mathbf{A}_+) \end{bmatrix}, \text{ where } k = 2n \text{ in this case.}$$

Sign restrictions on those IRFs can be represented by matrices $\mathbf{S}_j$ for $1 \leq j \leq n$, where the number of columns in $\mathbf{S}_j$ is equal to the number of rows in $f(\mathbf{A}_0, \mathbf{A}_+)$. Usually, $\mathbf{S}_j$ will be a selection matrix and thus will have exactly one non-zero entry in each row, though the theory will work for arbitrary $\mathbf{S}_j$. If the rank of $\mathbf{S}_j$ is s_j, then s_j is the number of sign restrictions on the IRFs to the j-th structural shock. We only specify matrices $\mathbf{S}_j$ for the structural shocks that are identified using sign restrictions. Let us define $\mathcal{S} \subset \{1, \ldots, n\}$ to be the set of indices of structural shocks that are identified using sign restrictions. The total number of sign restrictions will be $s = \sum_{j \in \mathcal{S}} s_j$. Let $\mathbf{e}_j$ denote the j-th column of $\mathbf{I}_n$, where $\mathbf{I}_n$ is the identity matrix of dimension $n \times n$.

Definition 3. *Let $(\mathbf{A}_0, \mathbf{A}_+)$ be any value of structural parameters. These parameters satisfy the sign restrictions if and only if $\mathbf{S}_j f(\mathbf{A}_0, \mathbf{A}_+) \mathbf{e}_j > \mathbf{0}$, for $j \in \mathcal{S}$.*

From equation (2), it is easy to see that if $(\mathbf{A}_0, \mathbf{A}_+)$ is any value of structural parameters and the matrix $\mathbf{Q}$ is any element of $O(n)$, the set of orthogonal matrices, then $(\mathbf{A}_0, \mathbf{A}_+)$ and $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ are observationally equivalent. It is also well known, e.g., Geweke (1986), that a SVAR with sign restrictions is not identified, since for any $(\mathbf{A}_0, \mathbf{A}_+)$ that satisfies the sign restrictions, $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ will also satisfy the sign restrictions for all orthogonal matrices $\mathbf{Q}$ sufficiently close to the identity. Thus, the structural parameters are not identified but set identified and the set of structural parameters conditional on the sign restrictions will be an open set of positive measure in the set of all structural parameters. This suggests the following algorithm for sampling from the posterior of structural parameters conditional on the sign restrictions.

Algorithm 1.

1. *Draw $(\mathbf{A}_0, \mathbf{A}_+)$ from the unrestricted posterior.*

2. *Keep the draw if the sign restrictions are satisfied.*

3. *Return to Step 1 until the required number of draws from the posterior of structural parameters conditional on the sign restrictions has been obtained.*

By unrestricted posterior we mean the posterior distribution of all structural parameters before any identification scheme is considered. The only obstacle to implementing Algorithm 1 is an efficient technique to accomplish the first step. In the next subsection we develop a fast algorithm that produces independent draws from the unrestricted posterior.

3.3 Draws from the Unrestricted Posterior

In order to obtain independent draws from the unrestricted posterior, we will require independent draws of orthogonal matrices from the uniform distribution with respect to the Haar measure on $O(n)$.[1] Faust (1998), Canova and Nicoló (2002), Uhlig (2005), and Rubio-Ramírez, Waggoner and Zha (2010) propose algorithms to do that. However, Rubio-Ramírez, Waggoner and Zha's (2010) algorithm is the only computationally feasible one for moderately large SVAR systems (e.g., $n > 4$).[2] Rubio-Ramírez, Waggoner and Zha's (2010) results are based on the following theorem.

Theorem 1. *Let* $\mathbf{X}$ *be an* $n \times n$ *random matrix with each element having an independent standard normal distribution. Let* $\mathbf{X} = \mathbf{QR}$ *be the QR decomposition of* $\mathbf{X}$.[3] *The random matrix* $\mathbf{Q}$ *has the uniform distribution with respect to the Haar measure on* $O(n)$.

Proof. The proof follows directly from Stewart (1980). □

Theorem 1 produces independent draws of orthogonal matrices from the uniform distribution with respect to the Haar measure on $O(n)$. With this result in hand, we now develop an efficient algorithm to obtain independent draws from the unrestricted posterior using independent draws from the posterior of the reduced-form parameters. If the prior on the reduced-form is from the family of multivariate normal inverse Wishart distributions, then the posterior will be from the same family and there are efficient algorithms for obtaining independent draws from this distribution. Such priors are called conjugate. The popular Minnesota prior will be from this family. However, we need draws from the unrestricted posterior, not from the posterior of reduced-form parameters. We now show that there is a way to convert a draw from the posterior of the reduced-form parameters, together with a draw from the uniform distribution with respect to the Haar measure on $O(n)$, to a draw from the unrestricted posterior.

Let g denote the mapping from the structural parameters to the reduced-form parameters given by $g(\mathbf{A}_0, \mathbf{A}_+) = \left(\mathbf{A}_+\mathbf{A}_0^{-1}, (\mathbf{A}_0\mathbf{A}_0')^{-1}\right)$. Let h be any continuously differentiable mapping from the set of symmetric positive definite $n \times n$ matrices into the set of $n \times n$ matrices such that $h(\mathbf{X})'h(\mathbf{X}) = \mathbf{X}$. For instance, $h(\mathbf{X})$ could be the Cholesky decomposition of $\mathbf{X}$ such that $h(\mathbf{X})$ is upper triangular with positive diagonal. Alternatively, $h(\mathbf{X})$ could be the square root of $\mathbf{X}$, which is the unique symmetric

[1] The Haar measure is the unique measure on $O(n)$ that is invariant under rotations and reflections such that the measure of all of $O(n)$ is one. See Krantz and Parks (2008) for more details.

[2] See Rubio-Ramírez, Waggoner and Zha (2010) for details.

[3] With probability one the random matrix $\mathbf{X}$ will be non-singular and so the QR decomposition will be unique if the diagonal of $\mathbf{R}$ is normalized to be positive.

and positive definite matrix $\mathbf{Y}$ such that $\mathbf{YY} = \mathbf{Y}'\mathbf{Y} = \mathbf{X}$. Using h, we can define a function $\hat{h}$ from the product of the set of reduced-form parameters with the set of orthogonal matrices into the set of structural parameters by $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q}) = (h(\mathbf{\Sigma})^{-1}\mathbf{Q}, \mathbf{B}h(\mathbf{\Sigma})^{-1}\mathbf{Q})$. Note that $g(\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})) = (\mathbf{B}, \mathbf{\Sigma})$ for every matrix $\mathbf{Q} \in O(n)$. The function $\hat{h}$ will be continuously differentiable with a continuously differentiable inverse.

Given a draw from the posterior of the reduced-form parameters and a draw of orthogonal matrices from the uniform distribution with respect to the Haar measure on $O(n)$, we can use $\hat{h}$ to draw from the unrestricted posterior. At the same time, given a prior density π on the reduced-form parameters $\hat{h}$ induces a prior on the unrestricted structural parameters. The induced prior density will be

$$\hat{\pi}\left(\mathbf{A}_0, \mathbf{A}_+\right) = \pi\left(\mathbf{B}, \mathbf{\Sigma}\right)\left|det\left(\hat{h}'\left(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q}\right)\right)\right|^{-1},$$

where $(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q}) = \hat{h}^{-1}(\mathbf{A}_0, \mathbf{A}_+)$. Though we will not explicitly use it, the expression for the determinant is

$$\left|det\left(\hat{h}'\left(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q}\right)\right)\right| = 2^{-\frac{n(n+1)}{2}}\left|det\left(\mathbf{\Sigma}\right)\right|^{-\frac{2n+m+1}{2}}.$$

The following theorem formalizes the above argument

Theorem 2. *Let π be a prior density on the reduced-form parameters. If $(\mathbf{B}, \mathbf{\Sigma})$ is a draw from the reduced-form posterior and the matrix $\mathbf{Q}$ is a draw from the uniform distribution with respect to the Haar measure on $O(n)$, then $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$ is a draw from the unrestricted posterior with respect to the prior $\hat{\pi}\left(\mathbf{A}_0, \mathbf{A}_+\right) = \pi\left(\mathbf{B}, \mathbf{\Sigma}\right)\left|det\left(\hat{h}'\left(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q}\right)\right)\right|^{-1}$.*

Proof. The proof follows from the chain rule and the fact that if $(\mathbf{A}_0, \mathbf{A}_+) = \hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$, then the likelihood of the data given the structural parameters $(\mathbf{A}_0, \mathbf{A}_+)$ is equal to the likelihood of the data given the reduced-form parameters $(\mathbf{B}, \mathbf{\Sigma})$. □

The following algorithm shows how to use Theorem 2 together with Algorithm 1 to independently draw from the posterior of structural parameters conditional on the sign restrictions

Algorithm 2.

1. *Draw $(\mathbf{B}, \mathbf{\Sigma})$ from the posterior distribution of the reduced-form parameters.*

2. *Use Theorem 1 to draw an orthogonal matrix $\mathbf{Q}$ from the uniform distribution with respect to the Haar measure on $O(n)$.*

3. *Because of Theorem 2,* $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$ *will be a draw from the unrestricted posterior.*

4. *Keep the draw if* $\mathbf{S}_j f\left(h(\mathbf{\Sigma})^{-1}\mathbf{Q}, \mathbf{B}h(\mathbf{\Sigma})^{-1}\mathbf{Q}\right)\mathbf{e}_j > \mathbf{0}$ *are satisfied for* $j \in \mathcal{S}$*.*

5. *Return to Step 1 until the required number of draws from the posterior of structural parameters conditional on the sign restrictions has been obtained.*

As mentioned above, as long as the reduced-form prior is multivariate normal inverse Wishart, we have an efficient algorithm to obtain the independent draws required in step one of Algorithm 2. Theorem 1 gives an efficient algorithm to obtain the independent draws required in step two and step three is justified by Theorem 2. In practice, $h(\mathbf{\Sigma})$ is normally the Cholesky decomposition $\mathbf{\Sigma}$ such that $h(\mathbf{\Sigma})$ is upper triangular with positive diagonal. In any case, Theorem 2 also shows that other mappings are possible.

3.4 A Recursive Formulation of Theorem 1

At this point it is useful to understand how Theorem 1 works, and more important, how it can be implemented recursively. While it is more efficient to obtain the orthogonal matrix $\mathbf{Q}$ in a single step via the QR decomposition, the fact that it can be obtained recursively will be of use when analyzing the zero restrictions. Furthermore, note that the recursive formulation allows a faster implementation of Algorithm 2 in those cases in which the researcher is interested in identifying less than n shocks.

Theorem 3. *Let* $\mathbf{X} = [\mathbf{x}_1 \ \cdots \ \mathbf{x}_n]$ *be an* $n \times n$ *random matrix with each element having an independent standard normal distribution. Let*

$$\mathbf{q}_j = \frac{[\mathbf{N}_{j-1} \ \ \mathbf{0}_{n,j-1}] \cdot \mathbf{x}_j}{\| \ [\mathbf{N}_{j-1} \ \ \mathbf{0}_{n,j-1}] \cdot \mathbf{x}_j \ \|} \text{ for } 1 \leq j \leq n,$$

where $\mathbf{N}_{j-1}$ *is any* $n \times (n-j+1)$ *matrix whose columns form an orthonormal basis for the null space of the matrix* $\mathbf{Q}'_{j-1}$*,* $\mathbf{Q}_{j-1} = [\mathbf{q}_1 \ \cdots \ \mathbf{q}_{j-1}]$*, and* $\mathbf{0}_{n,j-1}$ *is the* $n \times (j-1)$ *matrix of zeros. We follow the convention that* $\mathbf{Q}_0$ *is the* $n \times 0$ *empty matrix and* $\mathbf{N}_0$ *is the* $n \times n$ *identity matrix. The random matrix* $\mathbf{Q} = [\mathbf{q}_1 \ \cdots \ \mathbf{q}_n]$ *has the uniform distribution with respect to the Haar measure on* $O(n)$*.*

Proof. The proof follows from a reformulation of Theorem 1 using the Gram-Schmidt process. □

Alternatively, note that it is possible to write

$$\mathbf{q}_j = \frac{\mathbf{N}_{j-1}\mathbf{y}_j}{\| \ \mathbf{N}_{j-1} \cdot \mathbf{y}_j \ \|} \text{ for } 1 \leq j \leq n,$$

where $\mathbf{y}_j = \mathbf{N}'_{j-1}\mathbf{x}_j$ is a standard normal draw from $\mathbb{R}^{n-j+1}$ and $\mathbf{y}_j / \parallel \mathbf{y}_j \parallel$ is a draw from the uniform distribution on the unit sphere centered at the origin in $\mathbb{R}^{n-j+1}$, which is denoted by S^{n-j}. Because the columns of $\mathbf{N}_{j-1}$ are orthonormal, multiplication by $\mathbf{N}_{j-1}$ is a rigid transformation of $\mathbb{R}^{n-j+1}$ into $\mathbb{R}^n$. From this alternative geometric representation, one can see why Theorems 1 and 3 produce uniform draws from $O(n)$. For $1 \leq j \leq n$, the vector $\mathbf{q}_j$, conditional on $\mathbf{Q}_{j-1}$, is a draw from the uniform distribution on S^{n-j}.[4]

3.5 Algorithm with Sign and Zero Restrictions

Let us now assume that we also want to impose zero restrictions at several horizons, both finite and infinite. Similar to the case of sign restrictions, we use the function $f(\mathbf{A}_0, \mathbf{A}_+)$ to stack the IRFs at the desired horizons. The function $f(\mathbf{A}_0, \mathbf{A}_+)$ will contain IRFs for both sign and zero restrictions. Zero restrictions can be represented by matrices $\mathbf{Z}_j$ for $1 \leq j \leq n$, where the number of columns in $\mathbf{Z}_j$ is equal to the number of rows in $f(\mathbf{A}_0, \mathbf{A}_+)$. If the rank of $\mathbf{Z}_j$ is z_j, then z_j is the number of zero restrictions associated with the j-th structural shock. We only specify matrices $\mathbf{Z}_j$ for the structural shocks that are identified using zero restrictions.[5] Without loss of generality we assume that we only impose zero restrictions on the first k structural shocks, where $1 \leq k \leq n$.[6] The total number of zero restrictions will be $z = \sum_{j=1}^{k} z_j$.

Definition 4. *Let $(\mathbf{A}_0, \mathbf{A}_+)$ be any value of structural parameters. These parameters satisfy the zero restrictions if and only if $\mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{e}_j = \mathbf{0}$ for $1 \leq j \leq k$.*

We can no longer use Algorithm 1 for sampling from the posterior of structural parameters conditional on the sign and the zero restrictions, since the set of structural parameters conditional on the zero restrictions will be of measure zero in the set of all structural parameters. As we show below, as long as there are not too many zero restrictions, we will be able to directly obtain draws from the posterior of structural parameters conditional on the zero restrictions instead of draws from the unrestricted posterior. This is important for the same reasons used to motivate Algorithm 1. The set of structural parameters conditional on the sign and the zero restrictions will be of positive measure in the set of structural parameters conditional on the zero restrictions. Thus, we will be able to use

[4]While draws from $O(n)$ can be obtained recursively by drawing from S^{n-j} for $1 \leq j \leq n$, $O(n)$ is not topologically equivalent to a product of spheres, i.e., there does not exist a continuous bijection from $O(n)$ to $\prod_{j=1}^{n} S^{n-j}$.

[5]In this section, we assume that there is, at least, one zero restriction. If there are no zero restrictions we just implement Algorithm 2.

[6]Ordering them first is going to be convenient for the notation to follow.

the following algorithm for sampling from the posterior of structural parameters conditional on the sign and the zero restrictions.

Algorithm 3.

1. *Draw* $(\mathbf{A}_0, \mathbf{A}_+)$ *from the posterior of structural parameters conditional on the zero restrictions.*

2. *Keep the draw if the sign restrictions are satisfied.*

3. *Return to Step 1 until the required number of draws from the posterior of structural parameters conditional on the sign and the zero restrictions has been obtained.*

As was the case before, the only obstacle to implementing Algorithm 3 is an efficient technique to accomplish the first step. We now present an algorithm that produces draws from the posterior of structural parameters conditional on the zero restrictions.

3.6 Draws from the Posterior Conditional on the Zero Restrictions

This subsection will show that there is also a way to convert a draw from the posterior of the reduced-form parameters to a draw from the posterior of structural parameters conditional on the zero restrictions. The difficulty resides in that we cannot use draws of orthogonal matrices from the uniform distribution with respect to the Haar measure on $O(n)$. Instead, we will require draws of orthogonal matrices from the uniform distribution with respect to the Haar measure on $O(n)$ conditional on the zero restrictions.

We will first describe the constraints that the zero restrictions impose on the orthogonal matrices. We will show that they are linear. Second we show how to draw from the needed conditional uniform distribution. Then, we will show that these draws allow us to travel from the posterior of the reduced-form parameters into the posterior of structural parameters conditional on the zero restrictions. Finally, we modify Algorithm 3 to incorporate our findings.

3.6.1 Linear Constraints on the Orthogonal Matrices

From Definition 4 it is easy to see that the zero restrictions impose non-linear constraints on the structural parameters. We now show how to transform these into linear constraints on the orthogonal matrix $\mathbf{Q}$. Given any value of the reduced-form parameters $(\mathbf{B}, \mathbf{\Sigma})$ and a mapping h zero restrictions on the IRFs can be converted into linear restrictions on the columns of the orthogonal matrix $\mathbf{Q}$. To see this, note that

$$\mathbf{Z}_j f\left(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q}\right)\mathbf{e}_j = \mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{Q}\mathbf{e}_j = \mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_j \text{ for } 1 \leq j \leq n,$$

where $(\mathbf{A}_0, \mathbf{A}_+) = \hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{I}_n) = (h(\mathbf{\Sigma})^{-1}, \mathbf{B}h(\mathbf{\Sigma})^{-1})$.

Therefore, given any value of reduced-form parameters and a mapping h, the zero restrictions associated with the j-th structural shock can be expressed as linear restrictions on the j-th column of the orthogonal matrix $\mathbf{Q}$. Thus, the zero restrictions will hold if and only if

$$\mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_j = \mathbf{0} \text{ for } 1 \leq j \leq k, \tag{4}$$

where $(\mathbf{A}_0, \mathbf{A}_+) = \hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{I}_n) = (h(\mathbf{\Sigma})^{-1}, \mathbf{B}h(\mathbf{\Sigma})^{-1})$. In addition to equation (4), we need the resulting matrix $\mathbf{Q}$ to be orthonormal. This condition imposes extra linear constraints on the columns of $\mathbf{Q}$.

At this point two things should be clear. First, given a mapping h, any value of the reduced-form parameters $(\mathbf{B}, \mathbf{\Sigma})$ implies the following value of the structural parameters $(\mathbf{A}_0, \mathbf{A}_+) = \hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{I}_n) = (h(\mathbf{\Sigma})^{-1}, \mathbf{B}h(\mathbf{\Sigma})^{-1})$. Thus, in what follows and for simplicity, we will mostly write any value of the structural parameters $(\mathbf{A}_0, \mathbf{A}_+)$ instead of any value of the reduced-form parameters $(\mathbf{B}, \mathbf{\Sigma})$ and a mapping h, and $(\mathbf{A}_0, \mathbf{A}_+)$ instead of either $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{I}_n)$ or $(h(\mathbf{\Sigma})^{-1}, \mathbf{B}h(\mathbf{\Sigma})^{-1})$.

Second that, for any value of the structural parameters $(\mathbf{A}_0, \mathbf{A}_+)$, stating that the orthogonal matrix $\mathbf{Q}$ is such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions is equivalent to saying that the orthogonal matrix $\mathbf{Q}$ is such that $\mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_j = \mathbf{0}$ for $1 \leq j \leq k$. Hence, we will interchange both statements.

3.6.2 Finding Orthogonal Matrices Conditional on the Linear Restrictions

Given any value of the structural parameters $(\mathbf{A}_0, \mathbf{A}_+)$, the next theorem shows when and how we can find an orthogonal matrix $\mathbf{Q}$ such that $\mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_j = \mathbf{0}$ for $1 \leq j \leq k$. Since we have ordered the shocks such that we only impose zero restrictions on the first k structural shocks; hence we only impose zero restrictions on the first k columns of $\mathbf{Q}$, i.e. $\mathbf{q}_1 \ldots \mathbf{q}_k$.

Theorem 4. *For any value of the structural parameters* $(\mathbf{A}_0, \mathbf{A}_+)$*,* $\mathbf{Q}$ *is an orthogonal matrix such* *that* $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ *satisfies the zero restrictions if and only if* $\| \mathbf{q}_j \| = 1$,

$$\mathbf{R}_j\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_j = \mathbf{0}, \textit{ for } 1 \leq j \leq k, \textit{ where} \tag{5}$$

$$\mathbf{R}_j\left(\mathbf{A}_0, \mathbf{A}_+\right) = \left[\begin{array}{ccc} \mathbf{q}_1 \ldots \mathbf{q}_{j-1} & \mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right)' & \mathbf{q}_{j+1} \ldots \mathbf{q}_k \end{array}\right]', \textit{ and}$$

$$\mathbf{Q}'_{j-1}\mathbf{q}_j = \mathbf{0}, \textit{ for } k+1 \leq j \leq n.$$

Furthermore, if z_j *is less or equal than* $n-k$ *for* $1 \leq j \leq k$, *then there will be non-zero solutions of equation* (5) *for all values of* $\mathbf{q}_1 \ldots \mathbf{q}_{j-1}, \mathbf{q}_{j+1} \ldots \mathbf{q}_k$ *and for all* $1 \leq j \leq k$.

Proof. The first statement follows easily from the fact that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions if and only if $\mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_j = \mathbf{0}$ and the matrix $\mathbf{Q}$ is orthogonal if and only if $\| \mathbf{q}_j \| = 1$ and $\mathbf{Q}'_{j-1}\mathbf{q}_j = \mathbf{0}$. The second statement follows from the fact that the rank of $\mathbf{R}_j\left(\mathbf{A}_0, \mathbf{A}_+\right)$ is less than or equal to $z_j + k - 1$. Thus, if $z_j \leq n-k$, then the rank of $\mathbf{R}_j\left(\mathbf{A}_0, \mathbf{A}_+\right)$ will be strictly less than n and there will be non-zero solutions of equation (5). □

When considering sign and zero restrictions, one usually only wants to have a small number of zero restrictions. Hence the condition that z_j must be less or equal than $n-k$ for $1 \leq j \leq k$ will almost always be satisfied in practice. If the condition holds, it is also the case that zero restrictions impose no constraints on the reduced-form parameters but will impose constraints on the orthogonal matrix $\mathbf{Q}$.

Theorem 4 just defines the set of orthogonal matrices that we need to draw from. The next two corollaries of Theorem 3 are important ingredients that will allow us to use the set defined in Theorem 4 to obtain draws of orthogonal matrices from the uniform distribution with respect to the Haar measure on $O(n)$ conditional on the zero restrictions.

Following the convention introduced above, we divide the columns of orthogonal matrix $\mathbf{Q}$ into those that are restricted (the first k) and those that are not (the last $n-k$). The first corollary will allow us to draw from the uniform distribution of $\mathbf{q}_{k+1} \ldots \mathbf{q}_n$ conditional on $\mathbf{Q}_k$ directly using results in Theorem 3. The second corollary will allow us to adapt the results in Theorem 3 to Gibbs sample from the uniform distribution of $\mathbf{Q}_k$ conditional on $\mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\mathbf{q}_j = 0$ for $1 \leq j \leq k$.

Corollary 1. *The vector* $\mathbf{q}_j = [\mathbf{N}_{j-1} \;\; \mathbf{0}_{n,j-1}] \cdot \mathbf{x}_j / \| [\mathbf{N}_{j-1} \;\; \mathbf{0}_{n,j-1}] \cdot \mathbf{x}_j \|$ *is an element of the unit sphere in* $\mathbf{N}_{j-1} \subset \mathbb{R}^n$ *and its distribution is uniform for* $k+1 \leq j \leq n$.

Corollary 1 just states that once we have obtained $\mathbf{Q}_k$, we can draw the rest of the columns of the orthogonal matrix $\mathbf{Q}$ by simply using the results in Theorem 3.

Corollary 2. *For any value of the structural parameters* $(\mathbf{A}_0, \mathbf{A}_+)$*, let* $\widehat{\mathbf{N}}_j$ *be the null space of the matrix* $\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+)$*. If* $\widehat{\mathbf{N}}_j$ *is non-trivial, the vector* $\mathbf{q}_j = [\widehat{\mathbf{N}}_j \;\; \mathbf{0}_{n,n-\widehat{n}_j}] \cdot \mathbf{x}_j / \parallel [\widehat{\mathbf{N}}_j \;\; \mathbf{0}_{n,n-\widehat{n}_j}] \cdot \mathbf{x}_j \parallel$ *is an element of the unit sphere in* $\widehat{\mathbf{N}}_j \subset \mathbb{R}^n$ *and its distribution is uniform for* $1 \leq j \leq k$*, where* $\widehat{n}_j$ *is the rank of* $\widehat{\mathbf{N}}_j$*.*

Corollary 2 just asserts that the columns of $\mathbf{Q}_k$ can also be drawn using the results in Theorem 3. One just need to be careful about defining the appropriate null space.

It should be clear from Corollaries 1 and 2 that for each $(\mathbf{A}_0, \mathbf{A}_+)$ there are many orthogonal matrices $\mathbf{Q}$ such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfy the zero restrictions and that the particular orthogonal matrix $\mathbf{Q}$ to be drawn will depend on the particular draw of $\mathbf{x}_j$ for $1 \leq j \leq n$.

We now show how to use the corollaries to obtain the draws of orthogonal matrices from the uniform distribution with respect to the Haar measure on $O(n)$ conditional on the zero restrictions. We will first describe a Gibbs sampler to obtain $\mathbf{Q}_k$, then we will show how to draw the rest of the orthogonal matrix $\mathbf{Q}$.

