Algorithm for Inference with Sign and Zero Restrictions

Very Preliminary: Please do not cite or distribute

Jonas E. Arias
Duke University

Juan F. Rubio-Ramírez*
Duke University
Federal Reserve Bank of Atlanta

Daniel F. Waggoner
Federal Reserve Bank of Atlanta

July 25, 2013

Abstract

We extend the methodology for inference in SVARs identified with sign restrictions developed by Rubio-Ramírez et al. (2010) to also allow for zero restrictions. Then we use this methodology 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 two questions have been previously studied in the literature and the answers have been quite definite. Unfortunately, we show that these sharp conclusions are due to shortcomings in the current methods used for inference since none of the existing methods correctly draws from the posterior distribution of structural parameters conditional on the sign and zero restrictions holding. Our methodology properly draws from this posterior distribution. Once draws are correctly done, it is very hard to support that either optimism shocks are an important source of business cycle fluctuations or deficit-financed tax cuts work best at improving output as the literature has previously claimed. Our method is not only correct but also faster than current ones.

*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, Deokwoo Nam, Andrew Mountford, and Jian Wang for sharing supplementary material with us, and for helpful comments. We also thank the support and help of Grátula Bedátula. Without her this paper would have been impossible.

1 Introduction

Are optimism shocks an important source of business cycle fluctuations? Are deficit-financed tax cuts better than deficit-spending to increase output? Several questions such as these have been previously studied in the literature using SVARs identified imposing sign and zero restrictions on impulse response functions and frequently the answers have been definite. Beaudry et al. (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 to stimulate economic activity. Unfortunately, we show that these sharp conclusions are due to shortcomings in the current methods used for inference: none of the existing methods correctly draws from the posterior distribution of structural parameters conditional on the sign and zero restrictions.

The most widely used method is Mountford and Uhlig (2009)'s penalty function approach. Instead of drawing from the posterior distribution of structural parameters such that sign and zero restrictions hold, the penalty function approach selects a single value of the structural parameters by minimizing a loss function. We show that this approach has several shortcomings that crucially affect inference. First, the penalty function approach imposes additional sign restrictions on variables that are seemingly unrestricted, which bias the results. Indeed, for a particular 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 penalty function approach creates artificially narrow confidence intervals that severely affect inference and the economic interpretation of the results. Baumeister and Benati (2010) and Benati (2013) provide another approach that, although does not minimize any loss function, suffers from similar limitations than the penalty function approach because it neither draws from the posterior distribution of structural parameters.

This paper presents an efficient algorithm for inference in SVARs identified with sign and zero restrictions that properly draws from the posterior distribution of structural parameters. We extend the sign restrictions methodology developed by Rubio-Ramírez et al. (2010) to allow for zero restrictions. As it was the case in Rubio-Ramírez et al. (2010) we put forward most of our results imposing sign and zero restrictions on the impulse response functions, but our algorithm allows for a larger class of 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 crucial step allows us to draw from the posterior distribution of structural parameters conditional on the sign and zero restrictions.

We show the capabilities of our methodology and the shortcomings of the penalty function approach by means of two applications previously analyzed in the literature using the penalty function approach. The first application is related to optimism shocks as in Beaudry et al. (2011). The aim of Beaudry et al. (2011) is to provide new evidence on the relevance of optimism shocks as the main driver of macroeconomic fluctuations. We show how their main economic conclusion dramatically changes once our methodology accounts for the shortcomings of the penalty function approach. While Beaudry et al. (2011) conclude that optimism shocks are associated with standard business cycle type phenomena, using our methodology we show that it is very hard to support such claim. The sharp results reported in Beaudry et al. (2011) are due to biased impulse response functions and artificially narrow confidence intervals associated with the penalty function approach.

The second application identifies fiscal policy shocks as in Mountford and Uhlig (2009) in order to analyze the effects of fiscal policy shocks on economic activity. Mountford and Uhlig (2009)'s main finding is that deficit-financed tax cuts work best among different fiscal policies aimed at improving GDP. Using our methodology we find no evidence to support their finding. Analogous to Beaudry et al. (2011), the tight results obtained in Mountford and Uhlig (2009) are due to biased impulse response functions and artificially narrow confidence intervals associated with the penalty function approach.

It also important to remark that, at least for the two applications studied in this paper, our algorithm is faster than the penalty function approach. Our methodology is between 1.6 times and 25 faster than the penalty function approach depending on the number of sign and zero restrictions. Finally, we wish to argue that the aim of this paper is neither dispute nor 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 Canova and Nicoló (2002), Uhlig (2005), and Rubio-Ramírez et al. (2010). Instead, our findings related to optimism and fiscal policy shocks just indicate that the respective sign and zero restrictions are not enough to accurately identify these particular structural shocks.

The paper is organized as follows. Section 2 presents the methodology. It is here where we describe our theoretical contributions and algorithms. Section 3 shows some examples. Section 4 describes the penalty function approach and highlights its shortcomings. Section 5 presents the first of our applications. Section 6 presents the second application. Section 7 concludes.

2 Our Methodology

This section is organized in three parts. First, we describe the model. Second, we review the efficient algorithm for inference using sign restrictions on impulse response functions (IRFs) developed in Rubio-Ramírez et al. (2010). Third, we extend this algorithm to also allow for zero restrictions. It is important to note that the algorithm proposed by Rubio-Ramírez et al. (2010) and our extension can be embedded in a classical or bayesian framework. In this paper we follow the latter.

2.1 The Model

Consider the structural vector autoregression (SVAR) with the general form as in Rubio-Ramírez et al. (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 a $n \times 1$ vector of endogenous variables, ε_t an $n \times 1$ vector of exogenous structural shocks, $\mathbf{A}_\ell$ an $n \times n$ matrix of parameters for $0 \leq \ell \leq p$, $\mathbf{c}$ 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 $\mathbb{E}\left[\mathbf{u}_t\mathbf{u}_t'\right] = \boldsymbol{\Sigma} = \left(\mathbf{A}_0\mathbf{A}_0'\right)^{-1}$. The matrices $\mathbf{B}$ and $\boldsymbol{\Sigma}$ are the reduced-form parameters while $\mathbf{A}_0$ and $\mathbf{A}_+$ are the structural parameters; we assume that $\mathbf{A}_0$ is invertible.

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*

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

where

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

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 long-run IRF) corresponds to the element in row i and column j of the matrix*

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

2.2 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

$$f(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} \mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_\infty(\mathbf{A}_0, \mathbf{A}_+) \end{bmatrix},$$

where $k = 2n$ 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 $f(\mathbf{A}_0, \mathbf{A}_+)$. 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. The total number of sign restrictions will be $s = \sum_{j=1}^{n} s_j$. Let $\mathbf{e}_j$ denote the j-th column of $\mathbf{I}_n$.

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 $1 \leq j \leq n$.

From equation (2), it is easy to see that if $(\mathbf{A}_0, \mathbf{A}_+)$ is any set of structural parameters and $\mathbf{Q}$ is any element of $O(n)$, the set 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 satisfy 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. Therefore, the set of structural parameters satisfying 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 $(\mathbf{A}_0, \mathbf{A}_+)$ conditional on satisfying 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 posterior draws satisfying the sign restrictions have been obtained.*

By unrestricted posterior we mean the posterior distribution of unrestricted structural parameters, i.e. when no identification schemes are considered.

The only obstacle to implementing this algorithm is an efficient technique to accomplish the first step. For instance, one could use the Gibbs sampler described in Waggoner and Zha (2003a). However, this technique produces serially correlated draws. Thus, a different approach is needed. An efficient alternative is to exploit the fact that the space of all structural parameters is equivalent to the product of the space of all reduced form parameters and $O(n)$. This mapping is given by $(\mathbf{B}, \mathbf{\Sigma}, \mathbf{Q}) \longmapsto (\mathbf{T}^{-1}\mathbf{Q}, \mathbf{B}\mathbf{T}^{-1}\mathbf{Q})$, where $\mathbf{\Sigma} = \mathbf{T}'\mathbf{T}$ is the Cholesky decomposition of $\mathbf{\Sigma}$ such that $\mathbf{T}$ is upper triangular with positive diagonal and $\mathbf{Q} \in O(n)$. Because $(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1})$ is observationally equivalent to $(\mathbf{T}^{-1}\mathbf{Q}, \mathbf{B}\mathbf{T}^{-1}\mathbf{Q})$, the likelihood is flat over the space of orthogonal matrices. It is also the case that the likelihood at the reduced form parameters $(\mathbf{B}, \mathbf{\Sigma})$ will be equal to the likelihood at the structural parameters $(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1})$. Given a prior on the reduced form parameters, this implies that the unrestricted posterior is the product of the posterior distribution of the reduced form parameters with the uniform distribution with respect to the Haar measure on $O(n)$. If the prior on the reduced form parameters is conjugate, then the posterior will have the multivariate normal inverse Wishart distribution. There are efficient algorithms for obtaining independent draws from this distribution. So all that remains to be determined is an efficient algorithm for drawing from the uniform distribution with respect to the Haar measure on $O(n)$. Canova and Nicoló (2002), Uhlig (2005), and Rubio-Ramírez et al. (2010) propose algorithms to draw from that set. However, Rubio-Ramírez et al. (2010) is the only computationally feasible for moderately large SVAR systems (e.g. $n > 4$).[1] Rubio-Ramírez et al. (2010)'s 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* $\mathbf{X}$.[2] *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). □

The previous discussion and Theorem 1 motivates us to modify the first step in Algorithm 1 applied to $(\mathbf{A}_0, \mathbf{A}_+) = (\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1})$ to obtain the following efficient Algorithm.