3.6.3 A Gibbs Sampler

We need a Gibbs sampler to implement Corollary 2 in order to draw from the uniform distribution of $\mathbf{Q}_k$ conditional on $\mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\mathbf{q}_j = 0$ for $1 \leq j \leq k$. The following algorithm describes such sampler

Algorithm 4. *Let* $(\mathbf{A}_0, \mathbf{A}_+)$ *be any value of structural parameters. Given* $\mathbf{q}_1^{(i+1)}, \ldots, \mathbf{q}_{j-1}^{(i+1)}, \mathbf{q}_{j+1}^{(i)}, \ldots, \mathbf{q}_k^{(i)}$

1. *Let* $\widehat{\mathbf{N}}_j^{(i+1)}$ *be any matrix whose columns form an orthonormal basis for the null space of*

$$\left[\mathbf{q}_1^{(i+1)} \quad \cdots \quad \mathbf{q}_{j-1}^{(i+1)} \quad \mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)' \quad \mathbf{q}_{j+1}^{(i)} \quad \cdots \quad \mathbf{q}_k^{(i)}\right]'.$$

2. *Draw* $\mathbf{x}_j^{(i+1)}$ *from a standard normal and define*

$$\mathbf{q}_j^{(i+1)} = \frac{[\widehat{\mathbf{N}}_j^{(i+1)} \;\; \mathbf{0}_{n,n-\widehat{n}_j^{(i+1)}}] \cdot \mathbf{x}_j^{(i+1)}}{\parallel [\widehat{\mathbf{N}}_j^{(i+1)} \;\; \mathbf{0}_{n,n-\widehat{n}_j^{(i+1)}}] \cdot \mathbf{x}_j^{(i+1)} \parallel}.$$

   where $\widehat{n}_j^{(i+1)}$ *is the rank of* $\widehat{\mathbf{N}}_j^{(i+1)}$*.*

3. Repeat steps (1) and (2) L times to obtain $\mathbf{q}_1^{(L)}, \ldots, \mathbf{q}_k^{(L)}$.

For any admissible starting values $\mathbf{q}_1^{(0)}, \ldots, \mathbf{q}_k^{(0)}$, the distribution of $\mathbf{q}_1^{(L)}, \ldots, \mathbf{q}_k^{(L)}$ converges to the uniform distribution of $\mathbf{Q}_k$ conditional on $\mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\mathbf{q}_j = 0$ for $1 \leq j \leq k$.

To be admissible, the starting values must be orthonormal vectors satisfying $\mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\mathbf{q}_j^{(0)} = 0$ for $1 \leq j \leq k$. Furthermore, convergence of the Gibbs sampler is quicker if the starting values are drawn from a distribution that is close to the target distribution. Therefore, we recommend the following algorithm to randomly choose starting values

Algorithm 5. *Let* $(\mathbf{A}_0, \mathbf{A}_+)$ *be any value of structural parameters. Obtain* $\mathbf{q}_1^{(0)}, \ldots, \mathbf{q}_k^{(0)}$ *as follows.*

1. Let $\widetilde{\mathbf{N}}_j^{(0)}$ *be any matrix whose columns form an orthonormal basis for the null space of*

$$\left[\mathbf{q}_1^{(0)} \quad \cdots \quad \mathbf{q}_{j-1}^{(0)} \quad \mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)'\right]'.$$

2. Draw $\mathbf{x}_j^{(0)}$ *from a standard normal and define*

$$\mathbf{q}_j^{(0)} = \frac{[\widetilde{\mathbf{N}}_j^{(0)} \quad \mathbf{0}_{n,n-\widetilde{n}_j^{(0)}}] \cdot \mathbf{x}_j^{(0)}}{\| \, [\widetilde{\mathbf{N}}_j^{(0)} \quad \mathbf{0}_{n,n-\widetilde{n}_j^{(0)}}] \cdot \mathbf{x}_j^{(0)} \, \|},$$

where $\widetilde{n}_j^{(0)}$ *is the rank of* $\widetilde{\mathbf{N}}_j^{(0)}$.

There is also the question of how large L should be to obtain convergence. Experiments show that for the starting value given below, even $L = 1$ gives a good approximation of the desired distribution. In practice, increasing values of L can be used to determine when convergence has occurred.

The bound on the number of linear restrictions is important in order for the Gibbs sampler to converge. By construction, $\widehat{N}_j^{(i+1)}$ has dimension at least one since $\mathbf{q}_j^{(i)}$ is always in the null space. However, if the dimension of $\widehat{N}_j^{(i+1)}$ were exactly one, then no mixing could occur since it would always be the case that $\mathbf{q}_j^{(i+1)} = \pm\mathbf{q}_j^{(i)}$. Requiring z_j to be strictly less than $n - k$ ensures that the dimension of $\widehat{N}_j^{(i+1)}$ is always strictly larger than one.

3.6.4 Drawing the Rest of the Orthogonal Matrix Q

We could use Corollary 1 to draw the rest of the columns of the orthogonal matrix $\mathbf{Q}$ in a recursive manner. But, as mentioned, it is more efficient to obtain it in a single step via the QR decomposition. The next corollary of Theorem 1 shows how to do exactly that

Corollary 3. *Let* $(\mathbf{A}_0, \mathbf{A}_+)$ *be any value of structural parameters. Let* $\mathbf{Q}_k$ *be obtained using Algorithm 4 on* $(\mathbf{A}_0, \mathbf{A}_+)$. *Let* $\mathbf{X}$ *be an* $(n-k) \times (n-k)$ *random matrix with each element having an independent standard normal distribution. Let* $\mathbf{N}_k\mathbf{X} = \mathbf{Q}_{n-k}\mathbf{R}$ *be the QR decomposition of* $\mathbf{N}_k\mathbf{X}$. *The random matrix* $\mathbf{Q} = [\mathbf{q}_1 \ \ldots \ \mathbf{q}_k \ \mathbf{Q}_{n-k}]$ *has the uniform distribution with respect to the Haar measure on* $O(n)$ *conditional on* $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ *satisfying the zero restrictions.*

With Algorithm 4 and Corollary 3 in hand, we can develop an efficient algorithm to obtain draws from the posterior of structural parameters conditional on the zero restrictions. As before, under a class of priors, there are efficient algorithms for obtaining independent draws from the posterior of the reduced-form parameters. However, we need draws from the posterior of structural parameters conditional on the zero restrictions. We now show that there is a way to convert draws from the posterior of the reduced-form parameters, together with a draw from the uniform distribution with respect to the Haar measure on $O(n)$ conditional on the zero restrictions, to draws from the posterior of structural parameters conditional on the zero restrictions.

Theorem 5. *Let* π *be a prior density on the reduced-form parameters. If* $(\mathbf{B}, \mathbf{\Sigma})$ *is a draw from the reduced-form posterior and* $\mathbf{Q}$ *is a draw from the uniform distribution with respect to the Haar measure on* $O(n)$ *conditional on* $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$ *satisfying the zero restrictions, then* $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$ *is a draw from the posterior of the structural parameters conditional on the zero restrictions with respect to the prior* $\hat{\pi}(\mathbf{A}_0, \mathbf{A}_+) = \pi(\mathbf{B}, \mathbf{\Sigma}) \left| det\left(\hat{h}'(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})\right)\right|^{-1}$.

Proof. The proof follows from the chain rule and the fact that if $(\mathbf{A}_0, \mathbf{A}_+) = \hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$, then the likelihood of the data given the structural parameters $(\mathbf{A}_0, \mathbf{A}_+)$ is equal to the likelihood of the data given the reduced-form parameters $(\mathbf{B}, \mathbf{\Sigma})$. □

The following algorithm shows how to use Theorem 5 together with Algorithm 4 and Corollary 3 to draw from the posterior of structural parameters conditional on the sign restrictions

Algorithm 6.

1. *Draw* $(\mathbf{B}, \mathbf{\Sigma})$ *from the posterior distribution of the reduced-form parameters.*
2. *Use Algorithm 4, applied to* $(\mathbf{A}_0, \mathbf{A}_+) = \hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{I}_n) = (h(\mathbf{\Sigma})^{-1}, \mathbf{B}h(\mathbf{\Sigma})^{-1})$, *to draw* $\mathbf{Q}_k$.
3. *Use Corollary 3, applied to the* $\mathbf{Q}_k$ *obtained in Step 2, to draw an orthogonal matrix* $\mathbf{Q} = [\mathbf{q}_1 \ \ldots \ \mathbf{q}_k \ \mathbf{Q}_{n-k}]$ *from the uniform distribution with respect to the Haar measure on* $O(n)$ *conditional on* $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$ *satisfying the zero restrictions.*

4. *Because of Theorem 5,* $\hat{h}(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q})$ *will be a draw from the posterior of structural parameters conditional on the sign restrictions.*

5. *Keep the draw if* $\mathbf{S}_j f\left(h(\mathbf{\Sigma})^{-1}\mathbf{Q}, \mathbf{B}h(\mathbf{\Sigma})^{-1}\mathbf{Q}\right)\mathbf{e}_j > \mathbf{0}$ *is satisfied for* $j \in \mathcal{S}$*.*

6. *Return to Step 1 until the required number of posterior draws satisfying both the sign and zero restrictions have been obtained.*

From Algorithm 6 it is easy to see that, for each $(\mathbf{B}, \mathbf{\Sigma})$, there is a whole distribution of IRFs such that the restrictions hold. This observation is essential in interpreting the results in Sections 6 and 7. As was the case before, in practice, $h(\mathbf{\Sigma})$ is normally the Cholesky decomposition of $\mathbf{\Sigma}$ such that $h(\mathbf{\Sigma})$ is upper triangular with positive diagonal. In any case, Theorem 5 shows that other mappings are possible. Finally, a recursive implementation of Algorithm 6, where Corollary 1 is used instead of Corollary 3, would increase efficiency for cases in which the number of shocks to be identified is lower than the number of variables.

3.7 Efficiency and Normalization

In implementing Algorithm 6, it is critical that upon finding a $\mathbf{q}_j$ that violates the sign restrictions, one exits back to step 1 and obtains a new draw of the reduced-form parameters. It is tempting to implement the algorithm by simply making draws of $\mathbf{q}_j$ until we find one that satisfies the sign restrictions. However, this will usually lead to draws from the incorrect distribution. The easiest way to see this is to note that some draws of the reduced-form parameters may have large sets of orthogonal matrices $\mathbf{Q}$ that satisfy both the zero and sign restrictions, while other reduced-form parameters may have small sets of orthogonal matrices $\mathbf{Q}$ that satisfy both the zero and sign restrictions.[7] This difference should be reflected in the posterior draws, but if one draws $\mathbf{q}_j$ until one is accepted, this will not be true.

While it is not permissible to draw orthogonal matrices $\mathbf{Q}$ until acceptance, it is permissible to draw a fixed number of orthogonal matrices $\mathbf{Q}$ for each reduced-form draw and then keep all that satisfy the sign restrictions. However, because drawing from the reduced-form parameters is usually very efficient, it is often best to draw one orthogonal matrix $\mathbf{Q}$ for each reduced-form draw. One

[7] This is very different from zero restrictions only. For any reduced-form draw, the set of orthogonal matrices $\mathbf{Q}$ that satisfies only the zero restrictions lies on a unit sphere centered at the origin of dimension $n - j - z_j$, which are all of the same size.

instance in which it is always more efficient to make multiple draws of the orthogonal matrix $\mathbf{Q}$ is in the case of normalization.

If for the j-th shock there are no sign restrictions, then any $\mathbf{q}_j$ will trivially satisfy the sign restrictions. In this case, if $\mathbf{q}_j$ is the draw of the j-th column of the orthogonal matrix $\mathbf{Q}$, then both $\mathbf{q}_j$ and $-\mathbf{q}_j$ will satisfy the sign restrictions. If for the j-th shock there is exactly one sign restriction, then for any $\mathbf{q}_j$ either $\mathbf{q}_j$ or $-\mathbf{q}_j$ will satisfy the sign restriction. In this case, if $\mathbf{q}_j$ is the draw of the j-th column of the orthogonal matrix $\mathbf{Q}$ and $\mathbf{q}_j$ does not satisfy the sign restriction, then $-\mathbf{q}_j$ will. If for the j-th shock there is more than one sign restriction, then it may be the case that neither $\mathbf{q}_j$ nor $-\mathbf{q}_j$ will satisfy the sign restrictions. In this case, if $\mathbf{q}_j$ is the draw of the j-th column of the orthogonal matrix $\mathbf{Q}$ and $\mathbf{q}_j$ does not satisfy the sign, then $-\mathbf{q}_j$ may or may not satisfy the sign restrictions. Nevertheless, it will always improve efficiency to check both $\mathbf{q}_j$ and $-\mathbf{q}_j$ against the sign restrictions and keep all that satisfy the restrictions. Furthermore, the more shocks there are with zero or one sign restriction, the greater the efficiency gains.

If there are no sign restrictions on the j-th shock, and no additional normalization rule is added, we say that the shock is unnormalized. Unnormalized shocks will always have IRFs with distributions that are symmetric about zero. Thus, if we are interested in making inferences about an IRF, then the shock associated with such an IRF should always be normalized. A single sign restriction on a shock is a normalization rule. See Waggoner and Zha (2003) for a discussion of normalization in SVAR models and suggestions for a generic normalization rule. Finally, it is important to remember that, while it is true that normalization rules do not change the statistical properties of the reduced-form, it is the case that different normalization rules can lead to different economic interpretations.

3.8 A General Class of Restrictions

It is worth noting that although we have used the function $f(\mathbf{A}_0, \mathbf{A}_+)$ to stack the IRFs, the theorems and algorithms in this paper work for any $f(\mathbf{A}_0, \mathbf{A}_+)$ that satisfies the conditions described in Rubio-Ramírez, Waggoner and Zha (2010). Hence, our theory works for any $f(\mathbf{A}_0, \mathbf{A}_+)$ that is admissible, regular, and strongly regular as defined below.

Condition 1. The function $f(\mathbf{A}_0, \mathbf{A}_+)$ is admissible if and only if for any $\mathbf{Q} \in O(n)$, $f(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q}) = f(\mathbf{A}_0, \mathbf{A}_+)\mathbf{Q}$.

Condition 2. The function $f(\mathbf{A}_0, \mathbf{A}_+)$ is regular if and only if its domain is open and the transformation is continuously differentiable with $f'(\mathbf{A}_0, \mathbf{A}_+)$ of rank kn.

Condition 3. The function $f(\mathbf{A}_0, \mathbf{A}_+)$ is strongly regular if and only if it is regular and it is dense in the set of $k \times n$ matrices.

This highlights the fact that our theorems and algorithms allow us to consider two additional classes of restrictions (in addition to restrictions on IRFs). First, there are the commonly used linear restrictions on the structural parameters themselves $(\mathbf{A}_0, \mathbf{A}_+)$. This class of restrictions includes the triangular identification as described by Christiano, Eichenbaum and Evans (1996) and the non-triangular identification as described by Sims (1986), King et al. (1994), Gordon and Leeper (1994), Bernanke and Mihov (1998), Zha (1999), and Sims and Zha (2006). Second, there are the linear restrictions on the $\mathbf{Q}$'s themselves. For instance, in the case of the latter restrictions, one can define $f(\mathbf{A}_0, \mathbf{A}_+) = \mathbf{I}_n$. This final class will be useful in comparing our methodology with some existing methods of inference.

4 Example

In this section we present an example to illustrate how to use our theorems and algorithms. We assume some sign and zero restrictions and a draw from the posterior of the reduced-form parameters in order to show how Algorithm 2 allows us to draw a $\mathbf{Q}$ conditional on the sign restrictions, while Algorithm 6 allows us to draw a $\mathbf{Q}$ conditional on the sign and the zero restrictions. Consider a four-variable SVAR with one lag. The dimension and lag length of the SVAR are totally arbitrary. In this section, we will assume that $h(\mathbf{\Sigma})$ is normally the Cholesky decomposition $\mathbf{\Sigma}$ such that $h(\mathbf{\Sigma})$ is upper triangular with positive diagonal.

4.1 A Draw from the Posterior of the Reduced-Form Parameters

Let the following $\mathbf{B}$ and $\mathbf{\Sigma}$ be a particular draw from the posterior of the reduced-form parameters

$$\mathbf{B} = \begin{bmatrix} 0.7577 & 0.7060 & 0.8235 & 0.4387 \\ 0.7431 & 0.0318 & 0.6948 & 0.3816 \\ 0.3922 & 0.2769 & 0.3171 & 0.7655 \\ 0.6555 & 0.0462 & 0.9502 & 0.7952 \end{bmatrix} \text{ and } \mathbf{\Sigma} = \begin{bmatrix} 0.0281 & -0.0295 & 0.0029 & 0.0029 \\ -0.0295 & 3.1850 & 0.0325 & -0.0105 \\ 0.0029 & 0.0325 & 0.0067 & 0.0054 \\ 0.0029 & -0.0105 & 0.0054 & 0.1471 \end{bmatrix}.$$

Let the structural parameters be $(\mathbf{A}_0, \mathbf{A}_+) = (h(\mathbf{\Sigma})^{-1}, \mathbf{B}h(\mathbf{\Sigma})^{-1})$, hence

$$\mathbf{A}_0 = \begin{bmatrix} 5.9655 & 0.5911 & -1.4851 & -0.0035 \\ 0 & 0.5631 & -0.1455 & 0.0321 \\ 0 & 0 & 12.9098 & -2.2906 \\ 0 & 0 & 0 & 2.6509 \end{bmatrix} \text{ and } \mathbf{A}_+ = \begin{bmatrix} 4.5201 & 0.8454 & 9.4033 & -0.7034 \\ 4.4330 & 0.4572 & 7.8615 & -0.5815 \\ 2.3397 & 0.3878 & 3.4710 & 1.3104 \\ 3.9104 & 0.4135 & 11.2867 & -0.0694 \end{bmatrix}.$$

Assume that we want to impose restrictions on the IRFs at horizon zero, two, and infinity. Hence, we compute the respective IRFs and we stack them using function $f(\mathbf{A}_0, \mathbf{A}_+)$ as follows

$$f(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} \mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_2(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_\infty(\mathbf{A}_0, \mathbf{A}_+) \end{bmatrix} = \begin{bmatrix} 0.1676 & 0 & 0 & 0 \\ -0.1760 & 1.7760 & 0 & 0 \\ 0.0173 & 0.0200 & 0.0775 & 0 \\ 0.0173 & -0.0042 & 0.0669 & 0.3772 \\ 0.1355 & 1.9867 & 0.1828 & 0.5375 \\ 0.0259 & 1.3115 & 0.0828 & 0.2882 \\ 0.1377 & 2.1813 & 0.2131 & 0.6144 \\ 0.1069 & 2.0996 & 0.1989 & 0.6281 \\ 0.1091 & -0.3783 & -0.0847 & -0.2523 \\ -0.1170 & 1.2928 & -0.0599 & -0.2201 \\ -0.0422 & -0.7342 & 0.0006 & -0.1695 \\ -0.0575 & -1.1662 & 0.0362 & 0.2577 \end{bmatrix}.$$

4.2 The Restrictions

Assume that we want to impose a negative sign restriction at horizon two on the response of the third variable to the second structural shock, a positive sign restriction at horizon two on the response of the fourth variable to the second structural shock, a negative sign restriction at horizon zero on the response of the second variable to the third structural shock, a positive sign restriction at horizon zero, two, and infinity on the response of the first variable to the fourth structural shock, a zero restriction at horizon zero on the response of the third variable to the first structural shock, and a zero restriction at horizon infinity on the response of the fourth variable to the second structural shock. These restrictions can be enforced using the matrices $\mathbf{S}_j$ and $\mathbf{Z}_j$ for $1 \leq j \leq n$

$$\mathbf{S}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \mathbf{S}_3 = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{S}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \mathbf{Z}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\text{and } \mathbf{Z}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Since there are no sign restrictions associated with the first structural shock, we do not need to specify $\mathbf{S}_1$. Similarly, we do not specify $\mathbf{Z}_3$ and $\mathbf{Z}_4$.

4.3 Sign Restrictions

Let us start by discussing the sign restrictions that can be enforced using Algorithm 2. Assume that we draw

$$\mathbf{X} = \begin{bmatrix} 0.8110 & -1.8301 & -1.0833 & -1.7793 \\ -1.9581 & 0.5305 & -1.5108 & 1.0477 \\ 1.6940 & 0.4499 & -1.8539 & 1.0776 \\ -0.6052 & -0.2418 & -1.8677 & -0.1271 \end{bmatrix},$$

where each element is drawn from an independent standard normal distribution. Then, the orthogonal matrix $\mathbf{Q}$ associated with the QR decomposition is

$$\mathbf{Q} = \begin{bmatrix} 0.2917 & -0.8809 & -0.2226 & 0.2991 \\ -0.7044 & 0.0644 & -0.4764 & 0.5223 \\ 0.6094 & 0.4264 & -0.6430 & 0.1828 \\ -0.2177 & -0.1953 & -0.5569 & -0.7774 \end{bmatrix}.$$

Note that given $\mathbf{Q}$ the sign restrictions are satisfied since

$$\mathbf{S}_2 f(\mathbf{A}_0, \mathbf{A}_+)\,\mathbf{q}_2 = \begin{bmatrix} 0.0100 & 0.0032 \end{bmatrix}' > \mathbf{0}, \mathbf{S}_3 f(\mathbf{A}_0, \mathbf{A}_+)\,\mathbf{q}_3 = 0.8068 > 0,$$
$$\text{and } \mathbf{S}_4 f(\mathbf{A}_0, \mathbf{A}_+)\,\mathbf{q}_4 = \begin{bmatrix} 0.0501 & 0.6937 & 0.0157 \end{bmatrix}' > \mathbf{0}.$$

Nevertheless, there is no reason to expect the zero restrictions to be satisfied for such $\mathbf{Q}$. Indeed, in this case they do not hold,

$$\mathbf{Z}_1 f(\mathbf{A}_0, \mathbf{A}_+)\,\mathbf{q}_1 = \begin{bmatrix} 0.0382 \end{bmatrix} \neq \mathbf{0}, \text{ and } \mathbf{Z}_2 f(\mathbf{A}_0, \mathbf{A}_+)\,\mathbf{q}_2 = -0.0594 \neq 0.$$

4.4 Sign and Zero Restrictions

We now illustrate how to find a $\mathbf{Q}$ that satisfies the sign and zero restrictions based on Algorithm 6. Assume that in step 1 we use our draw from the posterior of the reduced-form parameters. Then, step 2 of Algorithm 6 is as follows:

1. Initialize the Gibbs Sampler:

   (a) Let $j = 1$.

   (b) Find a matrix $\widetilde{\mathbf{N}}_1^{(0)}$ whose columns form an orthonormal basis for the null space of

$$\begin{bmatrix}\mathbf{q}_1^{(0)} & \cdots & \mathbf{q}_1^{(0)} & \mathbf{Z}_1 f(\mathbf{A}_0, \mathbf{A}_+)'\end{bmatrix}', \widetilde{\mathbf{N}}_1^{(0)} = \begin{bmatrix} -0.2445 & 0.9506 & -0.1910 & 0 \\ -0.9463 & -0.1910 & 0.2607 & 0 \\ 0 & 0 & 0 & 1.0000 \end{bmatrix}'.$$

(c) Draw $\mathbf{x}_1^{(0)}$ from the standard normal distribution on $\mathbb{R}^4$, e.g. $\mathbf{x}_1^{(0)} = \begin{bmatrix} 0.2750 \\ -0.0020 \\ 0.4877 \\ -0.1961 \end{bmatrix}$.

(d) Let $\mathbf{q}_1^{(0)} = \frac{[\widetilde{\mathbf{N}}_1^{(0)} \ \mathbf{0}_{4,1}] \ \mathbf{x}_1^{(0)}}{\|[\widetilde{\mathbf{N}}_1^{(0)} \ \mathbf{0}_{4,1}] \ \mathbf{x}_1^{(0)}\|}$, $\mathbf{q}_1^{(0)} = \begin{bmatrix} -0.1167 & 0.4676 & -0.0948 & 0.8711 \end{bmatrix}'$.

(e) Similarly, when $j = 2$ we obtain the following matrices:

$$\widetilde{\mathbf{N}}_2^{(0)} = \begin{bmatrix} -0.0473 & 0.9886 \\ 0.0498 & -0.0156 \\ 0.9948 & 0.0369 \\ 0.0751 & 0.1449 \end{bmatrix}, \quad \mathbf{x}_2^{(0)} = \begin{bmatrix} 1.6144 \\ 0.1660 \\ 0.1966 \\ -0.1774 \end{bmatrix}, \quad \mathbf{q}_2^{(0)} = \begin{bmatrix} 0.0540 \\ 0.0479 \\ 0.9934 \\ 0.0896 \end{bmatrix}.$$

2. Apply the Gibbs Sampler L times. We set $L = 1$ in this example.

(a) Let $\ell = 1$. Set $j = 1$, and compute $\widehat{\mathbf{N}}_1^1 = \begin{bmatrix} -0.8010 & 0.5774 & 0.0297 & -0.1553 \\ 0.0655 & 0.3464 & -0.1041 & 0.9300 \end{bmatrix}'$.

(b) Draw $\mathbf{x}_1^1$, e.g. $\mathbf{x}_1^{(1)} = \begin{bmatrix} 2.1702 & -1.1935 & 1.4193 & 0.2916 \end{bmatrix}'$.

(c) Let $\mathbf{q}_1^{(1)} = \frac{[\widehat{\mathbf{N}}_1 \ \mathbf{0}_{4,2}] \ \mathbf{x}_1^{(1)}}{\|[\widehat{\mathbf{N}}_1 \ \mathbf{0}_{4,2}] \ \mathbf{x}_1^{(1)}\|}$, $\mathbf{q}_1^{(1)} = \begin{bmatrix} -0.7334 & 0.3390 & 0.0762 & -0.5842 \end{bmatrix}'$.

(d) Set $j = 2$, and compute $\widehat{\mathbf{N}}_2^1 = \begin{bmatrix} 0.0950 & 0.0327 & 0.9945 & 0.0294 \\ -0.5477 & 0.2068 & 0.0215 & 0.8104 \end{bmatrix}'$.

(e) Draw $\mathbf{x}_2^{(1)}$, e.g. $\mathbf{x}_2^{(1)} = \begin{bmatrix} -1.4300 & 0.5002 & 0.1978 & -0.8045 \end{bmatrix}'$.

(f) Let $\mathbf{q}_2^{(1)} = \frac{[\widehat{\mathbf{N}}_2 \ \mathbf{0}_{4,2}] \ \mathbf{x}_2^{(1)}}{\|[\widehat{\mathbf{N}}_2 \ \mathbf{0}_{4,2}] \ \mathbf{x}_2^{(1)}\|}$, $\mathbf{q}_2^{(1)} = \begin{bmatrix} -0.2705 & 0.0374 & -0.9316 & 0.2398 \end{bmatrix}'$.

Now that we have drawn the columns of $\mathbf{Q}$ associated with zero restrictions, we will draw those columns not associated with zero restrictions. This can be done using Corollary 3:

1. Compute $\mathbf{N}_2 = \begin{bmatrix} -0.4901 & -0.8477 & 0.1449 & 0.1423 \\ -0.3856 & 0.4064 & 0.3245 & 0.7622 \end{bmatrix}'$.

2. Draw $\mathbf{X}$, e.g. $\mathbf{X} = \begin{bmatrix} 1.1389 & -2.0445 \\ -1.2003 & -1.8605 \end{bmatrix}$.

3. Then applying the QR decomposition to $\mathbf{N}_2\mathbf{X}$ we obtain

$$\mathbf{Q}_2 = \begin{bmatrix} 0.0984 & 0.6158 \\ -0.7675 & 0.5428 \\ -0.2130 & -0.2844 \\ -0.5966 & -0.4952 \end{bmatrix}.$$

In this particular case we have chosen the random draws so that the sign restrictions also hold

$$\mathbf{S}_2 f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{q}_2 = \begin{bmatrix} 0.0068 & 0.0150 \end{bmatrix}' > \mathbf{0}, \mathbf{S}_3 f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{q}_3 = 1.3804 > 0,$$
$$\text{and } \mathbf{S}_4 f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{q}_4 = \begin{bmatrix} 0.1032 & 0.8436 & 0.0109 \end{bmatrix}' > \mathbf{0}.$$

Clearly, the fact that the sign restrictions hold depends on the draw of $\mathbf{x}_j^{\ell}$ for $1 \leq j \leq k$ and $\mathbf{X}$.

5 The Mountford and Uhlig Methodology

In this section, we discuss the PFA with sign and zero restrictions developed by Mountford and Uhlig (2009). First, we describe the algorithm. Second, we highlight how it selects one particular orthogonal matrix $\mathbf{Q}$ instead of drawing from the conditional uniform distribution derived in Subsection 3.5. We also analyze the consequences of this drawback. Third, we formally show how selecting a particular orthogonal matrix $\mathbf{Q}$ imposes additional sign restrictions on variables that are seemingly unrestricted.

5.1 Penalty Function Approach with Sign and Zero Restrictions

Let $(\mathbf{A}_0, \mathbf{A}_+)$ be any draw of the structural parameters. Consider a case where the identification of the j-th structural shock restricts the IRF of a set of variables indexed by $I_{j,+}$ to be positive and the IRF of a set of variables indexed by $I_{j,-}$ to be negative, where $I_{j,+}$ and $I_{j,-} \subset \{0, 1, \ldots, n\}$. Furthermore, assume that the restrictions on variable $i \in I_{j,+}$ are enforced during $H_{i,j,+}$ periods and the restrictions on variable $i \in I_{j,-}$ are enforced during $H_{i,j,-}$ periods. In addition to the sign restrictions, assume that the researcher imposes zero restrictions on the IRFs to identify the j-th structural shock. Let $\mathbf{Z}_j$ and $f(\mathbf{A}_0, \mathbf{A}_+)$ denote the latter. The PFA finds an orthogonal matrix $\bar{\mathbf{Q}}^* = \begin{bmatrix} \bar{\mathbf{q}}_1^* & \cdots & \bar{\mathbf{q}}_n^* \end{bmatrix}$ such that the IRFs come close to satisfying the sign restrictions, conditional on the zero restrictions being satisfied, according to a loss function.[8] In particular, for $1 \leq j \leq n$, this approach solves the following optimization problem

$$\bar{\mathbf{q}}_j^* = \text{argmin}_{\bar{\mathbf{q}}_j \in S} \quad \Psi\left(\bar{\mathbf{q}}_j\right)$$

subject to

$$\mathbf{Z}_j f\left(\mathbf{A}_0, \mathbf{A}_+\right) \bar{\mathbf{q}}_j = \mathbf{0} \text{ and } \bar{\mathbf{Q}}_{j-1}^{*'} \bar{\mathbf{q}}_j = 0$$

where

$$\Psi\left(\bar{\mathbf{q}}_j\right) = \sum_{i \in I_+} \sum_{h=0}^{H_{i,+}} g\left(-\frac{\mathbf{e}_i' \mathbf{L}_h\left(\mathbf{A}_0, \mathbf{A}_+\right) \bar{\mathbf{q}}_j}{\sigma_i}\right) + \sum_{i \in I_-} \sum_{h=0}^{H_{i,-}} g\left(\frac{\mathbf{e}_i' \mathbf{L}_h\left(\mathbf{A}_0, \mathbf{A}_+\right) \bar{\mathbf{q}}_j}{\sigma_i}\right),$$

$g(\omega) = 100\omega$ if $\omega \geq 0$ and $g(\omega) = \omega$ if $\omega \leq 0$, σ_i is the standard error of variable i, $\bar{\mathbf{Q}}_{j-1}^* = \begin{bmatrix} \bar{\mathbf{q}}_1^* & \cdots & \bar{\mathbf{q}}_{j-1}^* \end{bmatrix}$ for $1 \leq j \leq n$, and $S = S^0$. We follow the convention that $\bar{\mathbf{Q}}_0^*$ is the $n \times 0$ empty matrix.[9]

As before, if the prior on the reduced-form parameters is conjugate, then the posterior of the reduced-form parameters will have the multivariate normal inverse Wishart distribution. As mentioned, there are very efficient algorithms for obtaining independent draws from this distribution. In practice, the researcher will use the above algorithm where $(\mathbf{A}_0, \mathbf{A}_+) = (h(\mathbf{\Sigma})^{-1}, \mathbf{B}h(\mathbf{\Sigma})^{-1})$ with $h(\mathbf{\Sigma})$ set equal to the Cholesky decomposition of $\mathbf{\Sigma}$ such that $h(\mathbf{\Sigma})$ is upper triangular with positive diagonal. We will make this assumption throughout this section. This will also be the case in Section 6 and Section 7. For ease of exposition, in the rest of the paper we will assume the notational convention that $\mathbf{T}$ is equivalent to $h(\mathbf{\Sigma})$. Hence, $\mathbf{T}$ is the Cholesky decomposion of $\mathbf{\Sigma}$.

5.2 Choosing a Single Orthogonal Matrix Q

As mentioned above, the set of structural parameters satisfying the sign and zero restrictions is of positive measure on the set of structural parameters satisfying the zero restrictions. Conditional on a draw from the posterior of the reduced-form parameters, our Algorithm 6 uses this result to draw from the uniform distribution of orthogonal matrices conditional on the zero restrictions being satisfied. The PFA abstracts from using the result. Instead, given any draw of the reduced-form parameters,

[8] See Mountford and Uhlig (2009) for details.

[9] To obtain σ_i, we compute the standard deviation of the OLS residuals associated with the i-th variable.

$(\mathbf{B}, \boldsymbol{\Sigma})$, the penalty function chooses an optimal orthogonal matrix $\bar{\mathbf{Q}}^* = \begin{bmatrix} \bar{\mathbf{q}}_1^* & \cdots & \bar{\mathbf{q}}_n^* \end{bmatrix} \in O(n)$ that solves the following system of equations

$$\mathbf{Z}_j f\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j^* = \mathbf{0} \text{ and}$$

$$\Psi\left(\bar{\mathbf{q}}_j^*\right) = \sum_{i \in I_+} \sum_{h=0}^{H_{i,+}} g\left(-\frac{\mathbf{e}_i' \mathbf{L}_h \left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j^*}{\sigma_i}\right) + \sum_{i \in I_-} \sum_{h=0}^{H_{i,-}} g\left(\frac{\mathbf{e}_i' \mathbf{L}_h \left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j^*}{\sigma_i}\right),$$

for $1 \leq j \leq n$ where, in practice, it is also the case that $(\mathbf{A}_0, \mathbf{A}_+) = (\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1})$ and $\Psi\left(\bar{\mathbf{q}}_j^*\right)$ is the value of the loss function at the optimal value $\bar{\mathbf{q}}_j^*$. Of course, the optimal orthogonal matrix that solves the system of equations is the one that minimizes the loss function.

There are, at least, three possible issues with this approach. First, the optimal orthogonal matrix $\bar{\mathbf{Q}}^*$ that solves the system of equations may be such that the sign restrictions do not hold. Second, since only one orthogonal matrix is chosen, the researcher is clearly not considering all possible values of the structural parameters conditional on the sign and zero restrictions. In the applications, we will see how this issue greatly affects the confidence intervals. Third, it is easy to guess that by choosing a single orthogonal matrix to minimize a loss function, we may be introducing bias on the IRFs and other statistics of interest. Assume that the IRFs of two variables to a particular shock are correlated. Then, by choosing a particular orthogonal matrix that maximizes the response of one of the variables to the shock by minimizing the loss function, we are biasing the response of the other variable to the same shock. The PFA behaves as if there were additional sign restrictions on variables that are seemingly unrestricted and, hence, violates the agnosticism of any identification scheme being used. In general, it is hard to formally prove such a claim because the optimal orthogonal matrix, $\bar{\mathbf{Q}}^*$, is a function of the draw of the reduced-form parameters; hence, in most cases, we will just be able to look at the correlations between IRFs. These correlations are useful in understanding any bias that one could find, but they fall short of being a formal argument. Fortunately, there are exceptions. In the next subsection, we present a class of sign and zero restrictions where this claim can be formally proved. For this class of restrictions, we will formally show how choosing a single orthogonal matrix may impose additional restrictions on variables that are seemingly unrestricted. Nevertheless, even without a formal proof for a general class of sign and zero restrictions, this is a very serious drawback because the most attractive feature of sign restrictions is that one can be agnostic about the response of some variables of interest to some structural shocks. The applications will also highlight the dramatic economic implications of this final issue.

5.3 Is the Penalty Function Approach Truly Agnostic?

We now formally show how the PFA imposes additional sign restrictions on variables that are seemingly unrestricted. In this sense, the procedure is not truly agnostic and introduces bias in the IRFs and other statistics of interest. As argued above, choosing a single orthogonal matrix minimizing a loss function is likely to introduce some bias. Nevertheless, it is hard to formally prove this because the optimal orthogonal matrix depends on a given draw of reduced-form parameters. Fortunately, there is a class of sign and zero restrictions for which a formal proof is indeed possible because the optimal orthogonal matrix is independent of the draw of the reduced-form parameters.

Consider a structural vector autoregression with n variables, and assume that we are interested in imposing a positive sign restriction at horizon zero on the response of the second variable to the j-th structural shock, and a zero restriction at horizon zero on the response of the first variable to the j-th structural shock.[10] Let $(\mathbf{B}, \mathbf{\Sigma})$ be any draw from the posterior of the reduced-form parameters. Then, to find the optimal orthogonal matrix, $\bar{\mathbf{Q}}^*$, we need to solve the following problem

$$\bar{\mathbf{q}}_j^* = \text{argmin}_{\bar{\mathbf{q}}_j \in S} \quad \Psi\left(\bar{\mathbf{q}}_j\right)$$

subject to

$$\mathbf{e}_1' \mathbf{L}_0\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j = 0 \tag{6}$$

where

$$\Psi\left(\bar{\mathbf{q}}_j\right) = g\left(-\frac{\mathbf{e}_2' \mathbf{L}_0\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j}{\sigma_2}\right).$$

Note that we are identifying only one structural shock; therefore, we do not need to impose the orthogonality constraint between the different columns of $\bar{\mathbf{Q}}^*$.

Equation (6) implies that the optimal $\bar{\mathbf{q}}_j^*$ has to be such that $\mathbf{e}_1' \mathbf{L}_0\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j^* = \mathbf{e}_1' \mathbf{T}' \bar{\mathbf{q}}_j^* = \mathbf{t}_{1,1} \bar{\mathbf{q}}_{1,j}^* = 0$, where the next to last equality follows because $\mathbf{T}'$ is lower triangular. Thus, $\bar{\mathbf{q}}_{1,j}^* = 0$. To find the remaining entries of $\bar{\mathbf{q}}_j^*$, it is convenient to write $\mathbf{e}_2' \mathbf{L}_0\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j = \mathbf{e}_2' \mathbf{T}' \bar{\mathbf{q}}_j =$

[10]The order of the restrictions is not important. It is also the case that the results in this subsection hold when we have several zero restrictions and a single sign restriction identifying a particular structural shock. We choose to present the results for a single zero restriction to simplify the argument.

$\sum_{s=1}^{2} \mathbf{t}_{s,2}\bar{\mathbf{q}}_{s,j}$, where the last equality follows because $\mathbf{T}'$ is lower triangular. Substituting $\bar{\mathbf{q}}_{1,j}^{*} = 0$ into $\mathbf{e}_2'\mathbf{L}_0\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right)\bar{\mathbf{q}}_j$ yields $\mathbf{t}_{2,2}\bar{\mathbf{q}}_{2,j}$. If $-\mathbf{e}_2'\mathbf{L}_0\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right)\bar{\mathbf{q}}_j \geq 0$, then $f\left(-\frac{\mathbf{e}_2'\mathbf{L}_0\left(\mathbf{T}^{-1},\mathbf{B}\mathbf{T}^{-1}\right)\bar{\mathbf{q}}_j}{\sigma_2}\right) = -100\frac{\mathbf{t}_{2,2}\bar{\mathbf{q}}_{2,j}}{\sigma_2}$; otherwise, $f\left(-\frac{\mathbf{e}_2'\mathbf{L}_0\left(\mathbf{T}^{-1},\mathbf{B}\mathbf{T}^{-1}\right)\bar{\mathbf{q}}_j}{\sigma_2}\right) = -\frac{\mathbf{t}_{2,2}\bar{\mathbf{q}}_{2,j}}{\sigma_2}$. Since $\bar{\mathbf{q}}_{1,j}^{*} = 0$, and $\bar{\mathbf{q}}^{*}$ must belong to S, it is straightforward to verify that the criterion function is minimized at $\bar{\mathbf{q}}_j^{*} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}'$. Thus, the PFA imposes additional zero restrictions on the orthogonal matrix $\mathbf{Q}$. We now show that these zero restrictions also imply additional sign restrictions on the responses of variables of interest that are seemly unrestricted.

If the PFA were truly agnostic, it would impose no additional sign restrictions on the responses of other variables of interest to the j-th structural shock. In our example, this is not the case; the PFA introduces additional sign restrictions on the response of other variables to the j-th structural shock. To illustrate the problem, note that we have not introduced explicit sign restrictions on any variable except for the second. Nevertheless, the response at horizon zero of the i-th variable to the j-th structural shock for $i > 2$ does not depend on $\bar{\mathbf{q}}_j^{*}$ and it equals

$$\mathbf{e}_i'\mathbf{L}_0\left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right)\bar{\mathbf{q}}_j^{*} = \mathbf{t}_{2,i} \text{ for all } i > 2.$$

Thus, if the posterior distribution of $\mathbf{t}_{2,i}$ differs from the posterior distribution of the IRFs, the PFA will not recover the correct posterior distribution of the IRFs. In some cases, as we will show in our applications, the posterior distribution of $\mathbf{t}_{2,i}$ is such that the event $\mathbf{t}_{2,i} > 0$ ($\mathbf{t}_{2,i} < 0$) occurs more often than it should if correctly drawing from the posterior distribution of the IRFs. In that sense the PFA may introduce additional sign restrictions on seemingly unrestricted variables.

Finally, it is worth noting that the result that the criterion function is minimized at

$$\bar{\mathbf{q}}_j^{*} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}'$$

implies that, for this class of sign and zero restrictions, the Mountford and Uhlig (2009) methodology can be seen as a particular case of ours. Why? Because having the j-th column of the orthogonal matrix equal to $\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}'$ can always be enforced by zero restrictions on the j-th column of the orthogonal matrix. In Subsection 6.3.1 we will show how to implement those restrictions in the case of optimism shocks.

6 Application to Optimism Shocks

In this section, we use our methodology to study one application related to optimism shocks previously analyzed in the literature by Beaudry, Nam and Wang (2011) using the PFA. The aim of Beaudry, Nam and Wang (2011) is to contribute to the debate regarding the source and nature of business cycles. The authors claim to provide new evidence on the relevance of optimism shocks as the main driver of macroeconomic fluctuations using sign and zero restrictions to isolate optimism shocks. At least in their benchmark identification scheme, Beaudry, Nam and Wang (2011) want to be agnostic about the response of consumption and hours worked to optimism shocks. As we show below, the problem is that, by using the PFA, they are not being really agnostic about the response of these two variables.

After replicating their results, we repeat their empirical exercises using our methodology − that truly respects the agnosticism of the identification scheme − to show how their main economic conclusion substantially changes. While Beaudry, Nam and Wang (2011) conclude that optimism shocks are associated with standard business cycle type phenomena because they generate a simultaneous boom in output, investment, consumption, and hours worked, we show that, using our truly agnostic methodology, it is very hard to support such a claim. Moreover, they also find that optimism shocks account for a large share of the forecast error variance (FEV) of output, investment, consumption, and hours worked at several horizons. But again, once one uses our methodology such results are also substantially weakened. We also report how our methodology is not only correct, but faster than the PFA.

6.1 Data and Identification Strategy

Beaudry, Nam and Wang (2011) use two data sets. In the first one, they use data on TFP, stock price, consumption, the real federal funds rate, and hours worked. In the second one, they add investment and output. In both data sets, they consider the three identification strategies described in Table 1.

Identification 1 is the benchmark, where optimism shocks (sometimes called bouts of optimism) are identified as positively affecting stock prices and as being orthogonal to TFP at horizon zero. Identification 2 adds a positive response of consumption at horizon zero as an additional restriction to Identification 1. Finally, Identification 3 adds a positive response of the real interest rate at horizon zero to Identification 2. Appendix 9.1 gives details on the priors and the data sets. Identification 1 is agnostic about the response of consumption and hours worked to optimism shocks. As we will see

	Identification 1	Identification 2	Identification 3
Adjusted TFP	0	0	0
Stock Price	+	+	+
Consumption		+	+
Real Interest Rate			+
Hours Worked			
Investment			
Output			

Table 1: Identification Schemes Defined in Beaudry, Nam and Wang (2011)

below, the PFA will not respect this agnosticism.

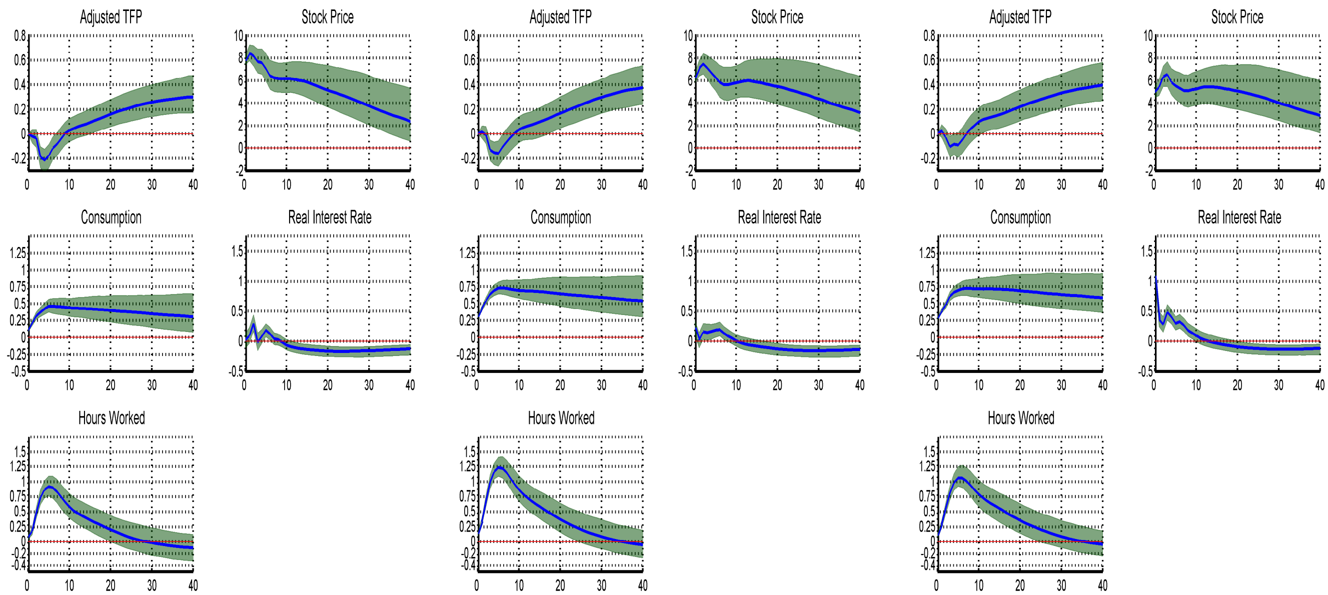
Next, we map these identification strategies to the function $f(\mathbf{A}_0, \mathbf{A}_+)$ and the matrices $\mathbf{S}$s and $\mathbf{Z}$s necessary to apply our methodology. Since the sign and zero restrictions are imposed at horizon zero, we have that $f(\mathbf{A}_0, \mathbf{A}_+) = \mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+)$ in both data sets. The matrices $\mathbf{S}$s and $\mathbf{Z}$s are a function of the number of variables used in the SVAR. In the smaller data set, when five variables are used, the $\mathbf{S}$s matrices are

$$\mathbf{S}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \mathbf{S}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \text{and } \mathbf{S}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

for Identifications 1, 2, and 3 respectively, while the $\mathbf{Z}$ matrix is $\mathbf{Z}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}$. In the larger data set, the sign and zero restrictions are defined analogously.

6.2 IRFs

We first show replications of the IRFs reported in Beaudry, Nam and Wang (2011) using the PFA. Then, we analyze how the results change once we use our methodology. Sometimes we will label our methodology the ARRW methodology. Panel (a) in Figure 3 shows the IRFs of TFP, stock price, consumption, the federal funds rate, and hours worked under Identification 1 when using the PFA on the first data set. This panel replicates the first block of Figure 1 in Beaudry, Nam and Wang (2011). The identified shocks generate a boom in consumption and hours worked. The response of hours worked is hump shaped. We also report 68 percent confidence intervals. Clearly, the confidence intervals associated with the IRFs do not contain zero for, at least, 20 quarters. Thus, it is easy to conclude that optimism shocks generate standard business cycle type phenomena. Panels (b) and



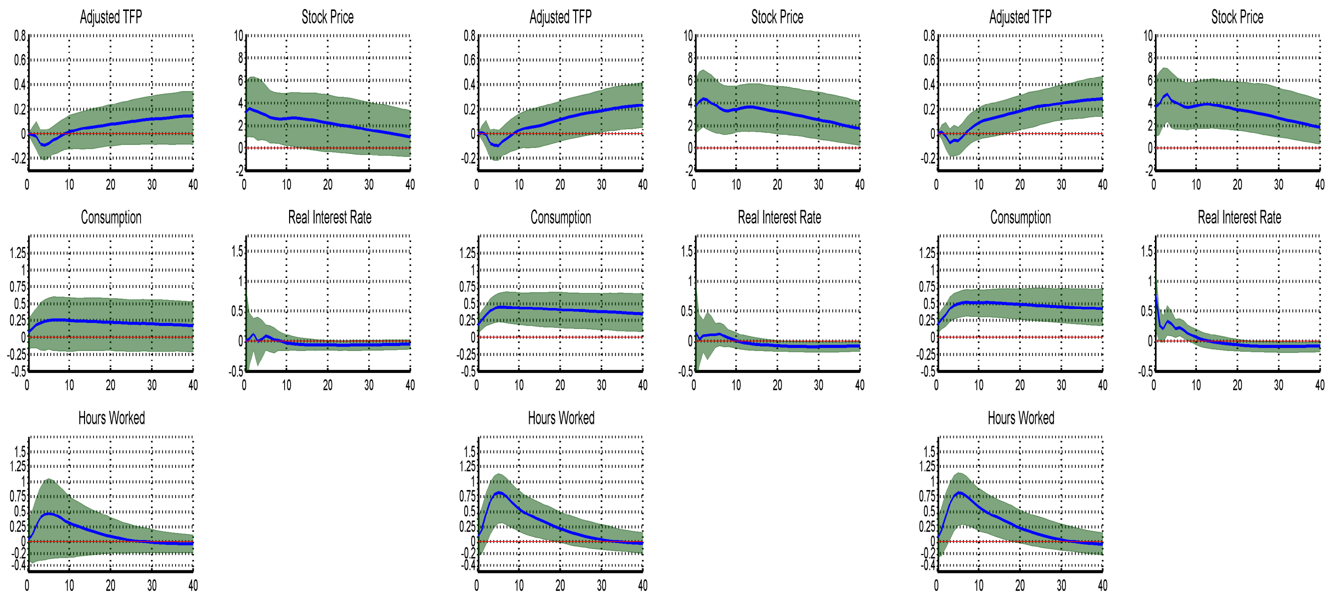
(a) Identification 1 (b) Identification 2 (c) Identification 3

Figure 3: IRFs to an Optimism Shock Using the PFA: Five-Variable SVAR

(c) in Figure 3 show the IRFs of TFP, stock price, consumption, the federal funds rate, and hours worked under Identifications 2 and 3. These panels replicate the second and third blocks of Figure 1 in Beaudry, Nam and Wang (2011). As expected, because of the addition of the sign restrictions on the IRF of consumption, the results are stronger. Using these two identification schemes we also find a positive response of consumption and a positive hump-shaped response of hours worked to optimism shocks. Furthermore, the positive responses last longer than under Identification 1 and the confidence intervals tell us that the IRFs are significantly different from zero. The findings reported in Figure 3 are robust to extending the number of variables. Figure 17 in the appendix shows the results when we consider the larger data set.

As expected, these IRFs are highlighted by Beaudry, Nam and Wang (2011). In theory, Identification 1 is agnostic about the response of consumption and hours worked to an optimism shock, while the identified shock generates a boom in consumption and hours worked. If correct, this conclusion would strongly support the view that optimism shocks are relevant for business cycle fluctuations. But, as we will show below, these IRFs are not correct. They do not reflect the IRFs associated with the agnostic identification scheme 1 because the PFA introduces sign restrictions in addition to the ones described in Table 1.

Once we use the ARRW methodology to compute the correct IRFs, the results highlighted by



(a) Identification 1 (b) Identification 2 (c) Identification 3

Figure 4: IRFs to an Optimism Shock Using the ARRW Methodology: Five-Variable SVAR

Beaudry, Nam and Wang (2011) basically disappear. Panel (a) in Figure 4 reports the results for the first data set using the ARRW methodology under Identification 1. There are three important differences with the results reported in Beaudry, Nam and Wang (2011). First, the PFA chooses a very large median response of stock prices in order to minimize the loss function. Second, the median IRFs for consumption and hours worked are closer to zero when we use the ARRW methodology. Third, the confidence intervals associated with the ARRW are much larger than the ones obtained with the PFA. As a consequence, using the PFA, there is an upward bias in the IRFs and artificially narrow confidence intervals.