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* $\mathbf{Q}$.
3. *Keep the draw if* $\mathbf{S}_j f\left(\mathbf{T}^{-1}\mathbf{Q}, \mathbf{B}\mathbf{T}^{-1}\mathbf{Q}\right)\mathbf{e}_j > \mathbf{0}$ *are satisfied for* $1 \leq j \leq n$.
4. *Return to Step 1 until the required number of posterior draws satisfying the sign restrictions have been obtained.*

[1] See Rubio-Ramírez et al. (2010) for details.

[2] 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.

Theorem 1 and Algorithm 2 are replications of Theorem 9 and Algorithm 2 in Rubio-Ramírez et al. (2010). As in Rubio-Ramírez et al. (2010), instead of working with the reduced form parameters, one could work directly with the structural parameters, although, in general, this is harder.

At this point it is useful to understand how Theorem 1 and Algorithm 2 work, and more importantly, how they can be implemented recursively. For $1 \leq j \leq n$, let $\mathbf{x}_j = \mathbf{X}\mathbf{e}_j$ and $\mathbf{q}_j = \mathbf{Q}\mathbf{e}_j$. The $\mathbf{q}_j$ can be obtained recursively by

$$\mathbf{q}_j = \frac{\left(\mathbf{I}_n - \mathbf{Q}_{j-1}\mathbf{Q}'_{j-1}\right)\mathbf{x}_j}{\left|\left|\left(\mathbf{I}_n - \mathbf{Q}_{j-1}\mathbf{Q}'_{j-1}\right)\mathbf{x}_j\right|\right|} = \frac{\mathbf{N}_{j-1}\mathbf{N}'_{j-1}\mathbf{x}_j}{\left|\left|\mathbf{N}_{j-1}\mathbf{N}'_{j-1}\mathbf{x}_j\right|\right|} = \mathbf{N}_{j-1}\frac{\mathbf{N}'_{j-1}\mathbf{x}_j}{\left|\left|\mathbf{N}'_{j-1}\mathbf{x}_j\right|\right|} \text{ for } 1 \leq j \leq n,$$

where $||\ ||$ is the Euclidean metric, $\mathbf{Q}_{j-1} = \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_{j-1} \end{bmatrix}$, and $\mathbf{N}_{j-1}$ is any $n \times (n-j+1)$ matrix whose columns form an orthornormal basis for the null space of $\mathbf{Q}'_{j-1}$.[3] We follow the convention that $\mathbf{Q}_0$ is the the $n \times 0$ empty matrix, $\mathbf{Q}_0\mathbf{Q}'_0$ is the $n \times n$ empty matrix, and $\mathbf{N}_0$ is the $n \times n$ identity matrix. Geometrically, $\mathbf{q}_j$ is the projection of $\mathbf{x}_j$ onto the null space of $\mathbf{Q}'_{j-1}$ normalized to be of unit length. Alternatively, $\mathbf{N}'_{j-1}\mathbf{x}_j$ is a standard normal draw from $\mathbb{R}^{n-j+1}$ and $\mathbf{N}'_{j-1}\mathbf{x}_j / \parallel \mathbf{N}'_{j-1}\mathbf{x}_j \parallel$ 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 orthornormal, 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 Algorithm 2 produces 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}. While it is more efficient to implement Algorithm 2 in a single step via the QR decomposition, the fact that it can be implemented recursively will be of use when there are zero restrictions.[4]

2.3 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

[3] The formula just described to obtain $\mathbf{q}_j$ recursively for $1 \leq j \leq n$ implicitly imposes the normalization that the diagonal of R is positive.

[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 $\prod_{j=1}^{n} S^{n-j}$.

of zero restrictions associated with the j-th structural shock. The total number of zero restrictions will be $z = \sum_{j=1}^{n} z_j$.

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 $1 \leq j \leq n$.

We can no longer use Algorithm 2 to obtain draws satisfying both the sign and the zero restrictions since the set of structural parameters satisfying 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 of structural parameters satisfying the zero restrictions. Because of the same arguments used to motivate Algorithm 1, the set of structural parameters satisfying both the sign and the zero restrictions will be of positive measure in the set of structural parameters satisfying the zero restrictions. Thus, we will be able to use the following algorithm to make draws from the posterior parameters satisfying both the sign and the zero restrictions.

Algorithm 3.

1. *Draw $(\mathbf{A}_0, \mathbf{A}_+)$ from the posterior satisfying the zero restrictions.*
2. *Keep the draw if the sign restrictions are satisfied.*
3. *Return to Step 1 until the required number of posterior draws satisfying both the sign and the zero restrictions have been obtained.*

The only obstacle to implementing this procedure is an efficient technique for the first step. Given a prior on the reduced form parameters, the posterior satisfying the zero restrictions is the product of the posterior distribution of the reduced form parameters with the uniform distribution with respect to the Haar measure on $O(n)$ conditional on the zero restrictions holding. All that remains to be determined is an efficient Algorithm for drawing from the uniform distribution with respect to the Haar measure on $O(n)$ conditional on the zero restrictions holding.

The first crucial step is to note that the zero restrictions on the IRFs can be converted into linear restrictions on the columns of the orthogonal matrix $\mathbf{Q}$. To see this, let $(\mathbf{A}_0, \mathbf{A}_+)$ be any value of structural parameters. Note that for any orthogonal matrix $\mathbf{Q}$, we have

$$\mathbf{Z}_j f(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})\, \mathbf{e}_j = \mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{Q}\mathbf{e}_j = \mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{q}_j$$

for $1 \leq j \leq n$. Therefore, the zero restrictions associated with the j-th structural shock can be expressed as linear restrictions on the j-th column of $\mathbf{Q}$. Thus, the zero restrictions will hold if and only if

$$\mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{q}_j = \mathbf{0} \tag{4}$$

for $1 \leq j \leq n$. 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 $\mathbf{Q}$.

Using these two insights the next theorem shows when and how, given any value of the structural parameters $(\mathbf{A}_0, \mathbf{A}_+)$, we can find a $\mathbf{Q}$ such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions.

Theorem 2. *Let $(\mathbf{A}_0, \mathbf{A}_+)$ be any value of structural parameters. The structural parameters $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$, where $\mathbf{Q}$ is orthogonal, satisfy the zero restrictions if and only if $\| \mathbf{q}_j \| = 1$ and*

$$\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{q}_j = \mathbf{0}, \tag{5}$$

for $1 \leq j \leq n$, where

$$\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} \mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{Q}'_{j-1} \end{bmatrix}.$$

Furthermore, if the rank of $\mathbf{Z}_j$ is less than or equal to $n - j$, then there will be non-zero solutions of equation (5) for all values of $\mathbf{Q}_{j-1}$.

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 $\mathbf{Z}_j f(\mathbf{A}_0, \mathbf{A}_+)\, \mathbf{q}_j = \mathbf{0}$ and $\mathbf{Q}$ 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 $\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+)$ is less than or equal to $z_j + j - 1$. Thus, if $z_j \leq n - j$, then the rank of $\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+)$ will be strictly less than n and there will be non-zero solutions of equation (5). □

Whether or not there will be non-zero solutions of equation (5) for all values of $\mathbf{Q}_{j-1}$, clearly depends on the ordering of the equations (columns) of the original system, which is arbitrary. We shall only consider zero restrictions such that the equations of the original system can be ordered so that $z_j \leq n - j$. Because, when considering zero restrictions together with sign restrictions, one usually only wants to have a small number of zero restrictions, this condition will almost always be satisfied in practice. If it is the case that the system can be ordered so that $z_j \leq n - j$, then Theorem 2 implies that for any value $(\mathbf{A}_0, \mathbf{A}_+)$ of the structural parameters one can always find an orthogonal matrix $\mathbf{Q}$ such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions. This implies that zero restrictions impose no constraints on the reduced form parameters, but will impose constraints on the orthogonal matrix $\mathbf{Q}$.

As mentioned above, drawing $(\mathbf{A}_0, \mathbf{A}_+)$ from the posterior satisfying the zero restrictions is equivalent to drawing $(\mathbf{A}_0, \mathbf{A}_+)$ from the unrestricted posterior and making uniform draws $\mathbf{Q}$ from the set of orthogonal matrices such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions. The next Theorem shows how to use the results in Theorem 2 to do exactly that.