We need to consider Identifications 2 and 3 (see Panels (b) and (c) in Figure 4), which force consumption to increase after an optimism shock, to find moderate evidence of positive IRFs of consumption and hours worked. But it is still the case that the median response of stock prices is weaker, the median IRFs of consumption and hours worked are closer to zero (i.e, the upward bias persists) and the confidence intervals are still quite wide when compared with the ones reported in Beaudry, Nam and Wang (2011). As reported in Figure 18, these findings are robust to considering a larger SVAR.

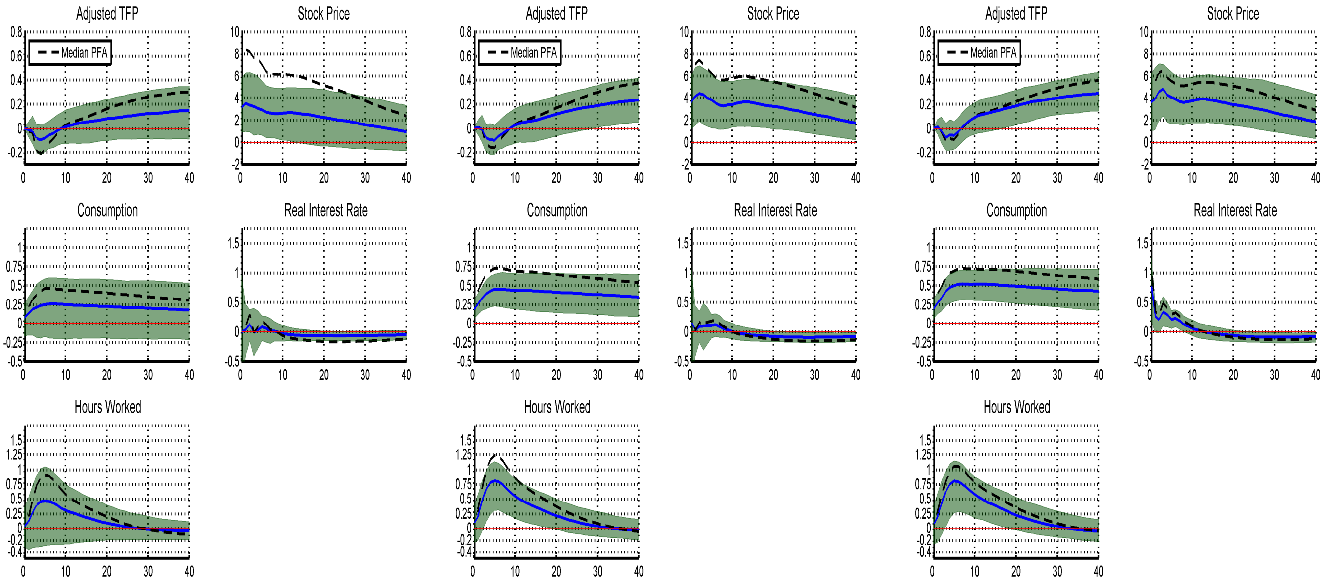
In summary, using the ARRW methodology it is hard to claim that optimism shocks trigger a boom in consumption and hours worked unless we impose a positive response of consumption at horizon zero.

Even after we impose this extra positive sign restriction, the results under the ARRW methodology are much weaker. The sharp results reported in Beaudry, Nam and Wang (2011) are, as indicated above, due to upward bias in the response of consumption and hours worked and artificially narrow confidence intervals associated with the PFA. Once we use the ARRW methodology to solve these two problems, the results disappear. Next, we show that the discrepancy has its origin in the fact that the PFA does not respect the agnosticism of the identification scheme by introducing additional sign restrictions on consumption and hours worked.

6.2.1 Understanding the Bias and the Artificially Narrow Confidence Intervals

We now shed some light on the upward bias and the artificially narrow confidence intervals. Let us begin with the upward bias. We will focus on the five-variable SVAR. In the appendix we show that the same conclusions are obtained using the seven-variable SVAR. Figure 5 plots the median IRFs and the 68 percent confidence intervals obtained using the ARRW methodology and compares them with the median IRFs obtained using the PFA. Figure 19, in the appendix, does the same for the larger SVAR. Clearly, the median IRFs constructed using the PFA are close to the 84-th confidence band constructed using the ARRW methodology. It is easy to observe that the PFA selects a large response of stock prices to optimism shocks in order to minimize the loss function. By choosing a large response of stock prices, the PFA also induces a positive response of consumption and hours worked because the three responses are positively correlated. For the five-variable SVAR the correlation between the IRF of stock prices to an optimism shock at horizon zero with the IRF of consumption to the same shock and horizon is 0.22. In the case of hours worked it is 0.13. The correlations are 0.26 and 0.12 in the larger SVAR. By inducing this positive response of consumption and hours worked, the PFA is introducing sign restrictions on these two variables and, thus, not respecting the agnosticism of the identification scheme.

Let us now consider the artificially narrow confidence intervals. We have repeated several times that the PFA selects a single orthogonal matrix instead of drawing from the conditional uniform distribution. As we mentioned when describing Algorithm 6, for each draw from the posterior distribution of the reduced-form parameters, there is a distribution of IRFs conditional on the sign and zero restrictions holding. By selecting a single orthogonal matrix, the PFA takes a single IRF from such a distribution. Figure 6 plots the 68 percent probability intervals from the distribution of IRFs such that the sign and zero restrictions hold at the OLS point estimate of the reduced-form parameters. These



(a) Identification 1 (b) Identification 2 (c) Identification 3

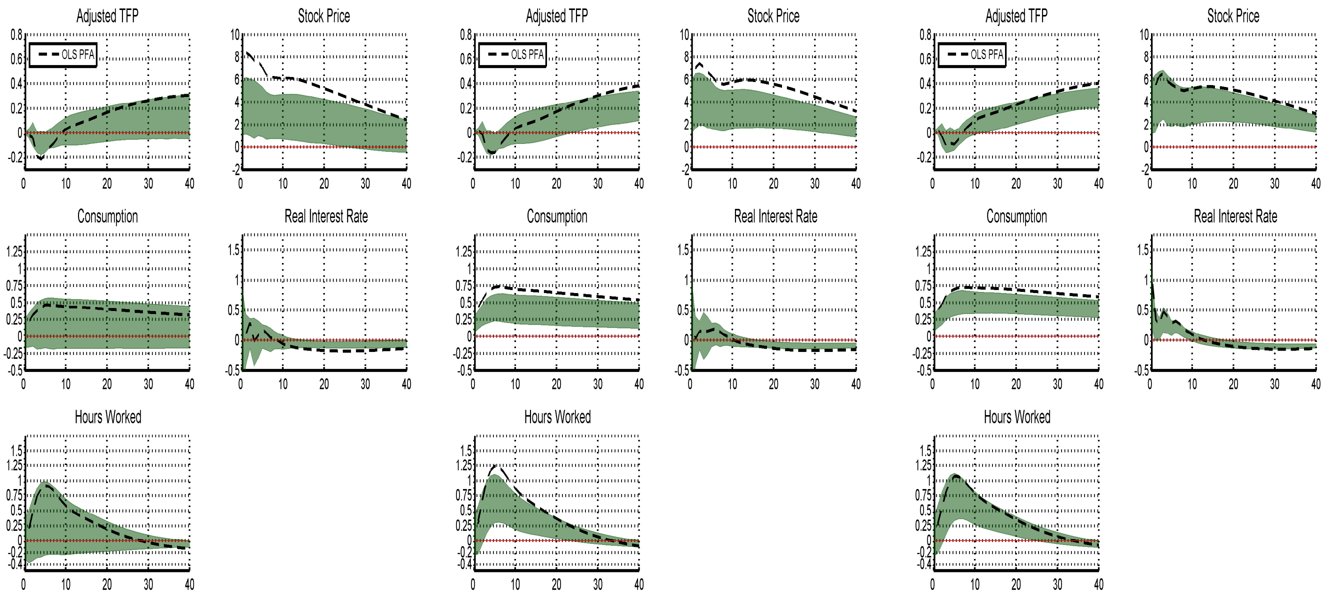
Figure 5: Comparison of IRFs to an Optimism Shock: Five-Variable SVAR

Note: Median PFA refers to the median IRF obtained using the PFA.

intervals are constructed using a single value of the structural parameters, obtained from the Cholesky decomposition and the OLS point estimate of the reduced-form parameters, and several draws of the conditional uniform distribution of the orthogonal matrix $\mathbf{Q}$. We have generated these draws repeating steps 2 and 3 of Algorithm 6 for the single value of the structural parameters. The probability intervals are compared with the single IRFs obtained with the PFA evaluated at the same value of the reduced-form parameters (i.e., the OLS point estimate).[11] In Figure 20 in the appendix we report the results for the seven-variable SVAR . The dashed line shows the value of the IRFs resulting from the PFA. The shadow area describes the 68 percent probability intervals obtained with our methodology. No uncertainty is considered when the PFA is used. In contrast, using the ARRW methodology we can see that there is an empirically relevant distribution of IRFs conditional on the sign and zero restrictions holding. Additionally, note that for some variables − such as stock price, consumption, and hours worked − the IRFs obtained using the PFA are close to the 84-th confidence band. Hence, once again, we can see how the PFA picks a large response of stock prices and there is an upward bias in the response of consumption and hours worked. The fact that the Mountford and Uhlig (2009) methodology does not consider the distribution of IRFs lies behind the narrower confidence intervals

[11] We present the results only for identification 1. Similar results are obtained for the other two identification schemes and alternative point estimates.

that Beaudry, Nam and Wang (2011) report.



(a) Identification 1 (b) Identification 2 (c) Identification 3

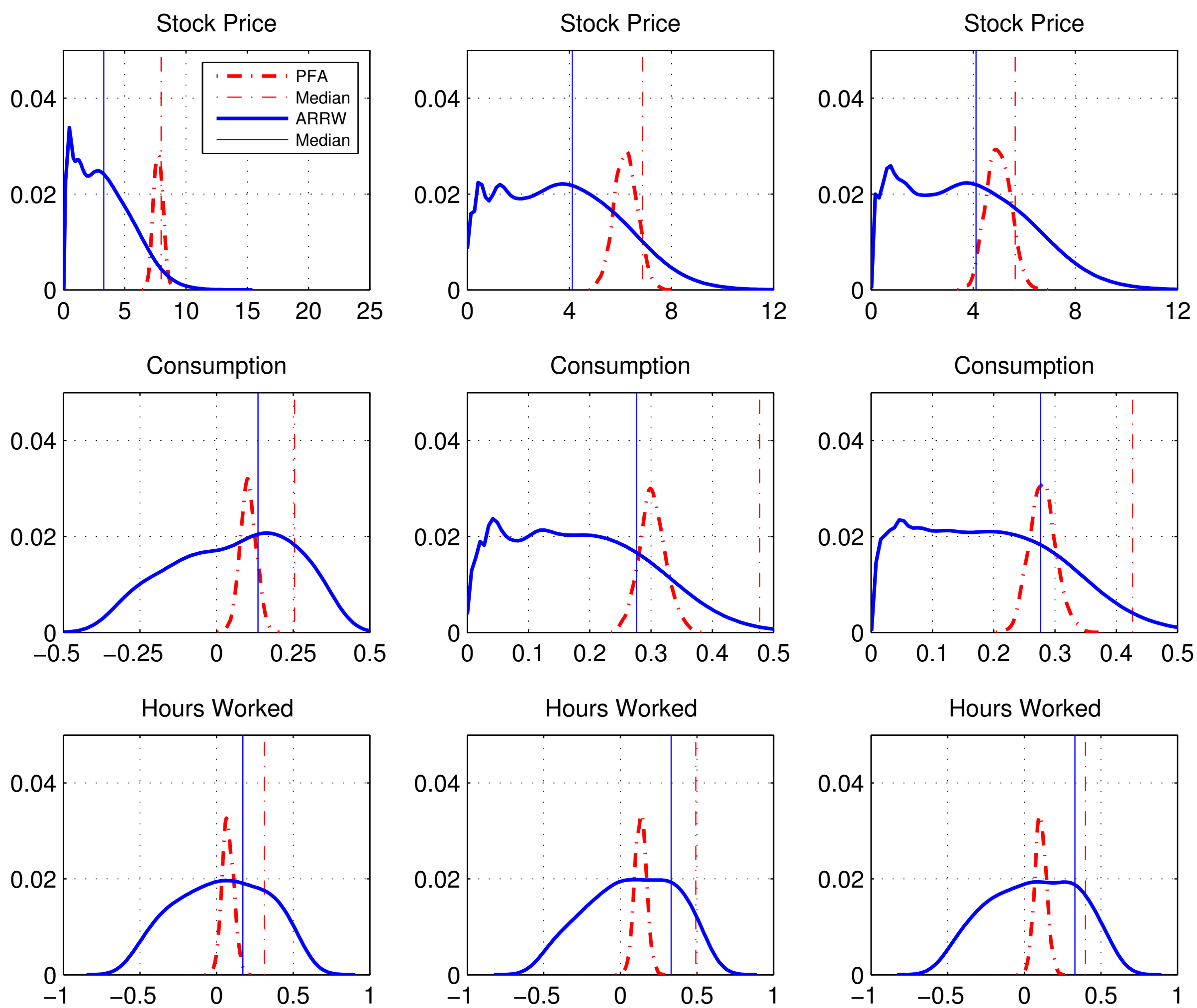
Figure 6: Distribution of IRFs with the ARRW Methodology vs. Single IRFs with the PFA: Five-Variable SVAR

Note: OLS PFA refers to the IRF obtained using the PFA and the OLS reduced-form estimates.

We can summarize our findings in Figure 7. Each column compares the posterior distributions of IRFs at horizon zero and the median IRFs for stock prices, consumption, and hours worked for each identification using both the PFA and the ARRW methodology.[12] The posterior distributions are approximated using a kernel smoothing function.[13] We have focused on the five-variable SVAR. Similar results apply to the seven-variable SVAR. Column 1 displays the results for Identification 1. Comparing the PFA and the ARRW methodology we reach the following conclusions. First, the posterior distribution of the IRFs of stock prices obtained using the PFA is centered around the right-hand tail of the distribution obtained using the ARRW methodology. The bias is even more clear looking at the median IRFs. Second, the PFA dramatically underestimates the variance of the posterior distribution of IRFs of stock prices. These two results were expected since the PFA maximizes the response of stock prices to optimism shocks in order to minimize the loss function. Since draws from the posterior distribution of the IRFs of consumption and hours worked are positively correlated with draws of the IRF of stock prices, we also observe artificially narrow and biased posterior distributions

[12]We report average median IRFs computed using horizons 0 to 3. The bias is larger using four periods and the results are emphasized. In any case, the bias persists even if we use only horizon 0.

[13]We use the MATLAB ksdensity function based on Bowman and Azzalini (1997).



(a) Identification 1 (b) Identification 2 (c) Identification 3

Figure 7: Density of IRFs at Horizon Zero and Median of IRFs from Horizons Zero to Three

of the IRFs for consumption and hours worked. Columns 2 and 3 show the results for Identifications 2 and 3. In both cases, we reach the same conclusion. The posterior distributions of IRFs at horizon zero for stock prices, consumption, and hours worked are artificially compressed and upwardly biased when computed using the PFA.

We now present additional evidence that the PFA introduces sign restrictions on variables that are seemingly unrestricted, thus not respecting the agnosticism of the identification scheme. Let's focus on the case of the five-variable SVAR. Similar results apply to the seven-variable SVAR. We begin with Identification 1. Table 2 compares the posterior probabilities that the IRFs for consumption and hours worked at horizon zero are negative for the two methodologies. The IRF of consumption is never negative when we use Mountford and Uhlig's (2009) methodology, while it is negative approximately 40 percent of the time under the ARRW methodology. The same is basically true for hours worked. These can also be seen by comparing the mean responses reported in Table 2. Therefore, Table 2 strongly supports the argument that the PFA does not respect the theoretical agnosticism of the identification scheme by introducing additional sign restrictions on these two variables. Another important result that can be found in the table is that the standard deviation of the IRFs at horizon zero is smaller under the PFA. This is, of course, related to the fact that confidence intervals are wider when using the ARRW methodology.

	The PFA			The ARRW Methodology		
	Mean	Std dev	Pr($\cdot < 0$)	Mean	Std dev	Pr($\cdot < 0$)
Consumption	0.1034	0.0260	0.0000	0.0532	0.1914	0.3980
Hours Worked	0.0736	0.0379	0.0250	0.0355	0.2891	0.4490

Table 2: Posterior Probabilities of Negative IRFs at Horizon Zero: Identification 1

Tables 3 and 4 repeat the exercise for Identifications 2 and 3. As we can see, the IRF of hours worked at horizon zero under Identification 2 is almost never negative using the PFA, while it is negative in approximately 40 percent of the draws using our methodology. Hence, the PFA also introduces additional sign restrictions on hours worked in the case of Identification 2.

	The PFA			The ARRW Methodology		
	Mean	Std dev	Pr($\cdot < 0$)	Mean	Std dev	Pr($\cdot < 0$)
Hours Worked	0.1325	0.0381	0.0010	0.0774	0.2793	0.3930

Table 3: Posterior Probabilities of Negative IRFs at Horizon Zero: Identification 2

	The PFA			The ARRW Methodology		
	Mean	Std dev	Pr($\cdot < 0$)	Mean	Std dev	Pr($\cdot < 0$)
Hours Worked	0.1325	0.0381	0.0010	0.0613	0.2833	0.4090

Table 4: Posterior Probabilities of Negative IRFs at Horizon Zero: Identification 3

6.3 FEV

The fact that the PFA adds sign restrictions on variables that are seemingly unrestricted and does not respect the agnosticism of the identification scheme is also reflected on the FEV reported in Beaudry, Nam and Wang (2011). Let us first analyze the SVAR with the smaller data set. We compare the contribution of optimism shocks to the FEV obtained using the ARRW methodology and the PFA. For ease of exposition, in Table 5 we focus on the contributions to the FEV at horizon 40.[14]

The ARRW Methodology			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.09	0.12	0.17
	[0.03 , 0.22]	[0.04 , 0.28]	[0.06 , 0.33]
Stock Price	0.16	0.26	0.31
	[0.03 , 0.47]	[0.07 , 0.58]	[0.09 , 0.62]
Consumption	0.17	0.28	0.40
	[0.02 , 0.49]	[0.06 , 0.59]	[0.13 , 0.66]
Real Interest Rate	0.18	0.20	0.23
	[0.07 , 0.39]	[0.08 , 0.40]	[0.09 , 0.44]
Hours Worked	0.18	0.27	0.29
	[0.04 , 0.48]	[0.07 , 0.55]	[0.07 , 0.57]
The PFA			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.17	0.22	0.28
	[0.08 , 0.30]	[0.10 , 0.37]	[0.14 , 0.43]
Stock Price	0.72	0.71	0.57
	[0.55 , 0.85]	[0.57 , 0.82]	[0.42 , 0.72]
Consumption	0.26	0.69	0.76
	[0.13 , 0.43]	[0.53 , 0.83]	[0.59 , 0.87]
Real Interest Rate	0.13	0.13	0.35
	[0.07 , 0.22]	[0.07 , 0.22]	[0.29 , 0.43]
Hours Worked	0.31	0.62	0.49
	[0.21 , 0.44]	[0.48 , 0.73]	[0.34 , 0.64]

Table 5: Share of FEV Attributable to Optimism Shocks at Horizon 40: Five-Variable SVAR

We first consider Identification 1. Using the ARRW methodology, the median contribution of optimism shocks to the FEV of consumption and hours worked is 17 and 18 percent, respectively. In contrast, using the PFA the median contributions are 26 and 31 percent, respectively. When Identification 2 is used, the median contribution of optimism shocks to the FEV of consumption and hours worked is 28 and 27 percent using our methodology, but it is equal to 71 and 62 percent using the PFA. Identification 3 yields the highest contribution of optimism shocks to the FEV of consumption and hours worked, 40 and 29 percent, respectively, when using our methodology. However, these values are moderate compared to the 76 and 49 percent that we found when using the PFA. Table 5 also

[14]Table 12 in the appendix reports the contributions to the FEV at additional horizons.

reports the 68 percent confidence intervals. As was the case with IRFs, the confidence intervals are much wider under the ARRW methodology. They are so wide that, in some cases, it is easy to argue that optimism shocks explain little of the FEV of most relevant variables.

The results for the seven-variable SVAR are reported in the appendix, Table 13 reports results at horizon 40 and Table 14 reports additional horizons. As expected, because of the increase in the number of variables, the contribution of optimism shocks declines relative to the case of five variables. For example, using the ARRW methodology the median contribution of optimism to the FEV of output is 12, 16, and 22 percent under Identifications 1, 2, and 3, respectively. In any case, these values are remarkably lower than the ones found using the PFA: 23, 59, and 60 percent, respectively. As before, confidence intervals are much wider when using the ARRW methodology.

Summarizing, using the ARRW methodology it is easy to conclude that optimism shocks explain a very small share of the FEV of any variable in the SVAR. This conclusion contrasts with the results obtained using the PFA. As was the case with the IRFs, since the PFA is not truly agnostic, it induces an upward bias in the median explained share of the FEV and artificially narrow confidence intervals. It is because of these two issues that Beaudry, Nam and Wang (2011) can claim that optimism shocks explain a large share of the FEV of some relevant variables. Once these two issues are corrected by the ARRW methodology, it is not possible to support such a claim. We have reported the results only for horizon 40 but Appendix 9.3 shows that these conclusions are true at any horizon.

6.3.1 Replicating the Penalty Function Approach Using the ARRW Methodolgy

In this subsection, we show that in the case of Identification 1 the PFA in Beaudry, Nam and Wang (2011) can be replicated using the ARRW methodology by considering some additional restrictions on the orthogonal matrix $\mathbf{Q}$.[15]

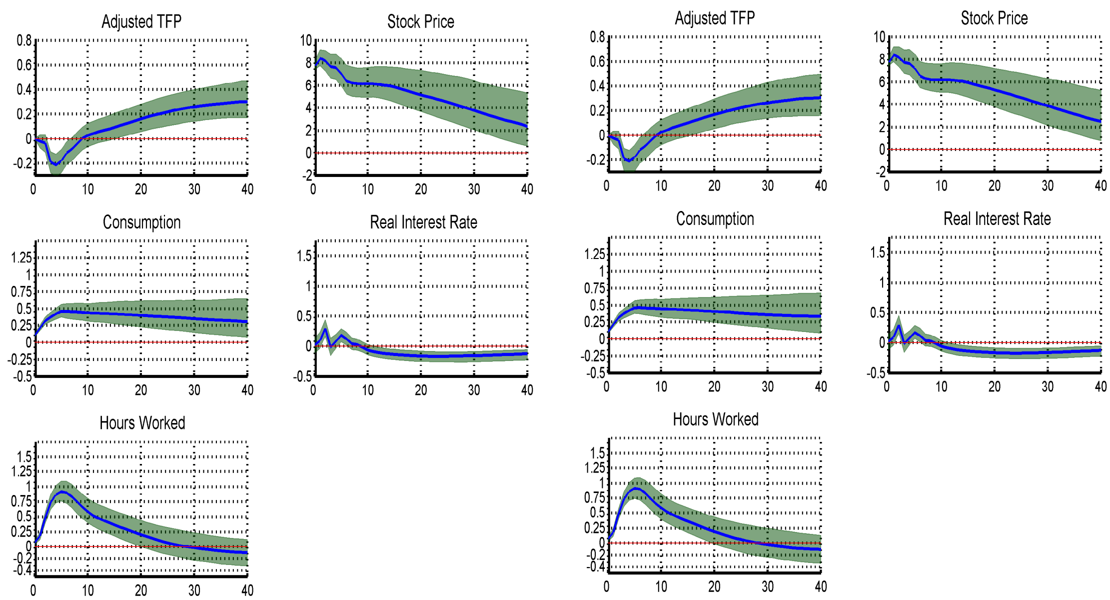
Consider Identification 1 and note that there exists a closed-form solution to the minimization problem embedded in the PFA, as shown in Subsection 5.3. Specifically, the penalty function is minimized when the first column of the orthogonal matrix equals $\bar{\mathbf{q}}_1^* = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}'$. Thus, we can replicate Beaudry, Nam and Wang (2011) using our methodology by imposing zero constraints on the first column of the orthogonal matrix $\mathbf{Q}$. In particular, let $f(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} \mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+) & \mathbf{I}_n \end{bmatrix}'$, where $\mathbf{I}_n$ allows us to put the zero constraints on the orthogonal matrix $\mathbf{Q}$. We define the matrices $\mathbf{S}$

[15] In fact, this is true for any identification that fulfills the conditions stated in Subsection 5.3.

and $\mathbf{Z}$ as follows

$$\mathbf{S}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \mathbf{Z}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The $\mathbf{S}$ matrix is identical to the one reported in Subsection 6.1, but the $\mathbf{Z}$ matrix has changed to reflect the additional restrictions on the first column of the orthogonal matrix $\mathbf{Q}$.



(a) The PFA (b) ARRW with Additional Restrictions

Figure 8: Replicating the PFA using the ARRW Methodology: Five-Variable SVAR

Panel (b) in Figure 8 plots the IRFs of TFP, stock price, consumption, the federal funds rate, and hours worked under Identification 1 using the ARRW methodology with the additional restrictions on the first column of the orthogonal matrix $\mathbf{Q}$. The results are identical to those reported in Panel (a) of Figure 3, which are reproduced in Panel (a) of Figure 8. Thus, we have shown that (under Identification 1), the PFA is a particular case of our methodology with additional restrictions on the first column of the orthogonal matrix $\mathbf{Q}$. This also applies to the seven-variable SVAR as shown in Figure 21 in the appendix. Since putting constraints on the orthogonal matrix $\mathbf{Q}$ is equivalent to imposing additional sign restrictions, the PFA does not respect the agnosticism of the identification scheme and it naturally produces artificially narrow confidence intervals and biased IRFs. Unfortunately, we can not show that the PFA is a particular case of our methodology for all identification schemes. Nevertheless, it should be clear that it always introduces additional restrictions (though they can not always be mapped into our methodology) that create artificially narrow confidence intervals and may introduce bias.

6.4 Computational Time

Our methodology is faster than the PFA. Table 6 reports the results for the case of optimism shocks using the five-variable SVAR. The PFA is approximately ten times slower than our methodology. Similar results can be found in the case of the seven-variable SVAR. Note that the computational time in Mountford and Uhlig's (2009) methodology is a function of the non-linear solver used to solve the minimization of the penalty function. We start the non-linear optimization from eight random starting points and then we pick the best one. Mountford and Uhlig (2009) follow a similar approach in order to avoid finding a local minimum. Of course, Identification 1 could have been solved faster with the PFA using the insights of Section 5.3.

	The PFA	The ARRW Methodology	Ratio
Identification 1	98.17	9.81	10.01
Identification 2	102.10	10.18	10.03
Identification 3	106.69	10.90	9.79

Table 6: Computational Time in Seconds: Five-variable SVAR

7 Fiscal Policy Shocks

Let us now focus on the second application. The aim of Mountford and Uhlig (2009) is to analyze the effects of fiscal policy using SVARs. They focus on unanticipated and anticipated fiscal policy shocks. They identify an unanticipated government revenue shock as well as an unanticipated government spending shock by imposing sign restrictions on the fiscal variables themselves as well as imposing orthogonality to a generic business cycle shock and a monetary policy shock. No sign restrictions are imposed on the responses of output, consumption, and investment to fiscal policy shocks. Thus, the identification remains agnostic with respect to the responses of these key variables of interest to fiscal policy shocks. The problem is, again, that the PFA is not really agnostic about the response of these variables.

They also consider three combined shocks (which are linear combinations of the unanticipated fiscal policy shocks): deficit-spending shocks, deficit-financed tax cut shocks, and balanced-budget spending shocks. These shocks are used to compare three fiscal policy scenarios of the same name. We proceed as in the case of the optimism shocks. We first replicate the results that Mountford and Uhlig (2009) obtain using the PFA. Then, we repeat their empirical work using our methodology, which is truly agnostic, to show how their main results significantly change.[16] Mountford and Uhlig (2009) conclude that deficit-financed tax cut shocks work best among the fiscal policy scenarios to improve GDP. In contrast, using our methodology we find no evidence to support such a claim. More generally, we find that it is very difficult to reach any conclusion about the effects of any of the three combined shocks (and therefore about the effects of any of the three associated fiscal policy scenarios) because of very wide confidence intervals around the median IRFs and the median fiscal multipliers associated with each scenario. As a consequence, any conclusion derived from Mountford and Uhlig's (2009) results relies on artificially narrow confidence intervals associated with the PFA.

Our findings also show that it is very hard to support any of Mountford and Uhlig's (2009) claims about the effects of the unanticipated fiscal policy shocks. Regarding unanticipated government revenue shocks, while Mountford and Uhlig (2009) report that GDP and consumption significantly decline in response to such shocks using the PFA, we find no support for such a claim using our methodology. The median IRFs of GDP, consumption, and non-residential investment to such shocks are negative using the PFA, but positive using our methodology. Furthermore, wide confidence intervals invalidate any conclusion. In the case of unanticipated government spending shocks, except for investment, the median IRFs from both methodologies are quite similar to each other, but wide confidence intervals make it very hard to reach any conclusion. As was the case with Beaudry, Nam and Wang (2011), we will argue that the problem behind the bias and the artificially narrow confidence intervals is that, by using the PFA, the authors are not being truly agnostic – additional sign restrictions are being imposed on output, consumption, and investment. Finally, for this application our methodology is also faster than the PFA.

7.1 Data and Identification Strategy

We use the same data set as Mountford and Uhlig (2009) in order to shed light on the implications of our methodology. The data set contains 10 U.S. variables at a quarterly frequency from 1955 to 2000:

[16] To keep our paper at a reasonable length, we omit the analysis of the anticipated fiscal policy shocks.

GDP, private consumption, total government spending, total government revenue, real wages, private non-residential investment, interest rate, adjusted reserves, producer price index of crude materials (PPIC), and GDP deflator. The identification strategy is described in Table 7. Appendix 9.2 gives details about the estimation procedure and the data set.

	Shocks			
	Business Cycle	Monetary Policy	Gov Revenue	Gov Spending
GDP	+			
Total Gov Spending				+
Total Gov Revenue	+		+	
Interest Rate		+		
Adjusted Reserves		−		
PPIC		−		
GDP Deflator		−		
Private Consumption	+			
Private Non-Res Investment	+			
Real Wages				

Table 7: Mountford and Uhlig (2009)

7.1.1 Unanticipated Fiscal Policy Shocks

We begin by describing the identification of the unanticipated fiscal policy shocks. Following Mountford and Uhlig (2009), we identify these shocks in three steps. In the first step, we identify a business cycle shock imposing four positive sign restrictions on GDP, private consumption, private non-residential investment, and total government revenue during four quarters − quarters zero to three − following the initial shock. In the second step, we identify a monetary policy shock imposing positive sign restrictions on interest rates, and negative sign restrictions on adjusted reserves, GDP deflator, and PPIC during the four quarters following the initial shock. In addition, the monetary policy shock is required to be orthogonal to the business cycle shock. In the third step, we identify the unanticipated fiscal shocks. The unanticipated government revenue shock is identified imposing positive sign restrictions on the response of total government revenue during the four quarters following the initial shock and requiring that the shock be orthogonal to the business cycle shock and the monetary policy shock. The unanticipated government spending shock is identified likewise. Hence, the identification is agnostic with respect to the responses of GDP, private consumption, and private non-residential investment to fiscal policy shocks. The PFA will not respect this agnosticism. Importantly, the unanticipated fiscal shocks are not required to be orthogonal between them.

As in the case of optimism shocks, it is instructive to map the identification strategy to our methodology. The function $f(\mathbf{A}_0, \mathbf{A}_+)$ and the matrices **S**s necessary to apply our methodology are

$$f(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} \mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_1(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_2(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_3(\mathbf{A}_0, \mathbf{A}_+) \end{bmatrix} \text{ and } \mathbf{S}_j = \begin{bmatrix} \mathbf{S}_{j0} & \mathbf{0}_{m(j),n} & \mathbf{0}_{m(j),n} & \mathbf{0}_{m(j),n} \\ \mathbf{0}_{m(j),n} & \mathbf{S}_{j1} & \mathbf{0}_{m(j),n} & \mathbf{0}_{m(j),n} \\ \mathbf{0}_{m(j),n} & \mathbf{0}_{m(j),n} & \mathbf{S}_{j2} & \mathbf{0}_{m(j),n} \\ \mathbf{0}_{m(j),n} & \mathbf{0}_{m(j),n} & \mathbf{0}_{m(j),n} & \mathbf{S}_{j3} \end{bmatrix} \text{ for } j = 1, \ldots, 4.$$

where $\mathbf{0}_{m(j),n}$ is an $m(j)$ times n matrix of zeros and $m(j) = 4$ if $j = 1$ or 2 and $m(j) = 1$ otherwise, and

$$\mathbf{S}_{1t} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \mathbf{S}_{2t} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{S}_{3t} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{and } \mathbf{S}_{4t} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

for $t = 0, \ldots, 3$. In addition, we need to impose the orthogonality conditions between the shocks as described above. This is straightforward from the discussion in Section 3.2. The only challenge is that Mountford and Uhlig (2009) do not require orthogonality between the unanticipated fiscal shocks; thus, we need to accommodate our methodology to study this case. We accomplish this by requiring that the unanticipated government revenue (spending) shock, associated with the third (fourth) column of **Q**, be orthogonal to the first and second columns of **Q**, associated with the business cycle and monetary policy shocks, respectively, without restricting other columns of **Q**. For example, in the case of the unanticipated government spending shock, this is accomplished by modifying $\mathbf{R}_4(\mathbf{A}_0, \mathbf{A}_+)$ in Theorem 4 to be equal to

$$\mathbf{R}_4(\mathbf{A}_0, \mathbf{A}_+) = \mathbf{Q}_2'.$$

The reader should note that a direct application of Theorem 4 would make $\mathbf{R}_4(\mathbf{A}_0, \mathbf{A}_+)$ depend on $\mathbf{Q}_3'$ instead of $\mathbf{Q}_2'$. There are no zero restrictions; hence, we do not need to define any **Z** matrix.

Next, we describe the three linear combinations of the unanticipated fiscal policy shocks that are used to study the three fiscal policy scenarios.

7.1.2 Fiscal Policy Scenarios

The deficit-spending shocks (used to study the deficit-spending scenario) are a sequence of unanticipated fiscal policy shocks where total government spending rises by 1 percent and total government revenue remains unchanged during the four quarters following the initial shock. The deficit-financed tax cut shocks (used to study the deficit-financed tax cuts scenario) are a sequence of unanticipated fiscal policy shocks where total government spending remains unchanged and total government revenue falls by 1 percent during the four quarters following the initial shock. The balanced-budget spending shocks (used to study the balanced-budget spending scenario) are a sequence of unanticipated fiscal policy shocks where total government spending rises by 1 percent and total government revenue rises by 1.28 percent during the four quarters following the initial shock.[17] Let $(a_{s,t}, b_{s,t})$ for $t = 0, \ldots, 4$ and $s \in \{\text{DS,DTC,BB}\}$ denote the weights to be used in the linear combination of the unanticipated fiscal policy shocks to get the deficit-spending shocks (DS), the deficit-financed tax cut shocks (DTC), and the balanced-budget spending shocks (BB), respectively. For example, let us consider the case of a deficit-spending shock. Following Mountford and Uhlig (2009) we solve for such weights by solving the linear system of equations described below

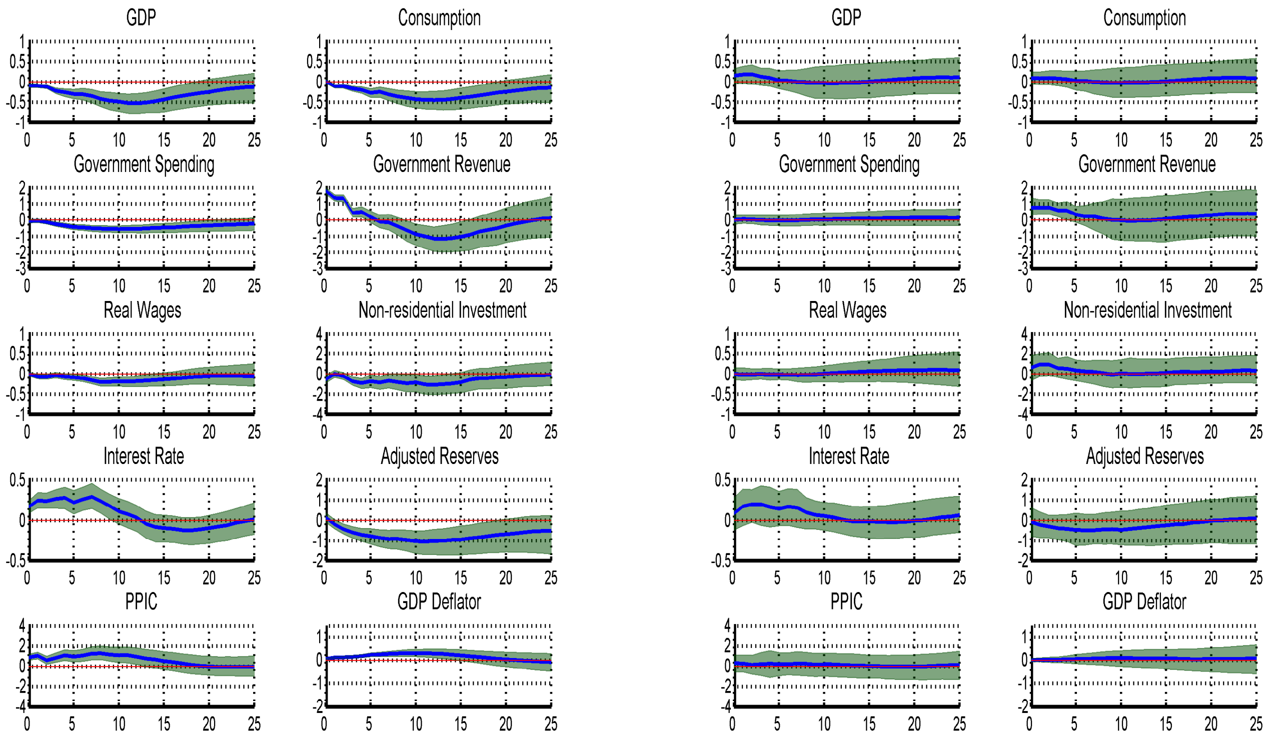
$$
\begin{aligned}
0.01 &= \sum_{t=0}^{\tau} \left(\mathbf{e}_{GS}' \mathbf{L}_{\tau-t}\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_4 a_{DS,t} + \mathbf{e}_{GS}' \mathbf{L}_{\tau-t}\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_3 b_{DS,t}\right) \quad \text{for } \tau = 0, \ldots, 3 \\
0 &= \sum_{t=0}^{\tau} \left(\mathbf{e}_{GR}' \mathbf{L}_{\tau-t}\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_4 a_{DS,t} + \mathbf{e}_{GR}' \mathbf{L}_{\tau-t}\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_3 b_{DS,t}\right) \quad \text{for } \tau = 0, \ldots, 3
\end{aligned}
$$

where $\mathbf{e}_{GS}$ ($\mathbf{e}_{GR}$) is a unit vector with a one at the entry associated with total government spending (government revenue) in the SVAR and zeros otherwise. Then, we can use the weights $(a_{DS,t}, b_{DS,t})$ for $t = 0, \ldots, 4$ to build the column vector associated with the deficit-spending shocks as $\mathbf{q}_{DS} = \mathbf{q}_4 a_{DS,t} + \mathbf{q}_3 b_{DS,t}$ for $t = 0, \ldots, 3$. In a similar fashion, we can construct weights for the other two combined shocks and obtain the column vectors $\mathbf{q}_{DTC}$ and $\mathbf{q}_{BB}$.

[17] The percentage increase in total government revenue is higher than the percentage increase in total government spending so that total government revenues and total government spending increase by the same amount during the four quarters following the initial shock.

7.2 IRFs to Unanticipated Fiscal Policy Shocks

Let us begin examining the IRFs. We first show the replications of the IRFs reported by Mountford and Uhlig (2009) using the PFA and then we analyze how the results change once we use our methodology. To save space, we do not report results on either business cycle or monetary policy shocks. Also, we refer to private consumption and to private non-residential investment as consumption and investment, respectively. Finally, and also because of space considerations, we just concentrate on the responses of GDP, consumption, and investment.

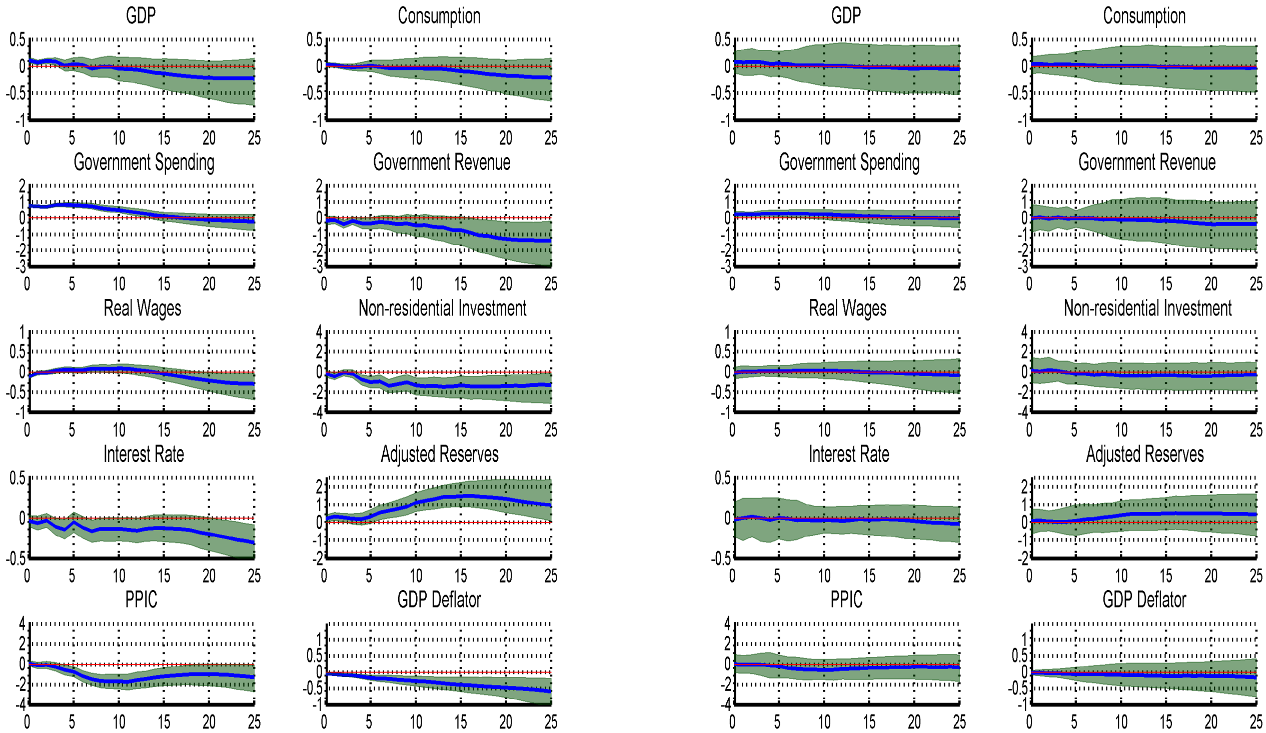


(a) The PFA

(b) The ARRW Methodology

Figure 9: IRFs to an Unanticipated Government Revenue Shock

Figure 9 plots the IRFs to an unanticipated government revenue shock. Panel (a) replicates the results reported in Figure 4 in Mountford and Uhlig (2009). This panel shows that using the PFA, the median IRFs of GDP, consumption, and investment are negative. Furthermore, the 68 percent confidence intervals are narrow and do not contain zero. Therefore, one can easily conclude that unanticipated government revenue shocks cause a decline in economic activity. In contrast, once we use our methodology (see Panel b), the sign of the median IRFs changes and the 68 percent confidence intervals are much wider. Thus, as in the case of optimism shocks, the PFA creates bias and artificially narrow confidence intervals. In addition, the PFA chooses a very large median response of total government revenue.



(a) The PFA (b) The ARRW Methodolgy

Figure 10: IRFs to an Unanticipated Government Spending Shock

Figure 10 plots the IRFs to an unanticipated government spending shock. Panel (a) replicates Figure 7 in Mountford and Uhlig (2009) and it shows the median IRFs of GDP, consumption, and investment. Using the penalty approach, the median response of GDP changes from positive to negative in period 10, the response of consumption changes from zero to negative around period 12, and the response of investment is always negative. Although less than in the case of the unanticipated government revenue shocks, the median IRFs change when we use the ARRW methodology (see Panel b). The changes are important for investment (whose response is positive for 5 periods). Nevertheless, the confidence intervals are wider under our methodology and they contain zero. Analogously to the case of unanticipated government revenue shocks, the PFA introduces downward bias (at least for several quarters) in the response of investment to unanticipated government spending shocks and creates artificially narrow confidence intervals. It is also the case that the PFA picks a large response of total government spending to minimize the loss function.

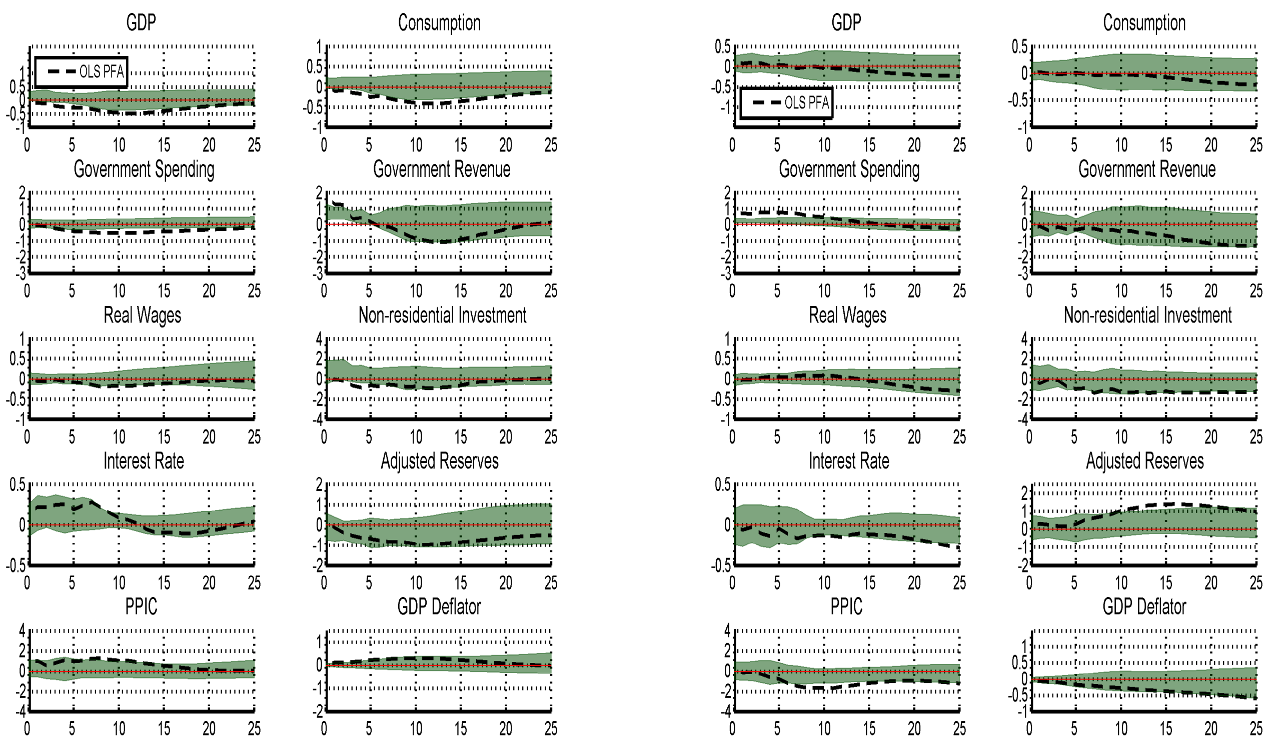
Summarizing, using the ARRW methodology we observe important changes in the median IRFs to unanticipated fiscal policy shocks with respect to the results reported in Mountford and Uhlig (2009). In what follows we show that the lack of agnosticism of the PFA introduces a bias in the response of some variables and delivers confidence intervals that are artificially narrow. The strong results reported in Mountford and Uhlig (2009) are because of the bias and the artificially narrow confidence intervals. The truly agnostic ARRW methodology amends these two problems and the results disappear.

7.2.1 Understanding the Bias and the Artificially Narrow Confidence Intervals

As before, we now shed some light on the biases and artificially narrow confidence intervals. Let us begin with the biases related to unanticipated government revenue shock. The PFA selects a large response of total government revenue to this shock in order to minimize the loss function. By selecting a large response of total government revenue, the PFA is implicitly forcing a negative response of GDP, consumption, and investment because their IRFs are negatively correlated with the IRF of total government revenue. The correlations of the IRF of total government revenue to unanticipated government revenue shocks at horizon zero with the IRFs of GDP, consumption, and investment to the same shock and horizon are -0.12, -0.11, and -0.01, respectively.

In the case of unanticipated government spending shocks the PFA also selects a large response of total government spending in order to minimize the loss function. By choosing a large response, the PFA is implicitly forcing a negative response of investment because its IRF is negatively correlated with the response of total government spending. The correlation between the IRF of total government spending to unanticipated government spending shocks at horizon zero with the IRF of investment to the same shock and horizon is -0.27. Additionally, the PFA is over-estimating the response of GDP and consumption to an unanticipated government spending shock because the correlations of the IRFs of GDP and consumption with the IRF of government spending are 0.49 and 0.28, respectively. By inducing these correlated responses the PFA is introducing additional sign restrictions on these variables and, thus, not respecting the agnosticism of the identification scheme.

Let us now focus on the artificially narrow confidence intervals generated by the PFA. We have shown that for each draw from the posterior distribution of the reduced-form parameters, there is a distribution of IRFs conditional on the sign and zero restrictions holding, and that the PFA selects a single orthogonal matrix instead of drawing from the conditional uniform distribution. What are the consequences of this? Not surprisingly at this juncture, the consequence is artificially narrow confidence intervals. To see this, we first examine the IRFs to an unanticipated government revenue shock evaluated at a value of reduced-form parameters given by the OLS point estimates. The dashed lines in Panel (a) of Figure 11 show the IRFs obtained using the Cholesky decomposition and the OLS point estimates. The shadow areas in Panel (a) of Figure 11 correspond to the 68% confidence bands of the distribution of IRFs at the OLS point estimate of the reduced-form parameters. These confidence bands characterize the distribution of IRFs consistent with the sign and zero restrictions, and they are constructed using the Cholesky decomposition, the OLS point estimate of the reduced-form parameters, and several draws from the conditional uniform distribution of orthogonal matrix $\mathbf{Q}$. The distribution of the IRFs is in sharp contrast with the single IRFs obtained using the PFA evaluated at the same value of the reduced-form parameters. Panel (b) Figure 11 shows that the same happens in the case of the unanticipated government spending shock.



(a) Unanticipated Government Revenue Shock (b) Unanticipated Government Spending Shock

Figure 11: Distribution of IRFs with the ARRW Methodology vs. Single IRFs with the PFA

Note: OLS PFA refers to the IRF obtained using the PFA and the OLS reduced-form estimates.