Theorem 3. *Let $1 \leq j \leq n$, let $\mathbf{Z}_j$ represent zero restrictions with the equations of the system given by (1) ordered so that $z_j \leq n - j$. Let $(\mathbf{A}_0, \mathbf{A}_+)$ be any value of the structural parameters. Let $\mathbf{Q}$ be obtained as follows.*

1. *Let $j = 1$.*
2. *Find a matrix $\mathbf{N}_{j-1}$ whose columns form an orthonormal basis for the null space of $\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+)$.*
3. *Draw $\mathbf{x}_j$ from the standard normal distribution on $\mathbb{R}^n$.*
4. *Let $\mathbf{q}_j = \mathbf{N}_{j-1}\left(\mathbf{N}'_{j-1}\mathbf{x}_j / \parallel \mathbf{N}'_{j-1}\mathbf{x}_j \parallel\right)$.*[5]
5. *If $j = n$ stop, otherwise let $j = j + 1$ and move to Step 2.*

Then, the random matrix $\mathbf{Q}$ has the uniform distribution with respect to the Haar measure on $O(n)$ such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions.

Proof. By Theorem 2, $\mathbf{Q}$ will be orthogonal and $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ will satisfy the zero restrictions. Let n_j be the number of columns in $\mathbf{N}_{j-1}$. Because $\mathbf{q}_j$, conditional on $\mathbf{Q}_{j-1}$, is a draw from the uniform distribution of the sphere whose dimension is $n_j - 1$, the distribution of $\mathbf{Q}$ will be uniform with respect to the Haar measure on $O(n)$ such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions. □

It should be clear from Theorem 3 that for each $(\mathbf{A}_0, \mathbf{A}_+)$ there are several $\mathbf{Q}$s such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions and that the particular $\mathbf{Q}$ to be drawn will depend

[5]Alternatively, we could draw $\mathbf{x}_j$ from the standard normal distribution on $\mathbb{R}^{n_j}$ and get $\mathbf{q}_j = \mathbf{N}_{j-1}\mathbf{x}_j / \parallel \mathbf{N}_{j-1}\mathbf{x}_j \parallel$, where n_j is the number of columns in $\mathbf{N}_{j-1}$, which is positive number. This implementation will results in an even faster procedure.

on the draw of $\mathbf{x}_j$ for $1 \leq j \leq n$. The fact that Theorem 3 allows us to draw $\mathbf{Q}$ from the uniform distribution with respect to the Haar measure on $O(n)$ such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions is key. This fact allows us to modify the first step in Algorithm 3 applied to $(\mathbf{A}_0, \mathbf{A}_+) = (\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1})$ to obtain posterior draws that satisfy both the zero and sign restrictions.

Algorithm 4.

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

2. *Use Theorem 3, applied to* $(\mathbf{A}_0, \mathbf{A}_+) = (\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1})$*, to draw an orthogonal matrix* $\mathbf{Q}$ *such that* $(\mathbf{T}^{-1}\mathbf{Q}, \mathbf{B}\mathbf{T}^{-1}\mathbf{Q})$ *satisfies the zero restrictions.*

3. *Keep the draw if* $\mathbf{S}_j f\left(\mathbf{T}^{-1}\mathbf{Q}, \mathbf{B}\mathbf{T}^{-1}\mathbf{Q}\right)\mathbf{e}_j > \mathbf{0}$ *are satisfied for* $1 \leq j \leq n$*.*

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

It is important to remember that, in the absence of sign restrictions to identify the j-th shock, we need to normalize the sign of $\mathbf{q}_j$ using a general normalization rule as described in Waggoner and Zha (2003b). From Algorithm 4 is also easy to see that, for each $(\mathbf{B}, \mathbf{\Sigma})$, there is a distribution of IRFs such that the restrictions hold. This is essential to interpret the results obtained when analyzing our applications in Sections 5 and 6.

Before presenting some examples of how to apply the algorithms presented in this section, 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 et al. (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 that our theorems and algorithms allow us to consider two additional class of restrictions (in addition to restrictions on IRFs). First, 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 et al. (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, 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 to compare our methodology to some existing methods of inference.

3 Examples

In this section we present two examples to illustrate how to use our theorems and algorithms. In both examples, 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 $\mathbf{Q}$ such that the sign restrictions hold while Algorithm 4 allows us to draw a $\mathbf{Q}$ such that both the sign and zero restrictions hold. The two examples will differ on their set of restrictions.

3.1 Example 1

Consider a five variables SVAR with one lag. The dimension and lag length of SVAR are totally arbitrary.

3.1.1 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

$$\mathbf{B} = \begin{bmatrix} 0.7577 & 0.7060 & 0.8235 & 0.4387 & 0.4898 \\ 0.7431 & 0.0318 & 0.6948 & 0.3816 & 0.4456 \\ 0.3922 & 0.2769 & 0.3171 & 0.7655 & 0.6463 \\ 0.6555 & 0.0462 & 0.9502 & 0.7952 & 0.7094 \\ 0.1712 & 0.0971 & 0.0344 & 0.1869 & 0.7547 \end{bmatrix}$$

and

$$
\mathbf{\Sigma} = \begin{bmatrix} 0.0281 & -0.0295 & 0.0029 & 0.0029 & 0.0024 \\ -0.0295 & 3.1850 & 0.0325 & -0.0105 & 0.0315 \\ 0.0029 & 0.0325 & 0.0067 & 0.0054 & 0.0030 \\ 0.0029 & -0.0105 & 0.0054 & 0.1471 & 0.0021 \\ 0.0024 & 0.0315 & 0.0030 & 0.0021 & 0.0140 \end{bmatrix}.
$$

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

$$
\mathbf{A}_0 = \begin{bmatrix} 5.9655 & 0.5911 & -1.4851 & -0.0035 & -0.4591 \\ 0 & 0.5631 & -0.1455 & 0.0321 & -0.0566 \\ 0 & 0 & 12.9098 & -2.2906 & -3.5385 \\ 0 & 0 & 0 & 2.6509 & 0.0072 \\ 0 & 0 & 0 & 0 & 8.9469 \end{bmatrix}
$$

and

$$
\mathbf{A}_+ = \begin{bmatrix} 4.5201 & 0.8454 & 9.4033 & -0.7034 & 1.0835 \\ 4.4330 & 0.4572 & 7.8615 & -0.5815 & 1.1879 \\ 2.3397 & 0.3878 & 3.4710 & 1.3104 & 4.4701 \\ 3.9104 & 0.4135 & 11.2867 & -0.0694 & 2.6867 \\ 1.0213 & 0.1559 & 0.1757 & 0.4192 & 6.5477 \end{bmatrix}.
$$

Assume that we want to impose restrictions at horizon zero, two, and infinity. These three IRFs are

$$
\mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} 0.1676 & 0 & 0 & 0 & 0 \\ -0.1760 & 1.7760 & 0 & 0 & 0 \\ 0.0173 & 0.0200 & 0.0775 & 0 & 0 \\ 0.0173 & -0.0042 & 0.0669 & 0.3772 & 0 \\ 0.0143 & 0.0192 & 0.0306 & -0.0003 & 0.1118 \end{bmatrix},
$$

$$
\mathbf{L}_2(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} 0.1468 & 2.1329 & 0.2138 & 0.5832 & 0.0522 \\ 0.0316 & 1.3934 & 0.0989 & 0.3142 & 0.0241 \\ 0.1447 & 2.2170 & 0.2294 & 0.6235 & 0.0473 \\ 0.1181 & 2.2576 & 0.2302 & 0.6779 & 0.0479 \\ 0.1405 & 2.5858 & 0.2838 & 0.7751 & 0.0952 \end{bmatrix},
$$

and

$$
\mathbf{L}_\infty(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} 0.1159 & -0.2625 & -0.0832 & -0.2330 & -0.0145 \\ -0.1149 & 1.3281 & -0.0594 & -0.2142 & -0.0044 \\ -0.0194 & -0.3461 & 0.0057 & -0.1048 & -0.0486 \\ -0.0449 & -0.9519 & 0.0389 & 0.2935 & -0.0268 \\ -0.0999 & -1.6985 & -0.0220 & -0.2832 & 0.2129 \end{bmatrix}.
$$

Then, the function $f(\mathbf{A}_0, \mathbf{A}_+)$ stacks these three IRFs as follows

$$
f(\mathbf{A}_0, \mathbf{A}_+) = \begin{bmatrix} \mathbf{L}_0(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_2(\mathbf{A}_0, \mathbf{A}_+) \\ \mathbf{L}_\infty(\mathbf{A}_0, \mathbf{A}_+) \end{bmatrix} = \begin{bmatrix} 0.1676 & 0 & 0 & 0 & 0 \\ -0.1760 & 1.7760 & 0 & 0 & 0 \\ 0.0173 & 0.0200 & 0.0775 & 0 & 0 \\ 0.0173 & -0.0042 & 0.0669 & 0.3772 & 0 \\ 0.0143 & 0.0192 & 0.0306 & -0.0003 & 0.1118 \\ 0.1468 & 2.1329 & 0.2138 & 0.5832 & 0.0522 \\ 0.0316 & 1.3934 & 0.0989 & 0.3142 & 0.0241 \\ 0.1447 & 2.2170 & 0.2294 & 0.6235 & 0.0473 \\ 0.1181 & 2.2576 & 0.2302 & 0.6779 & 0.0479 \\ 0.1405 & 2.5858 & 0.2838 & 0.7751 & 0.0952 \\ 0.1159 & -0.2625 & -0.0832 & -0.2330 & -0.0145 \\ -0.1149 & 1.3281 & -0.0594 & -0.2142 & -0.0044 \\ -0.0194 & -0.3461 & 0.0057 & -0.1048 & -0.0486 \\ -0.0449 & -0.9519 & 0.0389 & 0.2935 & -0.0268 \\ -0.0999 & -1.6985 & -0.0220 & -0.2832 & 0.2129 \end{bmatrix}.
$$

3.1.2 The Restrictions

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

$$\begin{aligned}
\mathbf{S}_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{S}_4 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{Z}_2 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{Z}_3 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Since there are no sign restrictions associated to the second, third, and fifth structural shocks, we do not need to specify $\mathbf{S}_2$, $\mathbf{S}_3$, and $\mathbf{S}_5$. The same is true for $\mathbf{Z}_1$, $\mathbf{Z}_4$, and $\mathbf{Z}_5$.

Sign Restrictions Next we enforce the sign restrictions using Algorithm 2. Assume that we draw

$$\mathbf{X} = \begin{bmatrix} 0.9848 & -0.7235 & -0.6087 & 1.4874 & 1.2671 \\ 0.3153 & -0.0684 & -1.7604 & 0.4111 & 0.2459 \\ -1.1912 & 1.5261 & -1.5656 & -1.8236 & -0.6564 \\ -0.5070 & -0.1686 & -0.1778 & -0.4927 & 1.6206 \\ -1.2656 & -0.9995 & 0.1588 & -1.1497 & 1.1970 \end{bmatrix},$$

where each element is drawn from an independent standard normal distribution. Then, the $\mathbf{Q}$

associated with the QR decomposition is

$$\mathbf{Q} = \begin{bmatrix} 0.4723 & -0.2394 & -0.4351 & 0.6775 & -0.2668 \\ 0.1512 & 0.0099 & -0.8032 & -0.5492 & 0.1741 \\ -0.5713 & 0.6350 & -0.3571 & 0.3061 & -0.2217 \\ -0.2432 & -0.1638 & -0.1180 & 0.3784 & 0.8700 \\ -0.6070 & -0.7159 & -0.1554 & -0.0493 & -0.3041 \end{bmatrix}.$$

Note that for this particular draw of $\mathbf{Q}$ the sign restrictions are satisfied

$$\mathbf{S}_1 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_1 = 0.0792 > 0,$$

and

$$\mathbf{S}_4 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_4 = 0.8156 > 0.$$

However, since the set of structural parameters satisfying the zero restrictions is of measure zero in the set of all structural parameters, there is no reason to expect the zero restrictions to be satisfied for such $\mathbf{Q}$. As it can be seen, in this case they do not hold,

$$\mathbf{Z}_2 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_3 = 0.0413 \neq 0,$$

and

$$\mathbf{Z}_3 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_5 = -0.0499 \neq 0.$$

Sign and Zero Restrictions We now illustrate how to find a $\mathbf{Q}$ that satisfies sign and zero restrictions based on Algorithm 4. Let us assume that in Step 1 we use our draw from the posterior of the reduced form parameters and, hence, our draw of the structural parameters. Then, Step 2 of Algorithm 4 is as follows.

1. Let $j = 1$.

2. Find a matrix $\mathbf{N}_{j-1}$ whose columns form an orthonormal basis for the null space of $\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+)$

$$\mathbf{N}_0 = \mathbf{I}_5.$$

3. Draw $\mathbf{x}_j$ from the standard normal distribution on $\mathbb{R}^n$

$$\mathbf{x}_1 = \begin{bmatrix} 1.0347 & 0.7269 & -0.3034 & 0.2939 & -0.7873 \end{bmatrix}'.$$

4. Let $\mathbf{q}_j = \mathbf{N}_{j-1}\left(\mathbf{N}'_{j-1}\mathbf{x}_j / \parallel \mathbf{N}'_{j-1}\mathbf{x}_j \parallel\right)$

$$\mathbf{q}_1 = \begin{bmatrix} 0.6683 & 0.4695 & -0.1960 & 0.1898 & -0.5085 \end{bmatrix}'.$$

5. If $j = n$ stop, otherwise let $j = j + 1$ and move to Step 2.

Thus, if we repeat the same steps until j equals 5, we get the following matrices.

$$\mathbf{N}_1 = \begin{bmatrix} 0.1752 & -0.3153 & 0.5347 \\ 0.0661 & 0.1248 & 0.0675 \\ 0.9788 & 0.0074 & -0.0477 \\ 0.0458 & 0.9395 & 0.1278 \\ -0.0689 & 0.0487 & 0.8312 \end{bmatrix}, \mathbf{N}_2 = \begin{bmatrix} 0.0503 & -0.6248 \\ -0.1998 & 0.7253 \\ -0.9299 & -0.1600 \\ 0.1142 & 0.2408 \\ 0.2825 & 0.0001 \end{bmatrix},$$

$$\mathbf{N}_3 = \begin{bmatrix} -0.1312 & 0.5582 \\ -0.5375 & -0.4404 \\ 0.1477 & 0.5736 \\ 0.6669 & -0.4043 \\ -0.4767 & -0.0449 \end{bmatrix}, \text{ and } \mathbf{N}_4 = \begin{bmatrix} 0.5726 \\ -0.3391 \\ 0.5387 \\ -0.5151 \\ 0.0395 \end{bmatrix}.$$

$$\mathbf{x}_2 = \begin{bmatrix} 0.8884 \\ -1.1471 \\ -1.0689 \\ -0.8095 \\ -2.9443 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 1.4384 \\ 0.3252 \\ -0.7549 \\ 1.3703 \\ -1.7115 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} -0.1022 \\ -0.2414 \\ 0.3192 \\ 0.3129 \\ -0.8649 \end{bmatrix}, \text{ and } \mathbf{x}_5 = \begin{bmatrix} -0.0301 \\ -0.1649 \\ 0.6277 \\ 1.0933 \\ 1.1093 \end{bmatrix}.$$

$$\mathbf{q}_2 = \begin{bmatrix} -0.3224 \\ -0.1382 \\ -0.2651 \\ -0.5962 \\ -0.6717 \end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0.3473 \\ -0.5269 \\ -0.7352 \\ -0.0170 \\ 0.2469 \end{bmatrix}, \mathbf{q}_4 = \begin{bmatrix} -0.0311 \\ -0.6065 \\ 0.2461 \\ 0.5856 \\ -0.4772 \end{bmatrix}, \text{ and } \mathbf{q}_5 = \begin{bmatrix} -0.5726 \\ 0.3391 \\ -0.5387 \\ 0.5151 \\ -0.0395 \end{bmatrix}.$$

By the theorems introduced in Section 2, the resulting $\mathbf{Q}$ is an uniform draw from the set of orthogonal matrices such that $(\mathbf{A}_0\mathbf{Q}, \mathbf{A}_+\mathbf{Q})$ satisfies the zero restrictions. In this case, the sign restrictions also hold

$$\mathbf{S}_1 f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_1 = 0.1120 > 0,$$

and

$$\mathbf{S}_4 f\left(\mathbf{A}_0, \mathbf{A}_+\right)\mathbf{q}_4 = 0.9501 > 0.$$

Of course, the fact that the sign restrictions hold depends on the draw of $\mathbf{x}_j$ for $1 \leq j \leq n$.

3.2 Example 2

Consider the SVAR and the reduced form parameters $\mathbf{B}$ and $\mathbf{\Sigma}$ from Example 1. In addition, for easy of exposition let us consider sign and zero restrictions that are imposed at the same horizons as in Example 1.

3.2.1 The Restrictions

Let us consider a different set of restrictions than in Example 1. 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 infinite on the response of the second variable to the fifth structural shock, a zero restriction at horizon zero on the response of the first variable to the first 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 two on the response of the fifth variable to the fourth structural shock. These restrictions can be enforced using the matrices $\mathbf{S}_j$ and $\mathbf{Z}_j$ for $1 \leq j \leq n$

$$
\begin{aligned}
\mathbf{S}_2 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{S}_3 &= \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{S}_5 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{Z}_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\
\mathbf{Z}_4 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Since there are no sign restrictions associated to the first and fourth structural shocks, we do not need to specify $\mathbf{S}_1$, and $\mathbf{S}_4$. Similarly, we do not specify $\mathbf{Z}_2$, $\mathbf{Z}_3$, and $\mathbf{Z}_5$.