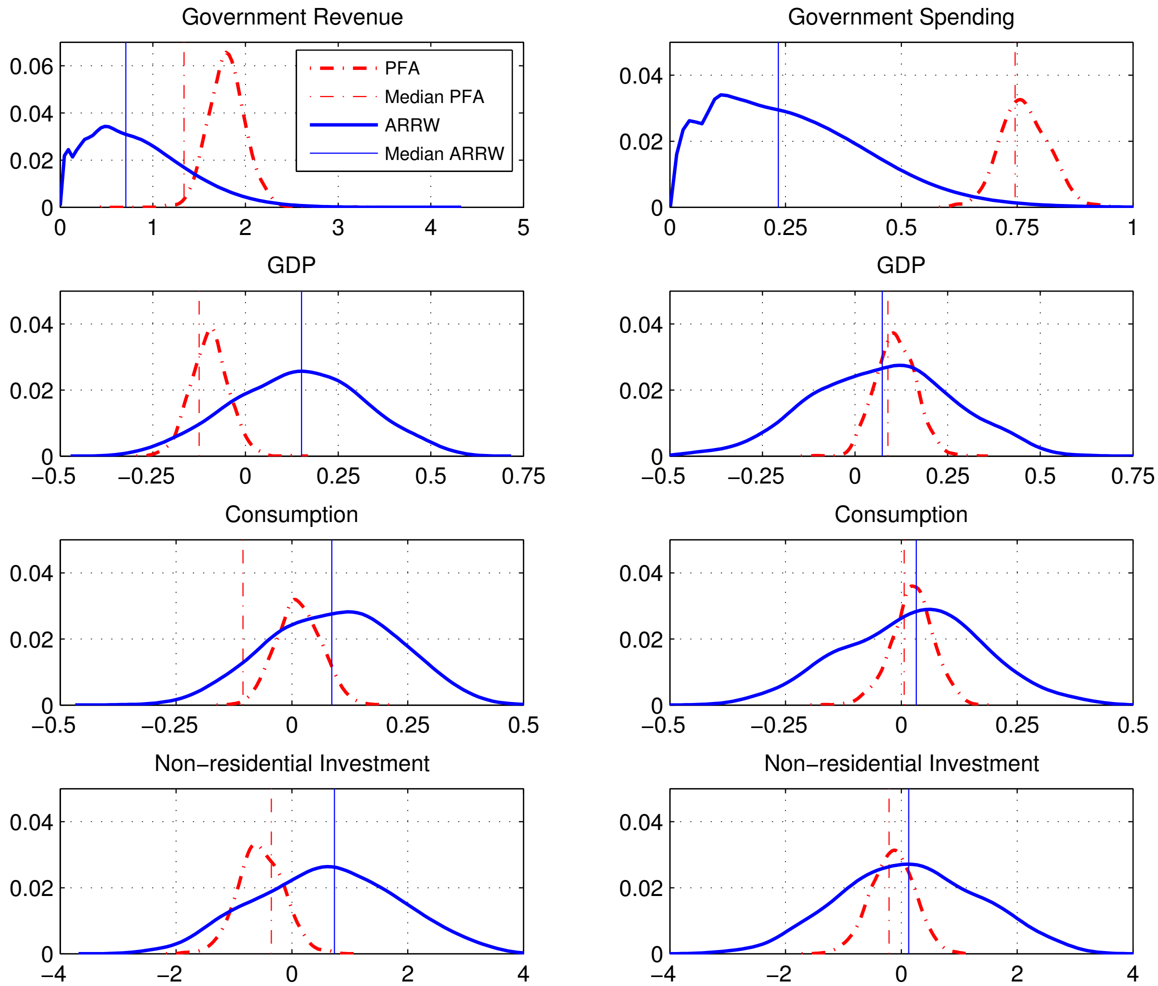
We summarize our findings in Figure 12. The first column compares the posterior distributions of IRFs at horizon zero and the median IRFs for total government revenue, GDP, consumption, and investment to an unanticipated government revenue shock using both the PFA and the ARRW methodology.[18] The posterior distributions are approximated using a kernel smoothing function. Comparing the PFA and the ARRW methodology, we reach the following conclusions. First, the posterior distribution of IRFs of total government revenue obtained using the PFA is centered around the right-hand tail of the distribution obtained using the ARRW methodology. The bias is even more clear, when we look at the median IRFs. Second, the PFA dramatically underestimates the variance of the posterior distribution of IRFs of total government revenue. These two results were expected since the PFA is maximizing the response of total government revenue in order to minimize the loss function. We also observe very narrow and downwardly biased posterior distributions of the IRFs for GDP, consumption, and investment. Column 2 does the same for the IRFs of total government spending, GDP, consumption, and investment to an unanticipated government spending shock. Again, the posterior distribution of IRFs of total government spending obtained using the PFA is centered around the right-hand tail of the distribution obtained using the ARRW methodology. We also observe very narrow and downwardly biased posterior distributions of the IRFs for investment, and there is an upward bias for the response of GDP.

We now present additional evidence that the PFA is not truly agnostic because it introduces additional sign restrictions on variables that are seemingly unrestricted. Table 8 compares the posterior probabilities that the IRFs for GDP, consumption, and investment to an unanticipated government revenue shock are negative in at least one of the first four horizons. The IRFs of the three variables are almost always negative when using the Mountford and Uhlig (2009) methodology, while it is negative only about 25 percent of the time when using the ARRW methodology. Hence, the PFA imposes additional sign restrictions on these three variables. Table 9 shows that the PFA also distorts the posterior probabilities of negative IRFs for GDP and investment in response to an unanticipated government expenditure shock .

7.3 Three Fiscal Policy Scenarios

Equipped with the unanticipated fiscal policy shocks, we can analyze the three fiscal policy scenarios. We first study how the IRFs associated with the deficit-spending, the deficit-financed tax cuts, and

[18] We report average median IRFs computed using horizons 0 to 3. The bias is larger using four periods and the results are emphasized. In any case, the bias basically persists using only horizon 0.



(a) Gov Revenue Shock

(b) Gov Spending Shock

Figure 12: Density of IRFs at Horizon Zero and Median of IRFs from Horizons Zero to Three

	The PFA			The ARRW Methodology		
	Mean	Std dev	$\Pr(\cdot < 0)$	Mean	Std dev	$\Pr(\cdot < 0)$
GDP	-0.1367	0.0741	0.9882	0.1528	0.2046	0.2400
Consumption	-0.0960	0.0854	0.8415	0.0869	0.1582	0.2990
Non-res Investment	-0.3658	0.4704	0.7752	0.1528	1.1037	0.2522

Table 8: Posterior Probabilities of Negative IRFs in at Least One of the First Four Horizons: Unanticipated Government Revenue Shock

the balanced-budget spending shocks change when using the ARRW methodology with respect to the results reported in Mountford and Uhlig (2009). Second, we do the same for the fiscal multipliers. As has been the case with unanticipated fiscal policy shocks, we just report results for GDP, consumption,

	The PFA			The ARRW Methodology		
	Mean	Std dev	Pr(· < 0)	Mean	Std dev	Pr(· < 0)
GDP	0.0882	0.0573	0.0565	0.0708	0.2187	0.3795
Consumption	0.0036	0.0523	0.4480	0.0270	0.1681	0.4203
Non-res Investment	-0.2157	0.3964	0.7073	0.0708	1.1502	0.4635

Table 9: Posterior Probabilities of Negative IRFs in at Least One of the First Four Horizons: Unanticipated Government Spending Shock

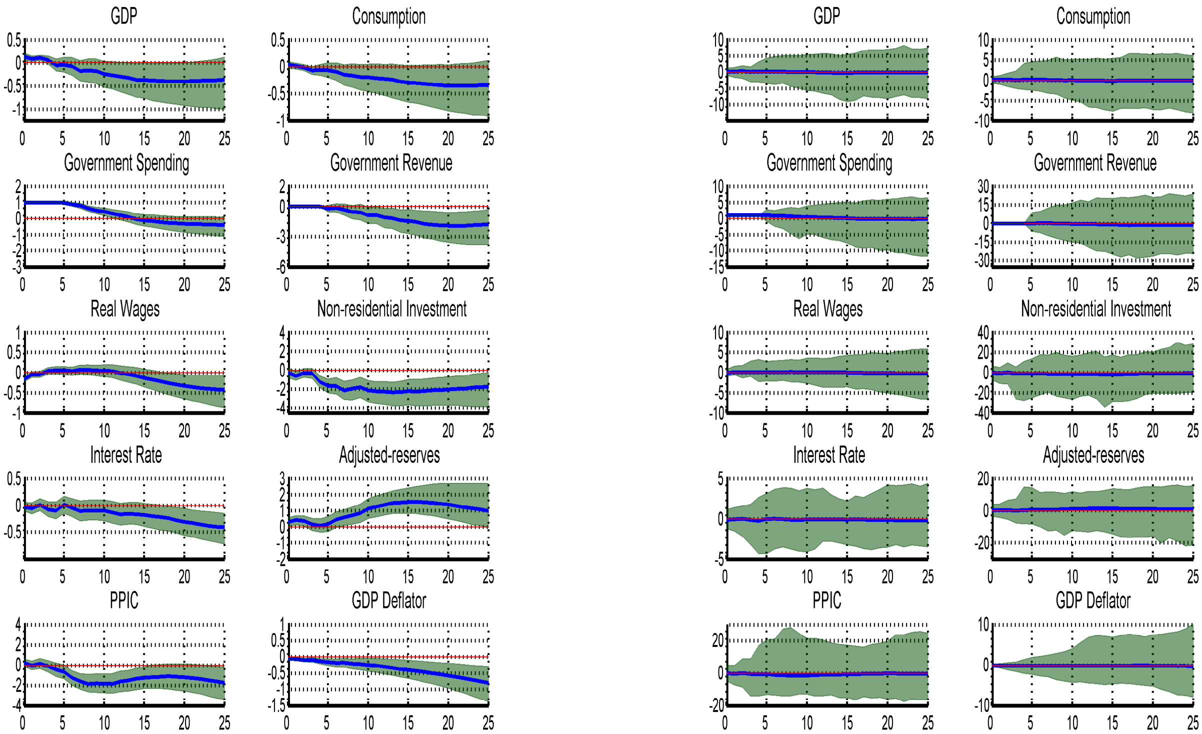
and investment.

7.3.1 IRFs

We begin with the deficit-spending shocks. Panel (a) in Figure 13 replicates the results in Figure 10 in Mountford and Uhlig (2009). Using the PFA, one could conclude that deficit-spending shocks produce a drop in GDP (after a few periods of a small increase), consumption, and investment (although the drop is statistically significant only for investment). However, once we use our methodology, these results disappear. There is a very wide range of IRFs that are consistent with these shocks, making it very hard to say anything about the effects of deficit-spending shocks. In most cases the confidence intervals reported using the ARRW methodology are at least five times bigger than the confidence intervals reported using the PFA. This means that, once we combine the unanticipated fiscal policy shocks, the confidence intervals are compounded and become even wider than before. Mountford and Uhlig's (2009) conclusions are based on artificially narrow confidence intervals.

Next, we study deficit-financed tax cut shocks. Panel (a) in Figure 14 replicates the results in Figure 11 in Mountford and Uhlig (2009). We can see that the median IRFs of GDP, consumption, real wages, and investment are positive and the tight 68 percent confidence intervals do not contain zero. Mountford and Uhlig (2009) use these results to claim that deficit-financed tax cut shocks work best to improve economic activity. On the contrary, the IRFs computed using the ARRW methodology do not provide evidence to support these findings. The median responses are negative for the first few periods but, again, very wide confidence intervals make the interpretation of the median IRFs very hard. The upward bias and the artificially narrow confidence intervals obtained using the PFA are behind Mountford and Uhlig's (2009) conclusions.

Finally, Mountford and Uhlig (2009) study a balanced-budget spending scenario. Panel (a) in Figure 15 replicates the results reported in Figure 12 in Mountford and Uhlig (2009). As can be seen, the median IRFs of GDP, consumption, and investment are (almost always) negative and the narrow



(a) The PFA

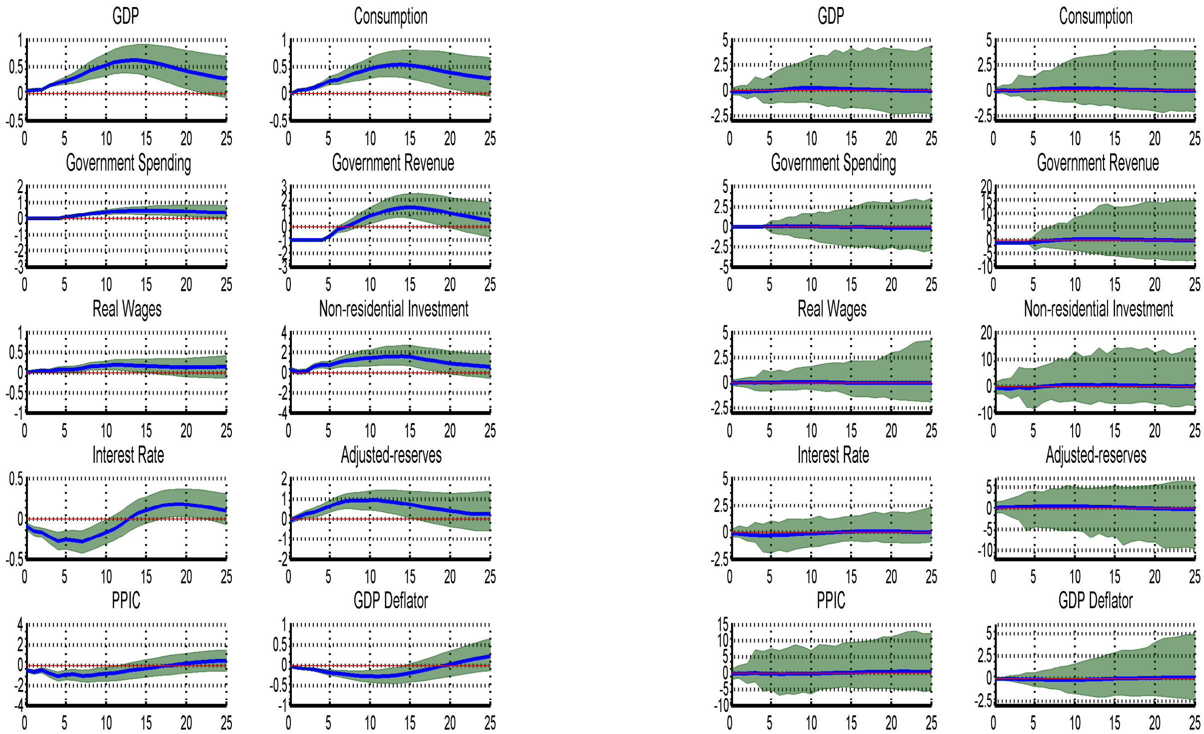
(b) The ARRW Methodology

Figure 13: IRFs to a Deficit-Spending Policy Shock

68 percent confidence intervals do not contain zero. Again, once we consider our methodology, there is no evidence to support these results. The median responses are positive for the first few periods, but the confidence intervals are so wide that it is hard to conclude anything at all. Downward bias and artificially narrow confidence intervals are behind any conclusions implied by Mountford and Uhlig (2009).

Our methodology paints a completely different picture than the one reported in Mountford and Uhlig (2009). The biases that we find are very hard to interpret because these shocks are linear combinations of shocks that are already biased. The PFA's lack of agnosticism is mostly reflected in extremely narrow confidence intervals.

The comparison between scenarios becomes even harder once we consider the cumulative discounted IRFs to either deficit-spending or deficit-financed tax cut shocks. The cumulative discounted IRFs at horizon τ of variable y to the combined shock s is $\sum_{t=0}^{\tau}(1+i)^{-t}\mathbf{e}_y'\mathbf{L}_t\left(\mathbf{A}_0,\mathbf{A}_+\right)\mathbf{q}_s$, where $\mathbf{e}_y$ is a unit vector that selects the IRF of the variable under analysis, $\mathbf{q}_s$ defines either the deficit-spending or deficit-financed tax cut shock depending on the value of $s \in \{\text{DS,DTC}\}$, and i denotes the average

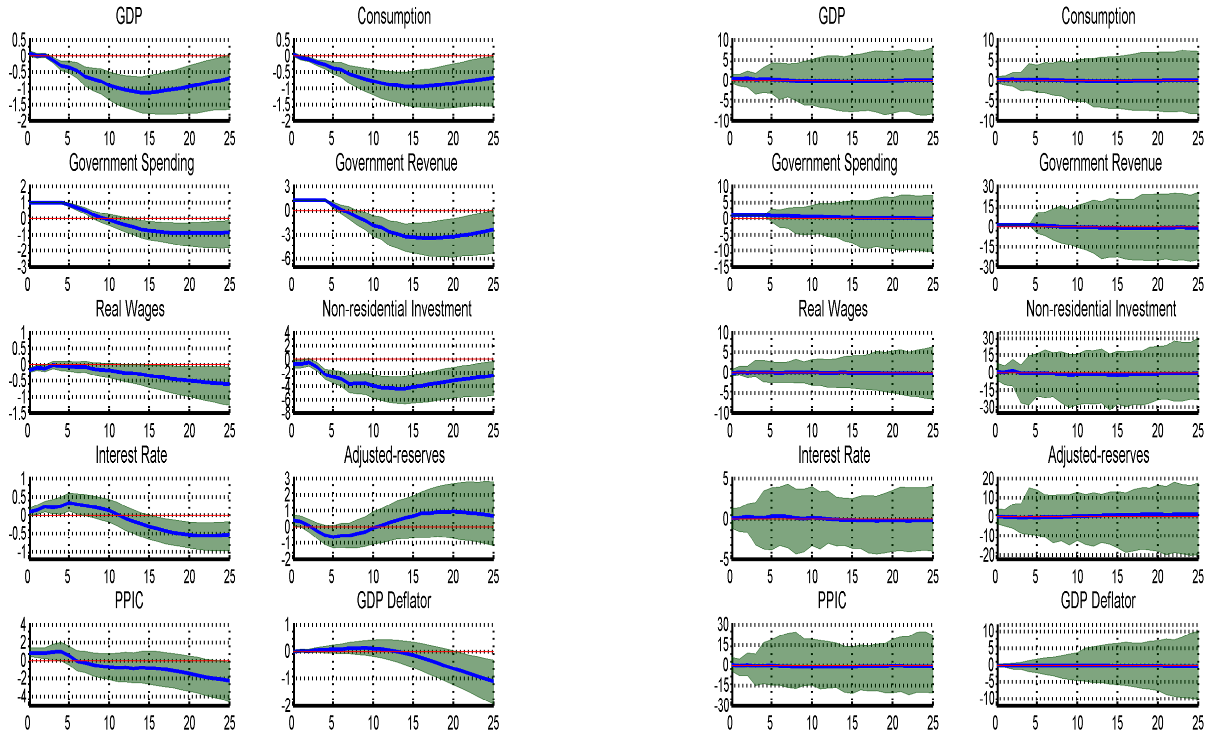


(a) The PFA (b) The ARRW Methodology

Figure 14: IRFs to a Deficit-Financed Tax Cut Shock

real interest rate over the sample. The real interest rate is computed as the difference between the federal funds rate and the inflation rate implied by the GDP deflator and, in our sample, equals 2.51 percent − annualized.

Panel (a) in Figure 16 replicates the results reported in Figure 13 in Mountford and Uhlig (2009). The panel shows that the median cumulative discounted IRF of GDP to a deficit-spending shock becomes negative after a few periods and in the case of a deficit-financed tax cut shock is always positive. Moreover, the 68 percent confidence intervals associated with the shocks are narrow and in the case of deficit-financed tax cut shocks do not contain zero. Based on this evidence, Mountford and Uhlig (2009) conclude that a deficit-financed tax cut scenario works best to improve GDP. Unfortunately, once we use the truly agnostic ARRW methodology, this result also disappears. The median cumulative discounted IRF of GDP to a deficit-spending shock is positive during 25 periods and it is negative for 10 periods for the case of deficit-financed tax cut shocks. As before, these biases are very hard to interpret because the deficit-spending and the deficit-financed tax cut shocks are linear combinations of shocks that are already biased. In any case, the correctly computed 68 percent confidence intervals



(a) The PFA (b) The ARRW Methodolgoy

Figure 15: IRFs to a Balanced-Budget Shock

contain zero for both IRFs and are at least five times larger than the ones reported using the PFA. Again, the PFA's lack of agnosticism is mostly reflected in extremely narrow confidence intervals.

7.3.2 Fiscal Multipliers

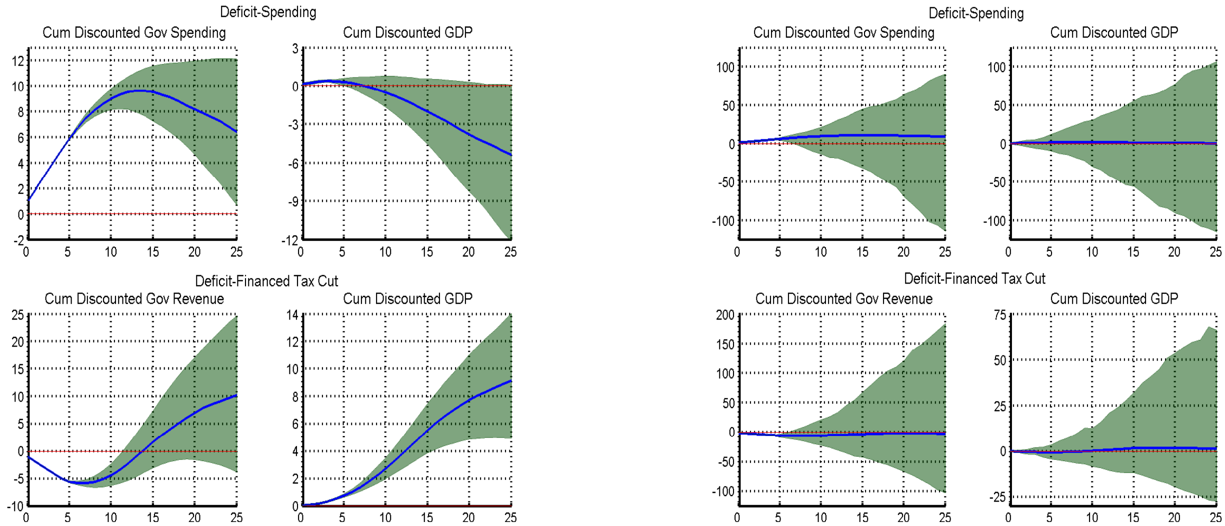
In addition to the IRF analysis, Mountford and Uhlig (2009) compute fiscal multipliers to compare the effects of deficit-spending shocks and deficit-financed tax cut shocks. Specifically, they compute the present value multipliers at horizon τ of the combined shock s on variable y

$$\frac{\sum_{t=0}^{\tau}(1+i)^{-t}\mathbf{e}_y'\mathbf{L}_t\left(\mathbf{A}_0,\mathbf{A}_+\right)\mathbf{q}_s}{\sum_{t=0}^{\tau}(1+i)^{-t}\mathbf{e}_f'\upsilon_{\mathbf{s},\mathbf{y}}\mathbf{L}_t\left(\mathbf{A}_0,\mathbf{A}_+\right)\mathbf{q}_s}\frac{1}{(f/GDP)}$$

and the impact multipliers at horizon τ of the combined shock s on variable y

$$\frac{\mathbf{e}_y'\mathbf{L}_\tau\left(\mathbf{A}_0,\mathbf{A}_+\right)\mathbf{q}_s}{\mathbf{e}_f'\upsilon_{\mathbf{s},\mathbf{y}}\mathbf{L}_0\left(\mathbf{A}_0,\mathbf{A}_+\right)\mathbf{q}_s}\frac{1}{(f/GDP)},$$

where $\mathbf{e}_y$ is a unit vector that selects the IRF of the variable under analysis, $\mathbf{e}_f$ is a unit vector



(a) The PFA

(b) The ARRW Methodology

Figure 16: Cumulative IRFs to Deficit-Spending and Deficit-Financed Tax Cut Shocks

that selects the IRF of the fiscal variable (total government revenue or total government spending), (f/GDP) denotes the average share of the selected fiscal variable in GDP over the sample, and $\upsilon_{\mathbf{s},\mathbf{y}}$ is equal to -1 if $s = DTC$ and $y \neq GDP$ and equals 1 otherwise. The indicator variable $\upsilon_{\mathbf{s},\mathbf{y}}$ is a normalization so that the multiplier of the deficit-financed tax cut shock can be interpreted as the increase in GDP in response to a decrease in total government revenue.

In the case of present value multipliers, y is GDP and f is total government spending (revenue) when $s =$ DS (DTC). In the case of impact multipliers, y can be GDP, total government spending, or total government revenue and f is total government spending (revenue) when $s =$ DS (DTC).

Table 10 reports the median multipliers. We also report the 68 percent confidence intervals. Also, quarter t in the table corresponds to horizon $t-1$ in the above formulas.

Panel (a) shows the present value multipliers associated with deficit-spending and deficit-financed tax cut shocks. The upper block of this panel replicates the results reported in Table 2 in Mountford and Uhlig (2009). The bottom block of this panel reports the results obtained using our methodology. Using the PFA, the median multipliers associated with deficit-financed tax cut shocks are positive for at least 12 quarters. In contrast, these median multipliers are negative at all horizons using the ARRW methodology. When we consider deficit-spending shocks, we find that while the median multipliers are negative after 12 quarters using the PFA, they are positive during 20 quarters using the ARRW methodology. Also, the median multipliers associated with the deficit-spending shocks are larger. At

	1 qrt	4 qrts	8 qrts	12 qrts	20 qrts	Max
(a) Present Value Multipliers						
The Penalty Function Approach						
DTC	0.30	0.53	1.49	4.75	-4.81	4.75
	[0.13 , 0.50]	[0.44 , 0.63]	[1.09 , 2.02]	[1.55 , 14.55]	[-11.35 , 2.35]	(qrt 12)
DS	0.62	0.45	0.02	-0.42	-1.94	0.62
	[0.26 , 0.99]	[0.33 , 0.57]	[-0.37 , 0.42]	[-1.21 , 0.37]	[-5.94 , 0.20]	(qrt 1)
The ARRW Methodology						
DTC	-0.75	-0.73	-1.04	-1.57	-1.83	-0.73
	[-2.88 , 1.19]	[-2.80 , 3.00]	[-4.18 , 2.30]	[-4.77 , 1.84]	[-3.49 , 0.12]	(qrt 4)
DS	1.12	1.03	1.37	1.59	2.99	4.03
	[-4.33 , 5.67]	[-7.68 , 5.86]	[-4.16 , 8.72]	[-5.32 , 8.07]	[-4.34 , 9.01]	(qrt 26)
(b) Impact Multipliers: DTC						
The Penalty Function Approach						
GDP	0.30	0.95	2.21	3.60	2.84	3.82
	[0.13 , 0.50]	[0.77 , 1.13]	[1.43 , 3.06]	[2.31 , 5.10]	[0.83 , 5.15]	(qrt 14)
Total Gov Revenue	-1.00	-1.00	-0.02	0.98	1.14	
	[-1.00 , -1.00]	[-1.00 , -1.00]	[-0.37 , 0.32]	[0.38 , 1.74]	[0.04 , 2.28]	
Total Gov Spending	0.00	0.00	0.23	0.44	0.47	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[0.15 , 0.32]	[0.25 , 0.62]	[0.15 , 0.83]	
The ARRW Methodology						
GDP	-0.75	-0.42	0.68	1.39	0.21	1.54
	[-2.88 , 1.19]	[-3.47 , 3.00]	[-7.35 , 2.30]	[-8.15 , 1.84]	[-12.55 , 26.18]	(qrt 11)
Total Gov Revenue	-1.00	-1.00	-0.19	0.37	0.10	
	[-1.00 , -1.00]	[-1.00 , -1.00]	[-3.07 , 4.99]	[-4.34 , 9.01]	[-7.10 , 14.33]	
Total Gov Spending	0.00	0.00	0.06	0.00	-0.12	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[-0.99 , 1.15]	[-1.91 , 1.87]	[-2.85 , 2.94]	
(c) Impact Multipliers: DS						
The Penalty Function Approach						
GDP	0.62	0.28	-0.84	-1.29	-1.93	0.62
	[0.26 , 0.99]	[-0.12 , 0.70]	[-1.87 , 0.21]	[-2.70 , 0.10]	[-4.28 , -0.04]	(qrt 1)
Total Gov Spending	1.00	1.00	0.80	0.32	-0.29	
	[1.00 , 1.00]	[1.00 , 1.00]	[0.64 , 0.95]	[0.03 , 0.57]	[-0.79 , 0.12]	
Total Gov Revenue	0.00	0.00	-0.33	-0.85	-1.88	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[-0.86 , 0.25]	[-1.70 , 0.00]	[-3.44 , -0.55]	
The ARRW Methodology						
GDP	1.12	0.46	0.61	-0.20	-0.52	1.46
	[-4.33 , 5.67]	[-14.31 , 7.64]	[-16.42 , 24.11]	[-28.28 , 26.19]	[-37.73 , 32.55]	(qrt 3)
Total Gov Spending	1.00	1.00	0.86	0.42	-0.12	
	[1.00 , 1.00]	[1.00 , 1.00]	[-3.14 , 3.07]	[-4.98 , 4.22]	[-9.93 , 5.66]	
Total Gov Revenue	0.00	0.00	0.36	-0.37	-1.19	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[-9.85 , 12.48]	[-15.72 , 18.84]	[-27.21 , 22.74]	

Table 10: Fiscal Multipliers

their maximum value, the multipliers of the deficit-spending shocks are five times larger than the ones reported using the PFA. As already mentioned, these biases are hard to interpret because the multipliers being analyzed correspond to shocks that are linear combinations of shocks that are already biased. Most important, the ARRW methodology reports confidence intervals that are huge relative to the ones obtained using the PFA.