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

$$
\mathbf{X} = \begin{bmatrix} 0.9195 & 0.1651 & 0.7871 & 0.0329 & 0.3847 \\ -0.2499 & -0.4216 & -0.4650 & -1.8634 & 1.0269 \\ -0.2079 & 0.2769 & 1.3521 & -0.2368 & -1.3322 \\ 0.9978 & 1.3410 & -0.2697 & 0.0062 & 0.4697 \\ -0.0693 & 1.7345 & 0.8953 & -0.2012 & 0.0055 \end{bmatrix},
$$

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

$$
\mathbf{Q} = \begin{bmatrix} 0.6582 & -0.2495 & 0.5362 & -0.1878 & 0.4263 \\ -0.1789 & -0.1192 & -0.2003 & -0.9551 & 0.0375 \\ -0.1488 & 0.2124 & 0.7167 & -0.1734 & -0.6238 \\ 0.7143 & 0.3059 & -0.3889 & -0.1094 & -0.4827 \\ -0.0496 & 0.8859 & 0.0866 & -0.1021 & 0.4413 \end{bmatrix}.
$$

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

$$
\begin{aligned}
\mathbf{S}_2 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_2 &= \begin{bmatrix} 0.0190 & 0.0002 \end{bmatrix}' > \mathbf{0} \\
\mathbf{S}_3 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_3 &= 0.4500 > 0 \\
\mathbf{S}_5 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_5 &= 0.1394 > 0.
\end{aligned}
$$

As in Example 1, there is no reason to expect the zero restrictions to be satisfied for such $\mathbf{Q}$. Indeed, in this case they do not hold,

$$
\mathbf{Z}_1 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_1 = \begin{bmatrix} 0.1103 & -0.0037 \end{bmatrix}' \neq \mathbf{0},
$$

and

$$
\mathbf{Z}_4 f\left(\mathbf{A}_0, \mathbf{A}_+\right) \mathbf{q}_4 = -0.0377 \neq 0.
$$

Sign and Zero Restrictions We now illustrate how to find a $\mathbf{Q}$ that satisfies sign and zero restrictions based on Algorithm 4. As in Example 1, assume that in Step 1 we use our draw from the posterior of the reduced form parameters and, hence, our draw of the structural parameters. Then, Step 2 of Algorithm 4 is as follows.

1. Let $j = 1$.

2. Find a matrix $\mathbf{N}_{j-1}$ whose columns form an orthonormal basis for the null space of $\mathbf{R}_j(\mathbf{A}_0, \mathbf{A}_+)$

$$\mathbf{N}_0 = \begin{bmatrix} 0 & -0.9682 & 0.2502 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 \\ 0 & 0 & 0 & 0 & 1.0000 \end{bmatrix}'.$$

3. Draw $\mathbf{x}_j$ from the standard normal distribution on $\mathbb{R}^n$

$$\mathbf{x}_1 = \begin{bmatrix} 0.3409 & -0.5418 & 1.5292 & 0.3320 & -0.4429 \end{bmatrix}'.$$

4. Let $\mathbf{q}_j = \mathbf{N}_{j-1}\left(\mathbf{N}'_{j-1}\mathbf{x}_j / \parallel \mathbf{N}'_{j-1}\mathbf{x}_j \parallel\right)$.

$$\mathbf{q}_1 = \begin{bmatrix} 0 & -0.8265 & 0.2135 & 0.3124 & -0.4168 \end{bmatrix}'.$$

5. If $j = n$ stop, otherwise let $j = j + 1$ and move to Step 2.

Thus, if we repeat the these steps until j equals 5, we get the following matrices.

$$\mathbf{N}_1 = \begin{bmatrix} 0.8265 & -0.2135 & -0.3124 & 0.4168 \\ 0.3169 & 0.1765 & 0.2582 & -0.3445 \\ 0.1765 & 0.9544 & -0.0667 & 0.0890 \\ 0.2582 & -0.0667 & 0.9024 & 0.1302 \\ -0.3445 & 0.0890 & 0.1302 & 0.8263 \end{bmatrix}, \mathbf{N}_2 = \begin{bmatrix} 0.4549 & -0.1767 & 0.8185 \\ 0.4177 & 0.3072 & -0.1995 \\ 0.6276 & -0.1331 & -0.1074 \\ 0.0468 & 0.9254 & 0.1984 \\ -0.4718 & 0.0164 & 0.4893 \end{bmatrix},$$

$$\mathbf{N}_3 = \begin{bmatrix} -0.6323 \\ -0.3924 \\ -0.5271 \\ -0.2887 \\ 0.2917 \end{bmatrix}, \text{ and } \mathbf{N}_4 = \begin{bmatrix} 0.6092 \\ -0.3678 \\ 0.0459 \\ -0.6484 \\ 0.2668 \end{bmatrix}.$$

$$\mathbf{x}_2 = \begin{bmatrix} -0.4423 \\ 2.0019 \\ 0.5116 \\ -0.7100 \\ 1.9563 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0.2203 \\ -0.1524 \\ 0.0247 \\ 0.7181 \\ 1.0279 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 0.3112 \\ 1.2880 \\ 0.2050 \\ -0.3948 \\ -1.0959 \end{bmatrix}, \text{ and } \mathbf{x}_5 = \begin{bmatrix} -0.3828 \\ -0.3661 \\ 0.3669 \\ 0.2647 \\ 0.8716 \end{bmatrix}.$$

$$\mathbf{q}_2 = \begin{bmatrix} -0.3033 \\ -0.0908 \\ 0.7289 \\ 0.0664 \\ 0.6034 \end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0.3704 \\ -0.1394 \\ -0.3783 \\ 0.6279 \\ 0.5532 \end{bmatrix}, \mathbf{q}_4 = \begin{bmatrix} 0.6323 \\ 0.3924 \\ 0.5271 \\ 0.2887 \\ -0.2917 \end{bmatrix}, \text{ and } \mathbf{q}_5 = \begin{bmatrix} -0.6092 \\ 0.3678 \\ -0.0459 \\ 0.6484 \\ -0.2668 \end{bmatrix}.$$

In this case, the sign restrictions also hold

$$\begin{aligned} \mathbf{S}_2 f(\mathbf{A}_0, \mathbf{A}_+) \mathbf{q}_2 &= \begin{bmatrix} 0.0082 & 0.0008 \end{bmatrix}' > \mathbf{0} \\ \mathbf{S}_3 f(\mathbf{A}_0, \mathbf{A}_+) \mathbf{q}_3 &= 0.3127 > 0 \\ \mathbf{S}_5 f(\mathbf{A}_0, \mathbf{A}_+) \mathbf{q}_5 &= 0.4235 > 0. \end{aligned}$$

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

4 The Mountford and Uhlig (2009) Methodology

In this section, we discuss the penalty function approach with sign and zero restrictions developed by Mountford and Uhlig (2009). First, we describe the methodology. Second, we highlight how the procedure selects one particular orthogonal matrix $\mathbf{Q}$ instead of drawing from the conditional uniform distribution derived in Subsection 2.3. We also analyze the consequences of this drawback. Third, we formally show how selecting a particular orthogonal matrix $\mathbf{Q}$ may impose additional sign restrictions on variables that are seemingly unrestricted.

In sections 5 and 6, we will analyze two applications that have been previously studied in the literature using the penalty function approach. First, the effects of optimism shocks as a source of business cycle fluctuations studied by Beaudry et al. (2011). Second, the effects of fiscal policy shocks on economic activity studied by Mountford and Uhlig (2009). We will compare their reported outcomes (that we replicate in this paper) with the results obtained using our methodology. The comparison will allow us to show how picking a single orthogonal matrix $\mathbf{Q}$ instead of drawing from their conditional uniform distribution generates artificially narrow confidence intervals and introduces bias on the IRFs and other statistics of interest.

4.1 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 to identify the j-th structural shock. Let Z_j and $f(\mathbf{A}_0, \mathbf{A}_+)$ denote the latter. The penalty function approach finds an orthogonal matrix $\bar{\mathbf{Q}}^* = \begin{bmatrix} \bar{\mathbf{q}}_1^* & \cdots & \bar{\mathbf{q}}_n^* \end{bmatrix}$ such that IRFs come close to satisfying the sign restrictions, conditional on the zero restrictions being satisfied, according to a loss function.[6] In particular, for $1 \leq j \leq n$, this approach solves the following optimization problem

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

subject to

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

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

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 the $n\times 0$ empty matrix.[7]

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. There are very efficient algorithms for obtaining independent draws from this distribution, hence, normally the research will use the above algorithm where $(\mathbf{A}_0,\mathbf{A}_+)=(\mathbf{T}^{-1},\mathbf{B}\mathbf{T}^{-1})$.

4.2 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 4 uses this result to draw from the uniform distribution of orthogonal matrices conditional on the zero restrictions being satisfied. The penalty function approach abstracts from using the result. Instead, given any draw of the reduced form parameters, $(\mathbf{B},\mathbf{\Sigma})$, 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.

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

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 holding. 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 variables to a particular shock are correlated. Then, by choosing a particular orthogonal matrix that maximizes the response of a set of variables to the shock by minimizing the loss function, we are biasing the response of the other variables to the same shock. In that sense, the penalty function approach will behave as if there were additional sign restrictions on variables that are seemingly unrestricted. In general, it is hard to formally prove such 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 to understand 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 one of the most attractive features 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.

4.3 Is the Penalty Function Approach Truly Agnostic?

We now formally show how the penalty function approach can impose additional sign constraints 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.[8] 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 only identifying 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 $\mathbf{e}_2' \mathbf{L}_0 \left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j = \mathbf{e}_2' \mathbf{T}' \bar{\mathbf{q}}_j = \sum_{s=1}^{2} \mathbf{t}_{s,2} \bar{\mathbf{q}}_{s,j}$, 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 $f\left(-\frac{\mathbf{e}_2' \mathbf{L}_0 \left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j}{\sigma_2}\right) = -100 \frac{\mathbf{t}_{2,2} \bar{\mathbf{q}}_{2,j}}{\sigma_2}$; else $f\left(-\frac{\mathbf{e}_2' \mathbf{L}_0 \left(\mathbf{T}^{-1}, \mathbf{B}\mathbf{T}^{-1}\right) \bar{\mathbf{q}}_j}{\sigma_2}\right) = -\frac{\mathbf{t}_{2,2} \bar{\mathbf{q}}_{2,j}}{\sigma_2}$. Since, $\bar{\mathbf{q}}_{1,j}^* = 0$, and $\bar{\mathbf{q}}^*$ must be belong to S, it is straightforward to verify that the criterion function is minimized at $\bar{\mathbf{q}}_j^* = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}'$.

If the penalty function approach 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 penalty function approach 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,

[8] 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 restrictions to simplify the argument.

the response at horizon zero of the i-th variable to the j-th structural shock for $i > 2$ is

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

Thus, if $\mathbf{t}_{2,i} > 0$ ($\mathbf{t}_{2,i} < 0$) the penalty function approach imposes an additional positive (negative) sign restriction on the response of the i-th variable to the j-th structural shock at horizon zero.

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 5.3.1 we will show how to implement those restrictions in the case of optimism shocks.

5 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 et al. (2011) using the penalty function approach. We show how using our methodology to correct the shortcomings of the penalty function approach affects inference dramatically and has important implications for the economic interpretation of the results. We also report how our methodology is much faster than the penalty function approach. In the next section, we will do the same for fiscal policy shocks.

Let us begin revisiting Beaudry et al. (2011). The aim of Beaudry et al. (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 by exploiting sign and zero restrictions as identification strategies to isolate optimism shocks using the penalty function approach. After replicating their results, we repeat their empirical exercises using our methodology to show how their main economic conclusion substantially changes once one accounts for the shortcomings of the penalty function approach. While Beaudry et al. (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 methodology, it is very hard to support such 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.

5.1 Data and Identification Strategy

Beaudry et al. (2011) use two datasets. In the first one, they use data on total factor productivity (TFP), stock price, consumption, real federal funds rate, and hours worked. In the second one, they add investment and output. In both datasets, they consider three identification strategies. The strategies are described in Table 1.

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

Table 1: Identification Schemes Defined in Beaudry et al. (2011)

Identification 1 is the benchmark, where optimism shocks (sometimes called bouts of optimism) are identified as innovations that affect positively the stock prices and that are 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 8.1 gives details on the priors and the datasets.

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 datasets. The matrices $\mathbf{S}$s and $\mathbf{Z}$s are a function of the number of variables used in the SVAR. In the smaller dataset, 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 dataset, when seven variables are used, the corresponding $\mathbf{S}$s matrices are

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

for Identifications 1, 2, and 3 respectively, while the $\mathbf{Z}$ matrix is $\mathbf{Z}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

5.2 IRFs

We first show replications of the IRFs reported in Beaudry et al. (2011) using the penalty function approach. Then, we analyze how the results change once we use our methodology. Sometimes we will label our methodology as the ARRW methodology. Panel (a) in Figure 1 shows the IRFs of TFP, stock price, consumption, federal funds rate, and hours worked under Identification 1 when using the penalty function approach on the first dataset. This panel replicates the first block of Figure 1 in Beaudry et al. (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 (c) in Figure 1 show the IRFs of TFP, stock price, consumption, federal funds rate, and hours worked under Identifications 2 and 3. These panels replicate the second and third blocks of Figure 1 in Beaudry et al. (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 shape 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 1 are robust to extending the number of variables. Figure 3 show the results when we consider the larger dataset. Consumption, hours worked, investment, and output have positive and long-lasting responses to optimism shocks identified as in Beaudry et al. (2011). Lack of response of these variables is clearly outside the 68 percent confidence intervals.

This is true for the three identification schemes.

Once we use the ARRW methodology, the results reported in Beaudry et al. (2011) basically disappear. Panel (a) in Figure 2 reports the results for the first dataset using the ARRW methodology under Identification 1. There are three important differences with the results reported in Beaudry et al. (2011). First, the penalty function approach 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 penalty function approach. As a consequence, using the penalty function approach, there is a upward bias in the IRFs and artificially narrow confidence intervals.

We need to consider Identifications 2 and 3 (see Panels (b) and (c)), that force consumption to increase after an optimism shock, to find moderate evidence of a 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 et al. (2011). As reported in Figure 4, 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 imposing this extra positive sign restriction, the results under the ARRW methodology are much weaker. The sharp results reported in Beaudry et al. (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 penalty function approach. Once we use the ARRW methodology to solve these two problems the results disappear.

5.2.1 Understanding the Bias and the Narrow Confidence Intervals

We now shed some light on the issues related to the upward bias and the artificially narrow confidence intervals. Let us begin by the upward bias. Figures 5 and 6 plot the median IRFs and the 68 percent confidence intervals obtained using the ARRW methodology and compares them with the median IRFs obtained using the penalty function approach. Figures 5 plots the IRFs for the five variables SVAR and Figure 6 does the same for the larger SVAR. In both cases the median IRFs constructed using the penalty function approach are close to the 84-th percentile band constructed using the ARRW methodology. It is easy to observe that the penalty function approach 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 penalty function approach is also inducing a large response of consumption and hours worked because the three responses are positively correlated. For the five variables 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 0.27. In the case of hours worked is 0.10. The correlations are 0.21 and 0.12 in the the larger SVAR.

Let us now consider the artificially narrow confidence intervals. We have repeated several times that the penalty function approach selects a single orthogonal matrix instead of drawing from their conditional uniform distribution. As mentioned when describing Algorithm 4, 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 penalty function approach takes a single IRF from such a distribution. Figure 7 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 intervals are constructed using a single value of the structural parameters (the one obtained from the Cholesky decomposition and the OLS point estimate of the reduced form parameters) and several draws of the conditional uniform distribution of orthogonal matrix $\mathbf{Q}$. We have generated these draws repeating steps 2 and 3 of Algorithm 4 for the single value of the structural parameters. The probability intervals are compared with the single IRFs obtained with the penalty function approach evaluated at the same value of the reduced form parameters (i.e, the OLS point estimate).[9] In panel (a) we describe the results for the five variables SVAR. The dashed line shows the value of the IRFs resulting from the penalty function approach. The shadow area describes the 68 percent probability intervals obtained with our methodology. No uncertainty is considered when the penalty function approach 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 penalty function approach are close to the 84 percent band. Hence, once again, we can see how the penalty function approach picks a large response of stock prices and there is an upward bias in the response of consumption and hours worked. In panel (b) we show that this analysis also holds when we use the seven variables SVAR.

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

The fact that the Mountford and Uhlig (2009) methodology does not consider the distribution of IRFs is behind the narrower confidence intervals that Beaudry et al. (2011) report.

We can summarize our findings in Figure 8. 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 penalty function approach and the ARRW methodology.[10] The posterior distributions are approximated using a kernel smoothing function.[11] We have focused in the five variables SVAR. Similar results apply to the seven variables SVAR. Column 1 displays the results for Identification 1. Comparing the penalty function approach and the ARRW methodology we reach the following conclusions. First, the posterior distribution of the IRFs of stock prices obtained using the penalty function approach is center 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 penalty function approach dramatically underestimates the variance of the posterior distribution of IRFs of stock prices. These two results were expected since the penalty function approach is maximizing 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 of the IRFs for consumption and hours worked. Columns 2 and 3 show results for Identifications 2 and 3. In both cases, we reach the same conclusion. The posterior distributions IRFs at horizon zero for stock prices, consumption, and hours worked are artificially compressed and upward biased when computed using the penalty function approach.

	The Penalty Function Approach			The ARRW Methodology		
	Mean	Std dev	$\Pr(\cdot < 0)$	Mean	Std dev	$\Pr(\cdot < 0)$
Consumption	0.1043	0.0264	0	0.0413	0.1900	0.4160
Hours Worked	0.0717	0.0397	0.0360	0.0199	0.2861	0.4700

Table 2: Posterior Probabilities of Negative IRFs at horizon zero: Five Variables SVAR and Identification 1

We now show that the artificially compressed and upward biased posterior distributions of IRFs affect the signs of the IRFs of variables that are seemingly unrestricted. Let us first focus on the five variables VAR and Identification 1. Table 2 compares the posterior probabilities that the IRFs

[10] 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 only use horizon 0.