Panel (b) presents the impact multipliers associated with deficit-financed tax cut shocks. The upper block of this panel replicates the results reported in Table 3 in Mountford and Uhlig (2009). The bottom block of the panel reports the results obtained using the ARRW methodology. While Mountford and Uhlig (2009) find positive GDP median multipliers for at least 20 quarters, we find negative ones during the four quarters following the initial shock. After four quarters the median multipliers associated with the ARRW methodology also become positive. In addition, even when they share a sign, the median multipliers associated with our methodology are much smaller than the median multipliers implied by the PFA. In any case, as before, the confidence intervals computed using the ARRW methodology are so large that it is very hard to say anything concrete about the sign and size of the multipliers.

Panel (c) presents the impact multipliers associated with deficit-spending shocks. The upper block of this panel replicates the results reported in Table 4 in Mountford and Uhlig (2009), and the bottom block of the panel reports the results obtained using the ARRW methodology. In this policy scenario, both methodologies find the same sign (except for the 12 quarters) for the median multiplier. However, the magnitudes are different. The absolute value of the GDP median multipliers resulting from the ARRW methodology is approximately twice as large as the one resulting from the PFA. But it is also the case that the confidence intervals computed using the ARRW methodology are so wide that it is very hard to reach any conclusion.

Summarizing, Mountford and Uhlig (2009) use their results regarding the fiscal multipliers to emphasize that deficit-financed tax cut shocks work best to increase economic activity. But as mentioned before, the PFA is not agnostic. Once we use our truly agnostic methodology, it is very hard to support Mountford and Uhlig's (2009) claims. Some median multipliers change sign; nevertheless, the correct confidence intervals are so wide that it is very hard to reach any conclusion from a statistical point of view.

7.4 Computational Time

Our methodology is also faster than the PFA in the case of unanticipated fiscal policy shocks. In this case, the PFA is approximately three times slower than our methodology. The results are closer, but the ARRW methodology is still faster. In order to avoid local minima when using the PFA, we start the non-linear optimization from eight random starting points and then we pick the best one. Mountford and Uhlig (2009) follow a similar strategy.

	The PFA	The ARRW Methodology	Ratio
Fiscal Policy Shocks	12498.06	3929.3	3.18

Table 11: Computational Time in Seconds

8 Conclusion

We have presented an efficient algorithm for inference in SVARs identified with sign and zero restrictions that properly draws from the posterior distribution of structural parameters. The algorithm extends the sign restrictions methodology developed by Rubio-Ramírez, Waggoner and Zha (2010) to allow for zero restrictions. Our key theoretical contribution shows how to efficiently draw from the uniform distribution with respect to the Haar measure on the set of orthogonal matrices conditional on some linear restrictions on their coefficients holding. This is the crucial step that allows us to draw from the posterior distribution of structural parameters conditional on the sign and zero restrictions. We have used this algorithm to answer the following questions. Are optimism shocks an important source of business cycle fluctuations? Are deficit-financed tax cuts better than deficit spending to increase output? These questions have been previously studied by Beaudry, Nam and Wang (2011) and Mountford and Uhlig (2009), respectively, using the PFA. These authors have provided very definitive answers. Unfortunately, we have shown that these sharp conclusions are due to shortcomings in the PFA. In particular, we have shown that the PFA (1) imposes additional sign restrictions on variables that are seemingly unrestricted that bias the results, and (2) it chooses a single value of structural parameters, instead of drawing from its posterior, creating artificially narrow confidence intervals that also affect inference and the economic interpretation of the results. These shortcomings appear because the PFA does not correctly draw from the posterior distribution of structural parameters conditional on the sign and zero restrictions. This problem is common to all of the existing methods. Our algorithm is also faster than the current methods.

9 Appendices

9.1 Appendix A. Estimation and Inference: Optimism Shocks

Following Beaudry, Nam and Wang (2011) we estimate equation (3) with four lags using Bayesian methods with a Normal-Wishart prior as in Uhlig (2005). Specifically, we take 1,000 parameter draws from the Normal-Wishart posterior of the reduced-form parameters $(\mathbf{B}, \mathbf{\Sigma})$ and from the conditional uniform distribution of $\mathbf{Q}$. We use the data set created by Beaudry, Nam and Wang (2011). This data set contains quarterly U.S. data for the sample period 1955Q1-2010Q4 and includes the following variables: TFP, stock price, consumption, real federal funds rate, hours worked, investment, and output. TFP is the factor-utilization-adjusted TFP series from John Fernald's website. Stock price is the Standard and Poor's 500 composite index divided by the CPI of all items from the Bureau of Labor Statistics (BLS). Consumption is real consumption spending on non-durable goods and services from the Bureau of Economic Analysis (BEA). The real federal funds rate corresponds to the effective federal funds rate minus the inflation rate as measured by the growth rate of the CPI all items from the BLS. Hours worked is the hours of all persons in the non-farm business sector from the BLS. Investment is real gross private domestic investment from the BEA. Output is real output in the non-farm business sector from the BLS. The series corresponding to stock price, consumption, hours worked, investment, and output are normalized by the civilian non-institutional population of 16 years and over from the BLS. All variables are logarithmic levels except for the real interest rate, which is in levels but not logged.

9.2 Appendix B. Estimation and Inference: Fiscal Policy Shocks

Following Mountford and Uhlig (2009) we estimate equation (3) with six lags using Bayesian methods with a Normal-Wishart prior as specified in Uhlig (2005). We take 1,000 parameter draws from the Normal-Wishart posterior $(\mathbf{B}, \mathbf{\Sigma})$ and from the conditional uniform distribution of $\mathbf{Q}$. We use the same data set as Mountford and Uhlig (2009). This data set contains quarterly U.S. data for the sample period 1955Q1-2010Q4 and includes the following variables: GDP, private consumption, total government spending, total government revenue, real wages, private non-residential investment, interest rate, adjusted reserves, producer price index of raw materials, and GDP deflator. All variables are logarithmic levels except for the interest rate, which is expressed in levels but not logged.

9.3 Appendix C. Tables and Figures

The ARRW Methodology																		
	Identification I						**Identification II**						**Identification III**					
	h=0	**h=4**	**h=8**	**h=16**	**h=24**	**h=40**	**h=0**	**h=4**	**h=8**	**h=16**	**h=24**	**h=40**	**h=0**	**h=4**	**h=8**	**h=16**	**h=24**	**h=40**
Adjusted TFP	0.00	0.01	0.02	0.04	0.06	0.09	0.00	0.01	0.02	0.04	0.07	0.12	0.00	0.01	0.02	0.05	0.09	0.17
	[0.00 , 0.00]	[0.00 , 0.04]	[0.01 , 0.06]	[0.01 , 0.09]	[0.02 , 0.14]	[0.03 , 0.22]	[0.00 , 0.00]	[0.00 , 0.04]	[0.01 , 0.06]	[0.01 , 0.10]	[0.02 , 0.16]	[0.04 , 0.28]	[0.00 , 0.00]	[0.00 , 0.03]	[0.01 , 0.05]	[0.02 , 0.11]	[0.03 , 0.20]	[0.06 , 0.33]
Stock Price	0.16	0.16	0.16	0.16	0.16	0.16	0.23	0.25	0.25	0.26	0.26	0.26	0.22	0.27	0.29	0.30	0.31	0.31
	[0.02 , 0.55]	[0.03 , 0.53]	[0.02 , 0.53]	[0.03 , 0.52]	[0.03 , 0.51]	[0.03 , 0.47]	[0.03 , 0.60]	[0.04 , 0.62]	[0.05 , 0.62]	[0.06 , 0.61]	[0.06 , 0.60]	[0.07 , 0.58]	[0.02 , 0.63]	[0.05 , 0.67]	[0.06 , 0.68]	[0.07 , 0.66]	[0.08 , 0.65]	[0.09 , 0.62]
Consumption	0.17	0.18	0.17	0.17	0.17	0.17	0.20	0.30	0.31	0.31	0.30	0.28	0.22	0.34	0.38	0.41	0.41	0.40
	[0.02 , 0.53]	[0.02 , 0.53]	[0.02 , 0.53]	[0.03 , 0.52]	[0.02 , 0.51]	[0.02 , 0.49]	[0.02 , 0.58]	[0.08 , 0.63]	[0.08 , 0.64]	[0.08 , 0.62]	[0.07 , 0.61]	[0.06 , 0.59]	[0.02 , 0.63]	[0.10 , 0.70]	[0.13 , 0.70]	[0.15 , 0.69]	[0.15 , 0.67]	[0.13 , 0.66]
Real Interest Rate	0.15	0.15	0.16	0.16	0.17	0.18	0.16	0.16	0.16	0.16	0.17	0.20	0.18	0.19	0.20	0.20	0.21	0.23
	[0.01 , 0.50]	[0.04 , 0.46]	[0.04 , 0.45]	[0.05 , 0.43]	[0.06 , 0.42]	[0.07 , 0.39]	[0.01 , 0.50]	[0.04 , 0.45]	[0.04 , 0.45]	[0.05 , 0.43]	[0.07 , 0.42]	[0.08 , 0.40]	[0.01 , 0.55]	[0.04 , 0.50]	[0.05 , 0.50]	[0.06 , 0.48]	[0.07 , 0.46]	[0.09 , 0.44]
Hours Worked	0.16	0.19	0.19	0.18	0.18	0.18	0.16	0.26	0.31	0.30	0.28	0.27	0.17	0.25	0.31	0.31	0.30	0.29
	[0.02 , 0.53]	[0.04 , 0.52]	[0.04 , 0.53]	[0.04 , 0.52]	[0.04 , 0.49]	[0.04 , 0.48]	[0.01 , 0.53]	[0.07 , 0.62]	[0.07 , 0.61]	[0.06 , 0.59]	[0.07 , 0.57]	[0.07 , 0.55]	[0.02 , 0.51]	[0.06 , 0.60]	[0.06 , 0.62]	[0.06 , 0.61]	[0.07 , 0.60]	[0.07 , 0.57]
The PFA																		
	Identification I						**Identification II**						**Identification III**					
	h=0	**h=4**	**h=8**	**h=16**	**h=24**	**h=40**	**h=0**	**h=4**	**h=8**	**h=16**	**h=24**	**h=40**	**h=0**	**h=4**	**h=8**	**h=16**	**h=24**	**h=40**
Adjusted TFP	0.00	0.03	0.04	0.05	0.08	0.17	0.00	0.02	0.03	0.04	0.08	0.22	0.00	0.01	0.02	0.05	0.10	0.28
	[0.00 , 0.00]	[0.01 , 0.07]	[0.01 , 0.09]	[0.02 , 0.09]	[0.04 , 0.14]	[0.08 , 0.30]	[0.00 , 0.00]	[0.01 , 0.04]	[0.01 , 0.07]	[0.02 , 0.08]	[0.03 , 0.16]	[0.10 , 0.37]	[0.00 , 0.00]	[0.00 , 0.03]	[0.01 , 0.04]	[0.02 , 0.10]	[0.04 , 0.20]	[0.14 , 0.43]
Stock Price	1.00	0.93	0.91	0.85	0.80	0.72	0.64	0.71	0.71	0.73	0.73	0.71	0.41	0.51	0.53	0.57	0.58	0.57
	[0.98 , 1.00]	[0.89 , 0.96]	[0.85 , 0.95]	[0.76 , 0.92]	[0.67 , 0.89]	[0.55 , 0.85]	[0.60 , 0.68]	[0.64 , 0.77]	[0.62 , 0.79]	[0.61 , 0.82]	[0.61 , 0.83]	[0.57 , 0.82]	[0.36 , 0.47]	[0.43 , 0.58]	[0.43 , 0.62]	[0.44 , 0.68]	[0.44 , 0.71]	[0.42 , 0.72]
Consumption	0.07	0.27	0.30	0.30	0.28	0.26	0.59	0.81	0.82	0.78	0.75	0.69	0.51	0.71	0.77	0.80	0.80	0.76
	[0.04 , 0.11]	[0.19 , 0.35]	[0.21 , 0.40]	[0.20 , 0.42]	[0.17 , 0.42]	[0.13 , 0.43]	[0.54 , 0.63]	[0.75 , 0.86]	[0.75 , 0.89]	[0.69 , 0.87]	[0.63 , 0.85]	[0.53 , 0.83]	[0.46 , 0.55]	[0.64 , 0.77]	[0.68 , 0.84]	[0.68 , 0.88]	[0.66 , 0.89]	[0.59 , 0.87]
Real Interest Rate	0.00	0.03	0.04	0.06	0.09	0.13	0.02	0.03	0.05	0.06	0.08	0.13	0.37	0.34	0.36	0.34	0.34	0.35
	[0.00 , 0.01]	[0.02 , 0.06]	[0.02 , 0.07]	[0.03 , 0.10]	[0.05 , 0.15]	[0.07 , 0.22]	[0.00 , 0.03]	[0.01 , 0.06]	[0.02 , 0.09]	[0.03 , 0.11]	[0.04 , 0.14]	[0.07 , 0.22]	[0.32 , 0.42]	[0.29 , 0.40]	[0.29 , 0.43]	[0.28 , 0.42]	[0.28 , 0.41]	[0.29 , 0.43]
Hours Worked	0.02	0.30	0.37	0.34	0.32	0.31	0.05	0.56	0.70	0.69	0.66	0.62	0.03	0.38	0.52	0.54	0.52	0.49
	[0.00 , 0.04]	[0.22 , 0.38]	[0.27 , 0.47]	[0.24 , 0.47]	[0.21 , 0.45]	[0.21 , 0.44]	[0.03 , 0.08]	[0.48 , 0.63]	[0.62 , 0.77]	[0.59 , 0.77]	[0.54 , 0.75]	[0.48 , 0.73]	[0.02 , 0.06]	[0.31 , 0.46]	[0.41 , 0.62]	[0.40 , 0.67]	[0.37 , 0.67]	[0.34 , 0.64]

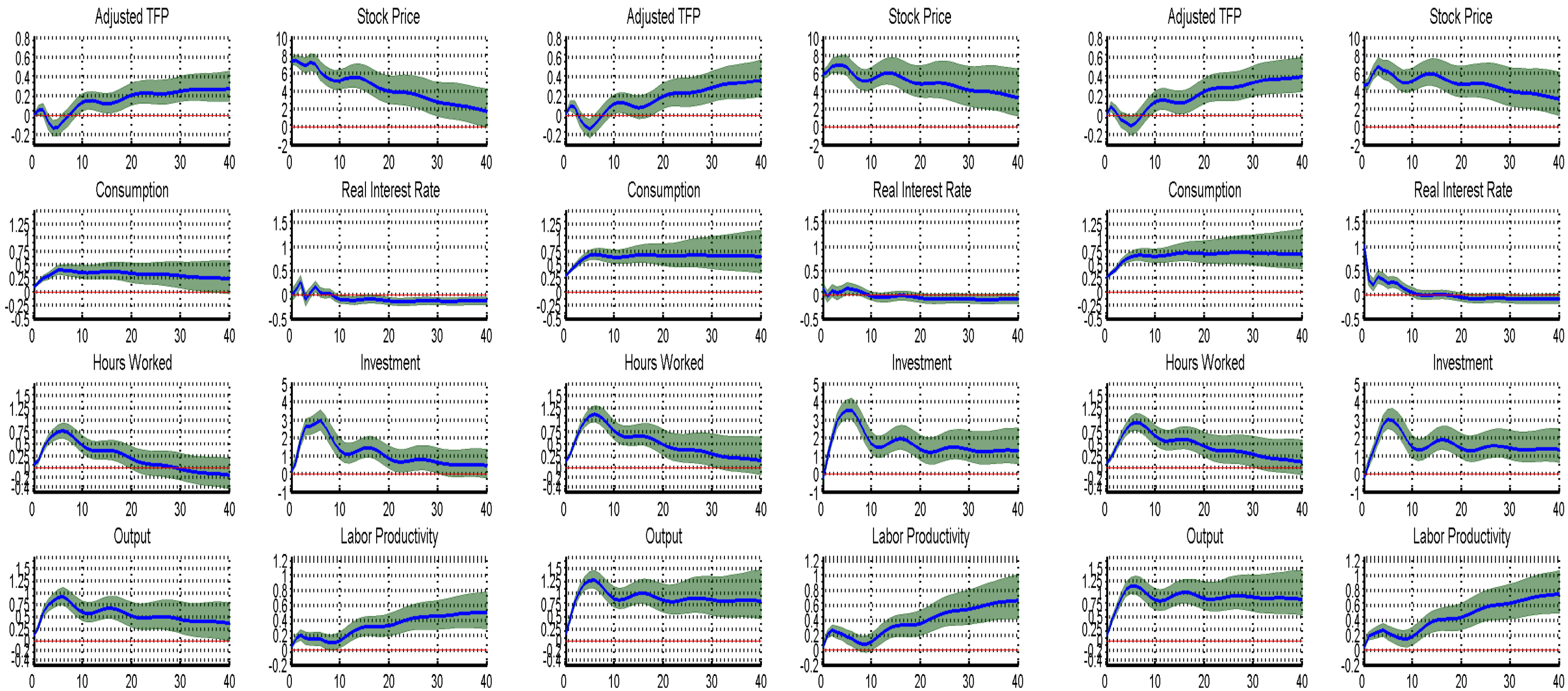
Table 12: Share of FEV Attributable to Optimism Shocks: Five-Variable SVAR

The ARRW Methodology			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.08 [0.03 , 0.21]	0.10 [0.03 , 0.23]	0.13 [0.05 , 0.28]
Stock Price	0.12 [0.03 , 0.34]	0.17 [0.04 , 0.39]	0.23 [0.07 , 0.46]
Consumption	0.11 [0.02 , 0.35]	0.14 [0.03 , 0.40]	0.21 [0.05 , 0.49]
Real Interest Rate	0.13 [0.06 , 0.27]	0.13 [0.05 , 0.27]	0.14 [0.06 , 0.28]
Hours Worked	0.12 [0.03 , 0.33]	0.14 [0.04 , 0.34]	0.16 [0.05 , 0.37]
Investment	0.13 [0.05 , 0.30]	0.15 [0.05 , 0.35]	0.18 [0.07 , 0.38]
Output	0.12 [0.03 , 0.33]	0.16 [0.05 , 0.39]	0.22 [0.07 , 0.46]
Labor Productivity	0.10 [0.02 , 0.29]	0.12 [0.03 , 0.34]	0.20 [0.04 , 0.44]
The PFA			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.17 [0.08 , 0.30]	0.21 [0.08 , 0.38]	0.29 [0.13 , 0.46]
Stock Price	0.52 [0.34 , 0.70]	0.62 [0.48 , 0.73]	0.57 [0.43 , 0.69]
Consumption	0.13 [0.06 , 0.28]	0.55 [0.39 , 0.69]	0.61 [0.43 , 0.76]
Real Interest Rate	0.10 [0.05 , 0.17]	0.06 [0.03 , 0.13]	0.26 [0.19 , 0.34]
Hours Worked	0.16 [0.09 , 0.26]	0.38 [0.25 , 0.53]	0.30 [0.18 , 0.44]
Investment	0.25 [0.16 , 0.37]	0.45 [0.34 , 0.57]	0.39 [0.26 , 0.52]
Output	0.23 [0.13 , 0.39]	0.59 [0.46 , 0.72]	0.60 [0.44 , 0.73]
Labor Productivity	0.24 [0.12 , 0.39]	0.37 [0.17 , 0.55]	0.49 [0.27 , 0.65]

Table 13: Share of FEV Attributable to Optimism Shocks at Horizon 40: Seven-Variable SVAR