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

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 (2009) methodology while it is negative approximately 40 percent of the time under the ARRW methodology. The same is basically true for hours worked. Hence, the penalty function approach imposes additional positive sign restrictions on variables that are seemingly unrestricted. These can also be seen by comparing the mean responses reported in Table 2. 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 penalty function approach. This is, of course, related to the fact that confidence intervals are wider when using the ARRW methodology.

	The Penalty Function Approach			The ARRW Methodology		
	Mean	Std dev	$\Pr(\cdot < 0)$	Mean	Std dev	$\Pr(\cdot < 0)$
Hours Worked	0.1331	0.0389	0	0.0695	0.2794	0.4110

Table 3: Posterior Probabilities of Negative IRFs at horizon zero: Five Variables SVAR and Identification 2

	The Penalty Function Approach			The ARRW Methodology		
	Mean	Std dev	$\Pr(\cdot < 0)$	Mean	Std dev	$\Pr(\cdot < 0)$
Hours Worked	0.1084	0.0379	0.0040	0.0723	0.2860	0.4100

Table 4: Posterior Probabilities of Negative IRFs at horizon zero: Five Variables SVAR and Identification 3

Tables 3 and 4 repeat the exercise for Identifications 2 and 3. As we can see, the IRF at horizon zero of hours worked under Identification 2 is never negative using the penalty function approach while they are negative in approximately 40 percent of the draws using our methodology. Hence, the penalty function approach also introduces additional positive sign restrictions on hours worked in the case of Identification 2. Something similar happens in the case of identification 3. The IRF at horizon zero of hours worked is almost never negative using the penalty function approach while it is negative in 40 percent of the draws using our methodology.

5.3 FVE

Let us now focus on the FVE. Let us first analyze the SVAR with the smaller dataset. We compare the contribution of optimism shocks to the FEV obtained using the ARRW methodology and the penalty function approach. For easy of exposition, in Table 5 we focus on the contributions to the FEV at horizon 40.[12]

The ARRW Methodology			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.09 [0.02 , 0.23]	0.11 [0.03 , 0.26]	0.16 [0.05 , 0.31]
Stock Price	0.15 [0.03 , 0.47]	0.25 [0.06 , 0.56]	0.29 [0.08 , 0.60]
Consumption	0.15 [0.02 , 0.49]	0.27 [0.05 , 0.59]	0.38 [0.12 , 0.67]
Real Interest Rate	0.19 [0.08 , 0.43]	0.20 [0.08 , 0.44]	0.22 [0.08 , 0.47]
Hours Worked	0.17 [0.04 , 0.46]	0.24 [0.07 , 0.56]	0.30 [0.08 , 0.59]
The Penalty Function Approach			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.18 [0.09 , 0.31]	0.23 [0.10 , 0.40]	0.28 [0.12 , 0.44]
Stock Price	0.73 [0.56 , 0.86]	0.72 [0.57 , 0.82]	0.58 [0.42 , 0.71]
Consumption	0.26 [0.14 , 0.42]	0.71 [0.55 , 0.84]	0.76 [0.58 , 0.87]
Real Interest Rate	0.14 [0.08 , 0.23]	0.13 [0.07 , 0.22]	0.36 [0.28 , 0.44]
Hours Worked	0.31 [0.20 , 0.43]	0.63 [0.49 , 0.74]	0.50 [0.35 , 0.63]

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 15 and 17 percent respectively. In contrast, using the penalty function approach 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 is 27 and 24 percent using our methodology, but it is equal to 71 and 63 percent using the penalty function approach. Identification 3 yields the highest contribution of optimism shocks to the FEV of consumption and hours, 38 and 30 percent respectively, when using our methodology. However, these values are moderate compared to the 76 and 50 percent that we found when using the penalty function approach. Table 5 also reports the 68 percent confidence

[12]Table 13 in the Appendix 8.3 reports the contributions to the FEV at additional horizons.

intervals. As it 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 ARRW Methodology			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.08 [0.03 , 0.21]	0.10 [0.03 , 0.24]	0.13 [0.04 , 0.28]
Stock Price	0.11 [0.03 , 0.31]	0.17 [0.04 , 0.39]	0.22 [0.06 , 0.43]
Consumption	0.10 [0.02 , 0.35]	0.15 [0.03 , 0.40]	0.22 [0.04 , 0.48]
Real Interest Rate	0.13 [0.05 , 0.28]	0.13 [0.05 , 0.27]	0.14 [0.06 , 0.30]
Hours Worked	0.10 [0.03 , 0.32]	0.14 [0.04 , 0.36]	0.15 [0.04 , 0.36]
Investment	0.12 [0.04 , 0.29]	0.16 [0.06 , 0.35]	0.17 [0.06 , 0.36]
Output	0.10 [0.02 , 0.32]	0.17 [0.04 , 0.39]	0.23 [0.06 , 0.44]
Labor Productivity	0.09 [0.02 , 0.29]	0.13 [0.03 , 0.34]	0.19 [0.06 , 0.43]
The Penalty Function Approach			
	Identification 1	**Identification 2**	**Identification 3**
Adjusted TFP	0.16 [0.07 , 0.28]	0.22 [0.09 , 0.38]	0.30 [0.14 , 0.46]
Stock Price	0.52 [0.36 , 0.70]	0.62 [0.48 , 0.74]	0.57 [0.43 , 0.69]
Consumption	0.14 [0.06 , 0.28]	0.55 [0.40 , 0.70]	0.61 [0.43 , 0.76]
Real Interest Rate	0.10 [0.05 , 0.17]	0.06 [0.03 , 0.13]	0.26 [0.20 , 0.34]
Hours Worked	0.16 [0.09 , 0.28]	0.40 [0.27 , 0.54]	0.31 [0.18 , 0.45]
Investment	0.26 [0.16 , 0.38]	0.46 [0.35 , 0.58]	0.40 [0.26 , 0.54]
Output	0.24 [0.13 , 0.38]	0.60 [0.46 , 0.72]	0.60 [0.45 , 0.73]
Labor Productivity	0.24 [0.11 , 0.38]	0.37 [0.17 , 0.56]	0.49 [0.28 , 0.65]

Table 6: Share of FEV Attributable to Optimism Shocks at Horizon 40. Seven Variable SVAR

The results for the SVAR with the larger dataset are reported in Table 6.[13] 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 are 10, 17, and 23 percent under Identifications 1, 2, and 3, respectively. In any case, these values are remarkably lower than the ones found using the penalty function approach: 24, 60, and 60 percent respectively. As before, confidence intervals are much wider when using the ARRW methodology.

Summarizing, using the ARRW methodology is easy to conclude that optimism shocks explain a very small share of the FEV of any variable in the SVAR. This conclusion contrast with the results obtained using the penalty function approach. As it was the case with the IRFs, the penalty function approach 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 et al. (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 only reported results for horizon 40 but Appendix 8.3 shows that these conclusions are true at any horizon.

5.3.1 Replicating the Penalty Function Approach using the ARRW Methodolgy

In this subsection, we show that in the case of Identification 1 the penalty function approach in Beaudry et al. (2011) can be replicated using the ARRW methodology by considering some additional restrictions on the orthogonal matrix $\mathbf{Q}$.[14] These additional restrictions on the orthogonal matrix are behind the upward bias and the artificially narrow confidence intervals reported above.

Consider Identification 1 and note that there exists a closed form solution to the minimization problem embedded in the penalty function approach.[15] Specifically, the penalty function is minimized when the first column of the orthogonal matrix equals $\bar{\mathbf{q}}_1^* = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}'$.[16] Thus, we can replicate Beaudry et al. (2011) using our methodology by imposing zero constraints on the first column of the orthogonal matrix $\mathbf{Q}$. This can be done defining

[13] As before, we only report results at horizon 40. Table 14 in the Appendix 8.3 reports additional horizons.

[14] In fact, this is true for any identification that fulfills the conditions stated in Subsection 4.3.

[15] This has been shown in Subsection 4.3.

[16] In this case, there is only one shock to be identified, hence we set $j = 1$.

$$f\left(\mathbf{A}_0, \mathbf{A}_+\right) = \begin{bmatrix} \mathbf{L}_0\left(\mathbf{A}_0, \mathbf{A}_+\right) \\ \mathbf{I}_n \end{bmatrix},$$

where $\mathbf{I}_n$ allows us to put the zero constraints on the orthogonal matrix $\mathbf{Q}$. We only present results for the smaller dataset, but it is straightforward to extend them to the larger dataset. We define the matrices $\mathbf{S}$ 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 5.1, but the $\mathbf{Z}$ matrix has changed to reflect the additional restrictions on the first column of the orthogonal matrix $\mathbf{Q}$.

Panel (b) in Figure 9 plots the IRFs of TFP, stock price, consumption, federal funds rate, and hours worked under Identification 1 using 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 1 that are reprinted in Panel (a) of Figure 9. Thus, we can conclude that, in some circumstances, the Mountford and Uhlig (2009) methodology can be seen as a special case of ours, where additional restrictions on the orthogonal matrix are considered.