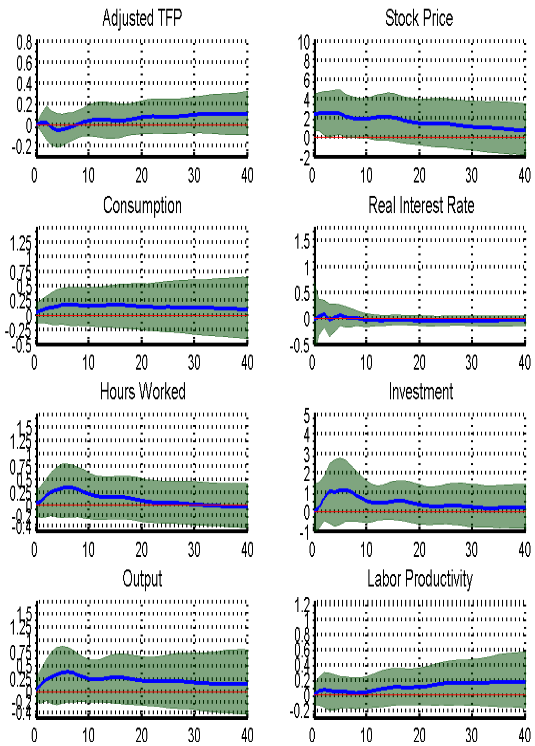
The ARRW Methodology																		
	Identification I						Identification II						Identification III					
	h=0	h=4	h=8	h=16	h=24	h=40	h=0	h=4	h=8	h=16	h=24	h=40	h=0	h=4	h=8	h=16	h=24	h=40
Adjusted TFP	0.00	0.02	0.03	0.05	0.06	0.08	0.00	0.02	0.03	0.05	0.07	0.10	0.00	0.02	0.03	0.06	0.08	0.13
	[0.00 , 0.00]	[0.01 , 0.07]	[0.01 , 0.08]	[0.02 , 0.13]	[0.02 , 0.16]	[0.03 , 0.21]	[0.00 , 0.00]	[0.01 , 0.07]	[0.01 , 0.09]	[0.02 , 0.13]	[0.02 , 0.16]	[0.03 , 0.23]	[0.00 , 0.00]	[0.01 , 0.07]	[0.01 , 0.09]	[0.02 , 0.14]	[0.03 , 0.19]	[0.05 , 0.28]
Stock Price	0.10	0.11	0.10	0.11	0.11	0.12	0.12	0.15	0.16	0.17	0.17	0.17	0.13	0.19	0.21	0.23	0.24	0.23
	[0.01 , 0.36]	[0.02 , 0.35]	[0.02 , 0.34]	[0.02 , 0.36]	[0.03 , 0.35]	[0.03 , 0.34]	[0.01 , 0.39]	[0.03 , 0.41]	[0.04 , 0.40]	[0.04 , 0.42]	[0.04 , 0.42]	[0.04 , 0.39]	[0.01 , 0.39]	[0.05 , 0.46]	[0.06 , 0.46]	[0.07 , 0.46]	[0.08 , 0.48]	[0.07 , 0.46]
Consumption	0.10	0.11	0.11	0.11	0.11	0.11	0.10	0.16	0.17	0.16	0.15	0.14	0.13	0.20	0.23	0.24	0.24	0.21
	[0.01 , 0.34]	[0.01 , 0.35]	[0.02 , 0.36]	[0.02 , 0.37]	[0.02 , 0.37]	[0.02 , 0.35]	[0.01 , 0.35]	[0.03 , 0.41]	[0.03 , 0.42]	[0.03 , 0.42]	[0.03 , 0.42]	[0.03 , 0.40]	[0.01 , 0.37]	[0.05 , 0.47]	[0.06 , 0.50]	[0.07 , 0.51]	[0.06 , 0.50]	[0.05 , 0.49]
Real Interest Rate	0.09	0.10	0.11	0.11	0.12	0.13	0.09	0.10	0.11	0.11	0.12	0.13	0.11	0.11	0.12	0.12	0.13	0.14
	[0.01 , 0.34]	[0.03 , 0.30]	[0.04 , 0.29]	[0.04 , 0.28]	[0.05 , 0.28]	[0.06 , 0.27]	[0.01 , 0.35]	[0.03 , 0.30]	[0.03 , 0.30]	[0.04 , 0.29]	[0.05 , 0.28]	[0.05 , 0.27]	[0.01 , 0.37]	[0.03 , 0.32]	[0.03 , 0.32]	[0.04 , 0.30]	[0.05 , 0.29]	[0.06 , 0.28]
Hours Worked	0.09	0.11	0.12	0.12	0.11	0.12	0.09	0.15	0.17	0.16	0.15	0.14	0.09	0.14	0.17	0.17	0.17	0.16
	[0.01 , 0.36]	[0.03 , 0.34]	[0.03 , 0.35]	[0.03 , 0.33]	[0.03 , 0.33]	[0.03 , 0.33]	[0.01 , 0.37]	[0.04 , 0.44]	[0.04 , 0.44]	[0.04 , 0.40]	[0.04 , 0.38]	[0.04 , 0.34]	[0.01 , 0.37]	[0.03 , 0.44]	[0.04 , 0.46]	[0.04 , 0.44]	[0.05 , 0.42]	[0.05 , 0.37]
Investment	0.08	0.11	0.11	0.12	0.12	0.13	0.08	0.13	0.15	0.16	0.16	0.15	0.07	0.12	0.16	0.17	0.18	0.18
	[0.01 , 0.26]	[0.03 , 0.31]	[0.03 , 0.32]	[0.04 , 0.31]	[0.04 , 0.31]	[0.05 , 0.30]	[0.01 , 0.26]	[0.04 , 0.36]	[0.05 , 0.38]	[0.05 , 0.36]	[0.05 , 0.35]	[0.05 , 0.35]	[0.01 , 0.25]	[0.04 , 0.34]	[0.05 , 0.39]	[0.05 , 0.38]	[0.06 , 0.38]	[0.07 , 0.38]
Output	0.04	0.10	0.10	0.11	0.12	0.12	0.04	0.15	0.17	0.18	0.17	0.16	0.04	0.17	0.21	0.23	0.23	0.22
	[0.00 , 0.13]	[0.02 , 0.32]	[0.02 , 0.33]	[0.03 , 0.32]	[0.03 , 0.33]	[0.03 , 0.33]	[0.00 , 0.14]	[0.03 , 0.39]	[0.04 , 0.40]	[0.05 , 0.40]	[0.05 , 0.40]	[0.05 , 0.39]	[0.00 , 0.14]	[0.03 , 0.40]	[0.06 , 0.44]	[0.07 , 0.46]	[0.08 , 0.46]	[0.07 , 0.46]
Labor Productivity	0.02	0.06	0.07	0.07	0.08	0.10	0.02	0.06	0.07	0.08	0.10	0.12	0.02	0.06	0.08	0.10	0.13	0.20
	[0.00 , 0.06]	[0.01 , 0.18]	[0.02 , 0.19]	[0.02 , 0.22]	[0.02 , 0.25]	[0.02 , 0.29]	[0.00 , 0.06]	[0.01 , 0.19]	[0.02 , 0.21]	[0.02 , 0.25]	[0.03 , 0.29]	[0.03 , 0.34]	[0.00 , 0.06]	[0.01 , 0.22]	[0.02 , 0.25]	[0.03 , 0.31]	[0.03 , 0.36]	[0.04 , 0.44]
The PFA																		
	Identification I						Identification II						Identification III					
	h=0	h=4	h=8	h=16	h=24	h=40	h=0	h=4	h=8	h=16	h=24	h=40	h=0	h=4	h=8	h=16	h=24	h=40
Adjusted TFP	0.00	0.02	0.02	0.05	0.09	0.17	0.00	0.02	0.03	0.05	0.09	0.21	0.00	0.01	0.02	0.05	0.12	0.29
	[0.00 , 0.00]	[0.01 , 0.03]	[0.01 , 0.05]	[0.03 , 0.10]	[0.04 , 0.18]	[0.08 , 0.30]	[0.00 , 0.00]	[0.01 , 0.03]	[0.02 , 0.05]	[0.02 , 0.09]	[0.04 , 0.18]	[0.08 , 0.38]	[0.00 , 0.00]	[0.01 , 0.03]	[0.01 , 0.05]	[0.03 , 0.11]	[0.05 , 0.24]	[0.13 , 0.46]
Stock Price	1.00	0.87	0.82	0.74	0.65	0.52	0.63	0.72	0.72	0.71	0.69	0.62	0.39	0.57	0.60	0.63	0.62	0.57
	[0.99 , 1.00]	[0.82 , 0.91]	[0.75 , 0.87]	[0.63 , 0.82]	[0.51 , 0.77]	[0.34 , 0.70]	[0.59 , 0.66]	[0.66 , 0.78]	[0.63 , 0.79]	[0.61 , 0.80]	[0.57 , 0.79]	[0.48 , 0.73]	[0.34 , 0.45]	[0.49 , 0.63]	[0.50 , 0.67]	[0.52 , 0.72]	[0.50 , 0.72]	[0.43 , 0.69]
Consumption	0.04	0.19	0.23	0.22	0.19	0.13	0.53	0.72	0.74	0.70	0.64	0.55	0.44	0.62	0.68	0.71	0.69	0.61
	[0.02 , 0.07]	[0.13 , 0.26]	[0.15 , 0.31]	[0.13 , 0.32]	[0.10 , 0.31]	[0.06 , 0.28]	[0.49 , 0.58]	[0.65 , 0.77]	[0.66 , 0.80]	[0.60 , 0.78]	[0.53 , 0.75]	[0.39 , 0.69]	[0.39 , 0.49]	[0.55 , 0.70]	[0.60 , 0.77]	[0.59 , 0.80]	[0.55 , 0.79]	[0.43 , 0.76]
Real Interest Rate	0.00	0.03	0.04	0.05	0.07	0.10	0.01	0.02	0.03	0.04	0.05	0.06	0.36	0.30	0.30	0.28	0.27	0.26
	[0.00 , 0.01]	[0.02 , 0.05]	[0.02 , 0.06]	[0.03 , 0.08]	[0.04 , 0.12]	[0.05 , 0.17]	[0.00 , 0.03]	[0.01 , 0.04]	[0.01 , 0.06]	[0.02 , 0.07]	[0.02 , 0.09]	[0.03 , 0.13]	[0.31 , 0.41]	[0.24 , 0.35]	[0.23 , 0.36]	[0.22 , 0.35]	[0.21 , 0.34]	[0.19 , 0.34]
Hours Worked	0.01	0.23	0.29	0.24	0.20	0.16	0.04	0.46	0.60	0.56	0.50	0.38	0.02	0.30	0.43	0.41	0.38	0.30
	[0.00 , 0.03]	[0.17 , 0.30]	[0.21 , 0.37]	[0.16 , 0.34]	[0.12 , 0.30]	[0.09 , 0.26]	[0.02 , 0.07]	[0.39 , 0.54]	[0.52 , 0.69]	[0.46 , 0.65]	[0.38 , 0.60]	[0.25 , 0.53]	[0.01 , 0.04]	[0.22 , 0.37]	[0.32 , 0.52]	[0.29 , 0.53]	[0.25 , 0.51]	[0.18 , 0.44]
Investment	0.00	0.27	0.35	0.33	0.30	0.25	0.01	0.38	0.51	0.50	0.49	0.45	0.01	0.21	0.33	0.36	0.39	0.39
	[0.00 , 0.01]	[0.20 , 0.34]	[0.26 , 0.43]	[0.25 , 0.42]	[0.22 , 0.40]	[0.16 , 0.37]	[0.00 , 0.02]	[0.30 , 0.45]	[0.41 , 0.59]	[0.40 , 0.60]	[0.39 , 0.59]	[0.34 , 0.57]	[0.00 , 0.02]	[0.15 , 0.28]	[0.25 , 0.44]	[0.25 , 0.47]	[0.26 , 0.50]	[0.26 , 0.52]
Output	0.01	0.28	0.35	0.33	0.31	0.23	0.03	0.56	0.67	0.69	0.66	0.59	0.02	0.40	0.53	0.60	0.63	0.60
	[0.01 , 0.02]	[0.21 , 0.35]	[0.26 , 0.43]	[0.24 , 0.45]	[0.21 , 0.43]	[0.13 , 0.39]	[0.02 , 0.05]	[0.48 , 0.62]	[0.60 , 0.74]	[0.60 , 0.76]	[0.56 , 0.75]	[0.46 , 0.72]	[0.01 , 0.03]	[0.33 , 0.48]	[0.44 , 0.63]	[0.49 , 0.70]	[0.50 , 0.73]	[0.44 , 0.73]
Labor Productivity	0.00	0.05	0.06	0.10	0.16	0.24	0.00	0.11	0.09	0.12	0.20	0.37	0.00	0.10	0.10	0.18	0.29	0.49
	[0.00 , 0.01]	[0.02 , 0.10]	[0.02 , 0.11]	[0.04 , 0.19]	[0.07 , 0.28]	[0.12 , 0.39]	[0.00 , 0.01]	[0.06 , 0.16]	[0.04 , 0.15]	[0.05 , 0.22]	[0.08 , 0.34]	[0.17 , 0.55]	[0.00 , 0.00]	[0.06 , 0.16]	[0.05 , 0.18]	[0.09 , 0.30]	[0.14 , 0.46]	[0.27 , 0.65]

Table 14: Share of FEV Attributable to Optimism Shocks: Seven-Variable SVAR

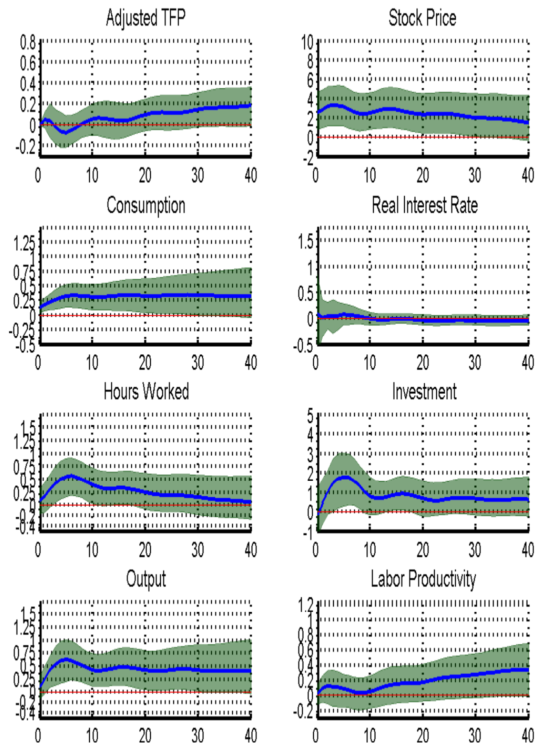


(a) Identification 1 (b) Identification 2 (c) Identification 3

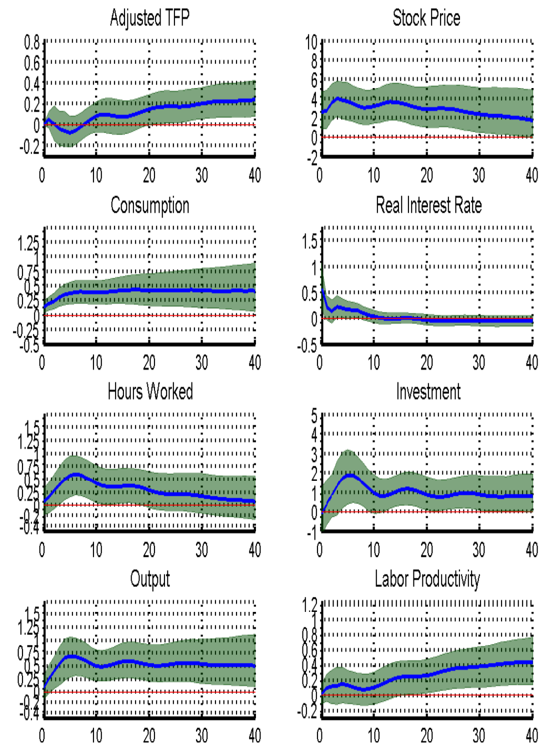
Figure 17: IRFs to an Optimism Shock Using the PFA: Seven-Variable SVAR



(a) Identification 1

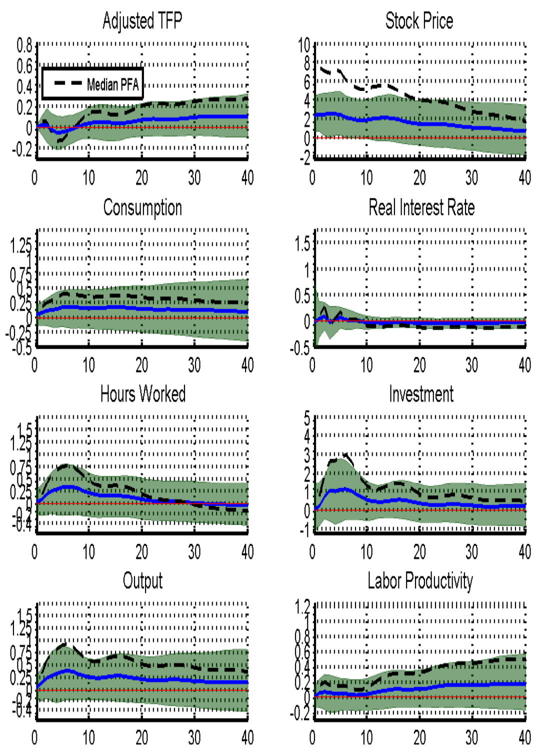


(b) Identification 2

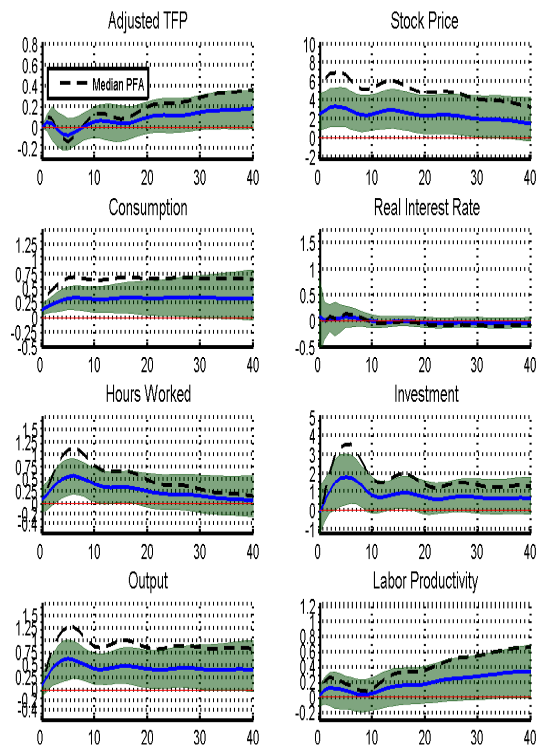


(c) Identification 3

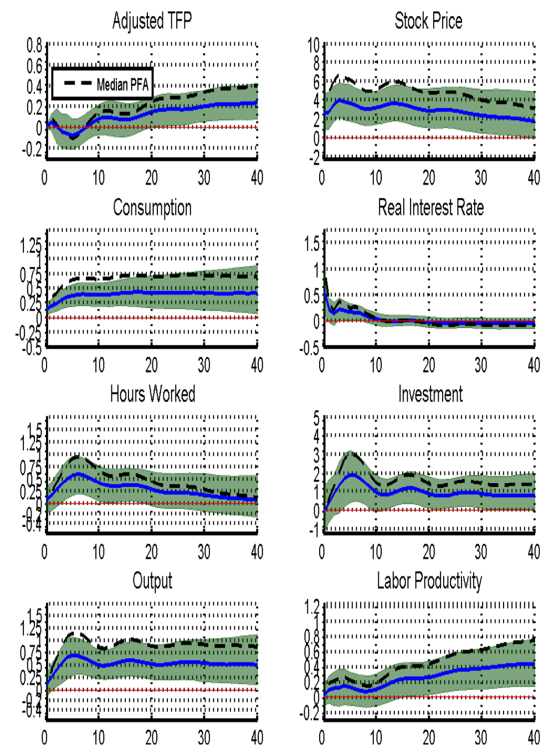
Figure 18: IRFs to an Optimism Shock Using the ARRW Methodology: Seven-Variable SVAR



(a) Identification 1



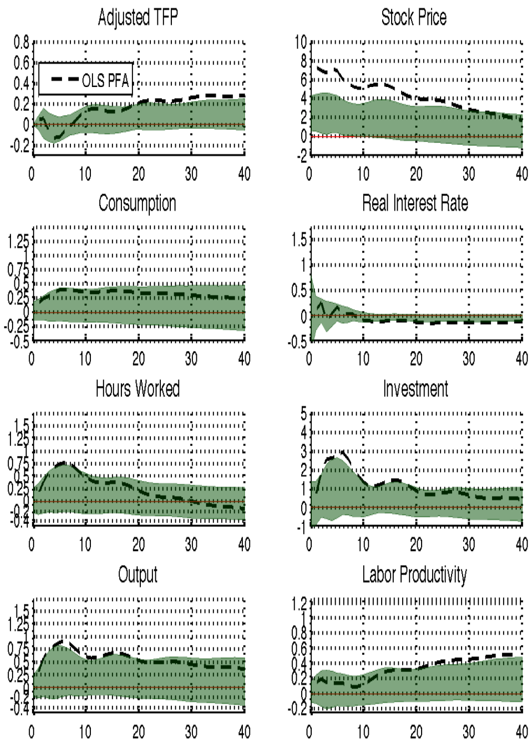
(b) Identification 2



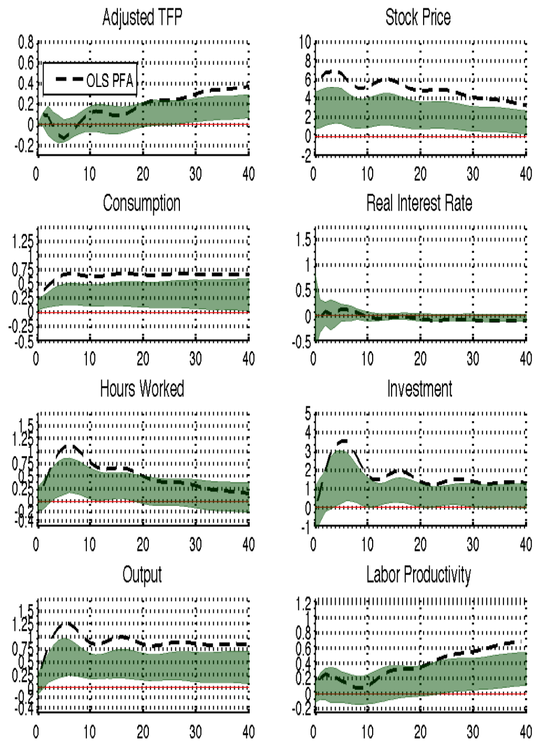
(c) Identification 3

Figure 19: Comparison of IRFs to an Optimism Shock: Seven-Variable SVAR

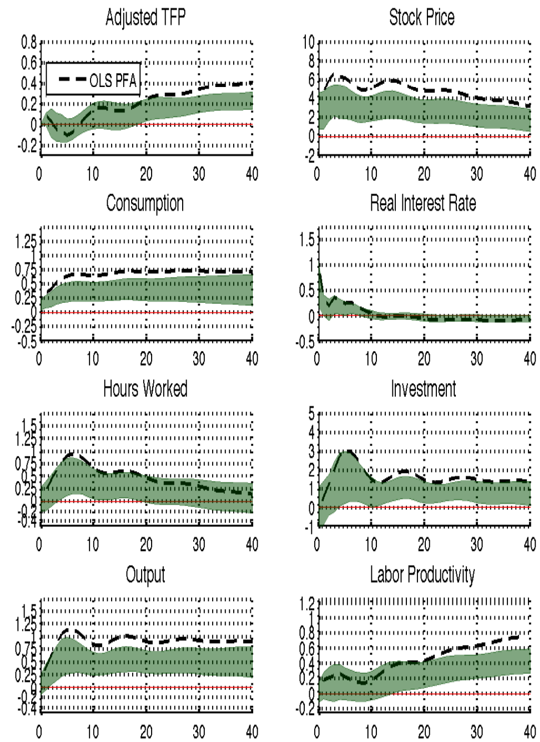
Note: Median PFA refers to the median IRF obtained using the PFA.



(a) Identification 1



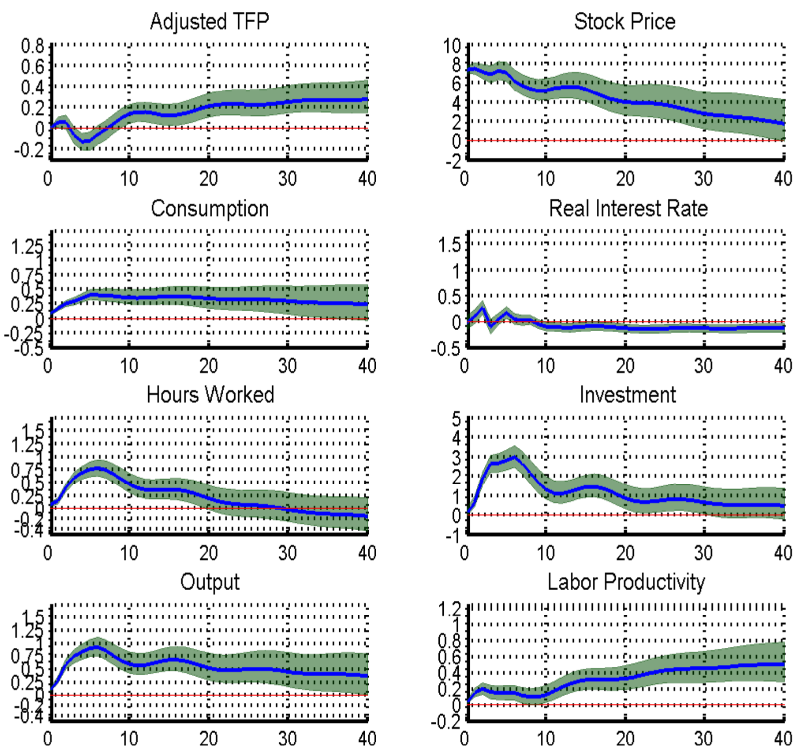
(b) Identification 2



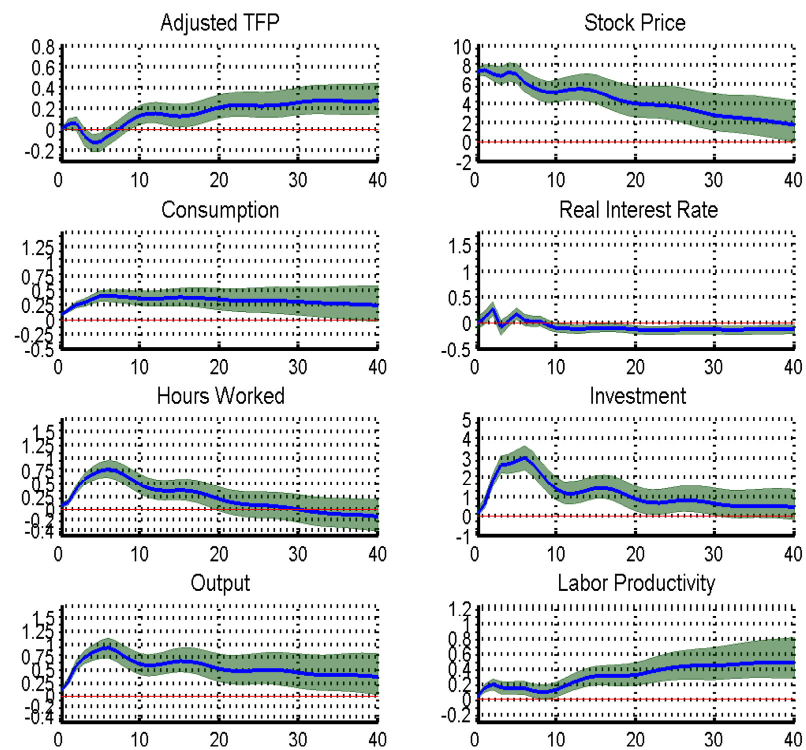
(c) Identification 3

Figure 20: Distribution of IRFs with the ARRW Methodology vs. Single IRFs with the PFA: Seven-Variable SVAR

Note: OLS PFA refers to the IRF obtained using the PFA and the OLS reduced-form estimates.



(a) The PFA



(b) ARRW with Additional Restrictions

Figure 21: Replicating the PFA Using the ARRW Methodology: Seven-Variable SVAR

References

Barsky, R. B. and E. R. Sims (2011). News Shocks and Business Cycles. *Journal of Monetary Economics 58*(3), 273–289.

Baumeister, C. and L. Benati (2013). Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound. *International Journal of Central Banking 9*(2), 165–212.

Baumeister, C. and J. Hamilton (2013). Sign Restrictions, Structural Vector Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information. *UCSD mimeo*.

Beaudry, P., D. Nam, and J. Wang (2011). Do Mood Swings Drive Business Cycles and is it Rational? *NBER Working Papers*.

Benati, L. (2013). Why are Recessions Associated with Financial Crises Different? *Working Paper*.

Bernanke, B. S. and I. Mihov (1998). Measuring Monetary Policy. *Quarterly Journal of Economics 113*(3), 869–902.

Binning, A. (2013). Underidentified SVAR Models: A Framework for Combining Short and Long-run Restrictions with Sign-restrictions. *Norges Bank Working Papers*.

Bowman, A. W. and A. Azzalini (1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford University Press.

Caldara, D. and C. Kamps (2012). The Analytics of SVAR: A Unified Framework to Measure Fiscal Multipliers. *Finance and Economics Discussion Series 2012-20*.

Canova, F. and G. D. Nicoló (2002). Monetary Disturbances Matter for Business Fluctuations in the G-7. *Journal of Monetary Economics 49*(6), 1131–1159.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1996). The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds. *Review of Economics and Statistics*, 16–34.

Faust, J. (1998). The Robustness of Identified VAR Conclusions about Money. *Carnegie-Rochester Conference Series on Public Policy 49*, 207–244.

Geweke, J. (1986). Exact Inference in the Inequality Constrained Normal Linear Regression Model. *Journal of Applied Econometrics 1*(2), 127–141.

Gordon, D. B. and E. M. Leeper (1994). The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification. *Journal of Political Economy*, 1228–1247.

King, R., C. I. Plosser, J. H. Stock, and M. W. Watson (1994). Stochastic Trends and Economic Fluctuations. *Journal of Political Economy*, 1228–1247.

Krantz, S. G. and H. R. Parks (2008). *Geometric Integration Theory*. Birkhauser.

Mountford, A. and H. Uhlig (2009). What are the Effects of Fiscal Policy Shocks? *Journal of Applied Econometrics 24*(6), 960–992.

Rubio-Ramírez, J., D. Waggoner, and T. Zha (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *Review of Economic Studies 77*(2), 665–696.

Sims, C. A. (1986). Are Forecasting Models Usable for Policy Analysis? *Federal Reserve Bank of Minneapolis Quarterly Review 10*(1), 2–16.

Sims, C. A. and T. Zha (2006). Were There Regime Switches in US Monetary Policy? *American Economic Review*, 54–81.

Stewart, G. (1980). The Efficient Generation of Random Orthogonal Matrices with an Application to Condition Estimators. *SIAM Journal on Numerical Analysis 17*(3), 403–409.

Uhlig, H. (2005). What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure. *Journal of Monetary Economics 52*(2), 381–419.

Waggoner, D. F. and T. Zha (2003). Likelihood Preserving Normalization in Multiple Equation Models. *Journal of Econometrics 114*(2), 329–347.

Zha, T. (1999). Block Recursion and Structural Vector Autoregressions. *Journal of Econometrics 90*(2), 291–316.