Summarizing, in this subsection we have shown that in the case of optimism shocks (under Identification 1) the Mountford and Uhlig (2009) methodology is a particular case of our methodology with additional restrictions on the first column of $\mathbf{Q}$. In that sense, it is natural that the penalty function approach produces artificially narrow confidence intervals and biased IRFs. Unfortunately, we can not show that the penalty function approach is a particular case of our methodology for all identification schemes. Nevertheless, it should be clear that the Mountford and Uhlig (2009) methodology always introduces additional restrictions (though not always can be mapped into our methodology) that create artificially narrow confidence intervals and may introduce bias.

5.4 Computational Time

Our methodology is faster than the penalty function approach. Table 7 reports the results for the case of optimism shocks. In the case of the 5 variables SVAR, the penalty function approach is approximately 24 times slower than our methodology using identification 1, and it is approximately 10 times slower than our methodology using identification 3. Similar results can be found in the case of the seven variables SVAR.

	The Penalty Function Approach	The ARRW Methodology	Ratio
Five variables SVAR			
Identification 1	792.74	32.89	24.10
Identification 2	788.36	49.31	15.98
Identification 3	748.97	78.33	9.56
Seven variables SVAR			
Identification 1	990.53	38.81	25.52
Identification 2	979.03	55.54	17.62
Identification 3	986.17	86.08	11.45

Table 7: Computational Time in Seconds

6 Fiscal Policy Shocks

Let us now focus on the second application. In this subsection, we revisit the results reported in Mountford and Uhlig (2009). 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 also consider three combined shocks (which are linear combinations of the unanticipated fiscal policy shocks): deficit-spending shocks, deficit-financed tax cuts 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 optimism shocks. We first replicate the results that Mountford and Uhlig (2009) obtain using the penalty function approach. Then, we repeat their empirical work using our methodology to show how their main results gravely change.[17]

Mountford and Uhlig (2009) conclude that deficit-financed tax cuts shocks work best among the fiscal policy scenarios to improve GDP. In contrast, using our methodology we find no evidence to support such claim. More generally, we find it is very difficult to reach any conclusion about

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

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 immense confidence intervals around the median IRFs and the median fiscal multipliers associated to each scenario. As a consequence, any conclusion derived from Mountford and Uhlig (2009)'s results relies on artificially narrow confidence intervals associated with the penalty function approach. In fact, once we use the ARRW methodology to correct for this shortcoming their results disappear.

Furthermore, our findings show that it is very hard to support any of Mountford and Uhlig (2009)'s claims about the effects of the unanticipated fiscal policy shocks. They focus on two types of these shocks: unanticipated government revenue shocks and unanticipated government spending 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 penalty function approach, we find no support for such claim using our methodology. The median IRFs of GDP, consumption, and non-residential investment to such shocks are negative using the penalty function approach, but positive using our methodology. Furthermore, wide confidence intervals invalidate any conclusion. Regarding unanticipated government spending shocks, the median IRFs from both methodologies are more similar to each other than in the case of unanticipated government revenue shocks, but the confidence intervals are also so wide under our methodology that the IRFs are not statistically different from zero.

Thus, we conclude that, if we use the identification schemes proposed by Mountford and Uhlig (2009) and described in Table 8 below, little can be said about the effects of unanticipated fiscal policy shocks (or the linear combinations of them). The sharp results reported in Mountford and Uhlig (2009) are mostly linked to their factitiously tight confidence intervals and some bias to be reported below. Once the ARRW methodology corrects these two issues it is very hard to support Mountford and Uhlig (2009)'s conclusions.

6.1 Data and Identification Strategy

We use the same dataset as Mountford and Uhlig (2009) in order to shed light on the implications of our methodology. The dataset contains 10 U.S. variables at a quarterly frequency from 1955 to 2000: 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 8. Appendix

8.2 gives details about the estimation procedure and the dataset.

	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	+			
Wage				

Table 8: Mountford and Uhlig (2009)

6.1.1 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 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 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. 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 $\mathbf{S}$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 a $m(j)$ times n matrix of zeros and $m(j) = 4$ if $j = 1$ or 2 and $m(j) = 1$ otherwise, and

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

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

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 2.2. 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 requiring that the unanticipated government revenue (spending) shock, associated with the third (fourth) column of $\mathbf{Q}$, be orthogonal to the first and second column of $\mathbf{Q}$, associated with the business cycle and monetary policy shocks respectively, without restricting other columns of $\mathbf{Q}$. For example, in the case of the unanticipated government spending shock this is accomplished modifying $\mathbf{R}_4(\mathbf{A}_0, \mathbf{A}_+)$ in Theorem 2 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 2 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 $\mathbf{Z}$ matrix. Next, we describe the three linear combinations of the unanticipated fiscal policy shocks that are used to study the three fiscal policy scenarios.

6.1.2 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 four quarters following the initial shock. The deficit-financed tax cuts 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 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 four quarters following the initial shock.[18] Let $(a_{s,t}, b_{s,t})$ for $t = 0, \ldots, 3$ 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 cuts shock (DTC), and the balanced-budget spending shocks (BB) respectively. For example, let us consider the case of a deficit-spending shock. We can solve for such weights by solving the following linear system of equations

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

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 $t = 0, \ldots, 3$ to build the column vector associated to 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 $\mathbf{q}_{BB}$.

[18] 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.

6.2 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 penalty function approach 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 non-residential investment, respectively. Finally, and also because of space considerations, we just concentrate on the responses of GDP, consumption, and non-residential investment. Very similar conclusions can be reached analyzing the responses of the other variables in the SVAR.

Figure 10 plots the IRFs to an unanticipated government revenue shock. Panel (a) replicates the results reported on Figure 4 in Mountford and Uhlig (2009). This panel shows that using the penalty function approach, the median IRFs of GDP, consumption, and non-residential investment are negative. Furthermore, the 68 percent confidence intervals are narrow and do not contain zero. Therefore, using the penalty function approach one can easily conclude that unanticipated government revenue shocks cause a decline in economic activity. In contrast, once we use our methodology, the sign of the median IRFs changes and the 68 percent confidence intervals are much wider − as in the case of optimism shocks the penalty function approach creates artificially narrow confidence intervals. In addition, the penalty function approach chooses a very large median response of total government revenue, and as a byproduct it introduces a downward bias in the response of GDP, consumption, and non-residential investment.

Figure 11 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 non-residential 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 non-residential investment is always negative. Although less than in the case of unanticipated government revenue shock, the median IRFs change when we use the ARRW methodology. The changes are very important for non-residential investment (whose response is positive for 5 periods before becoming also negative). Nevertheless, the confidence intervals are wider under our methodology and they contain zero. Analogously to the case of unanticipated government revenue shocks, the penalty function approach introduces an downward bias (at least for several quarters) in the response of non-residential investment to unanticipated government spending shocks and creates artificially narrow confidence intervals. It is also the case that the penalty function approach 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). As a consequence of choosing a large response of total government revenue and total government spending to minimize the respective loss functions, the penalty function approach introduces a downward bias in the response of some variables to both shocks. But the main difference between both methodologies is the width of the confidence intervals. The results obtained using the ARRW methodology show that the penalty function approach delivers confidence intervals that are artificially narrow. Once the ARRW methodology corrects for such a problem, it is very hard to say anything about the effects of either unanticipated government spending shocks or unanticipated government revenue shocks on GDP, consumption, or non-residential investment. The results under the ARRW methodology are very weak. The strong results reported in Mountford and Uhlig (2009) are due to the bias and artificially narrow confidence intervals associated with the penalty function approach. Once we use the ARRW methodology to amend these two problems the results disappear.

6.2.1 Understanding the Bias and the Narrow Confidence Intervals

As before, we now shed some light on the biases and artificially narrow confidence intervals. Let us begin by the biases related to unanticipated government revenue shock. The penalty function approach 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 penalty function approach is implicitly forcing a negative response of GDP, consumption, and non-residential 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 shock at horizon zero with the IRFs of GDP, consumption, and non-residential investment to the same shock and horizon are -0.13, -0.13, and -0.01 respectively.

In the case of unanticipated government spending shock the penalty function approach also selects a large response of total government spending in order to minimize the loss function. By choosing a large response, the penalty function approach is implicitly forcing a negative response of non-residential 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 shock at horizon zero with the IRF of non-residential investment to the same shock and horizon is -0.26. Additionally, the penalty function approach 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 0.50 and 0.28 respectively.

Let us now focus on the artificially narrow confidence intervals generated by the penalty function approach. 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 penalty function approach selects a single orthogonal matrix instead of drawing from their 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 the reduced form parameters described in Figure 12. To make our point transparent, we use two separate panels to illustrate that given a value of reduced form parameters no uncertainty is considered when the penalty function approach is used.[19] Panel (b) in Figure 12 plots the range (the interval between the maximum and the minimum IRF) of the distribution of IRFs at the OLS point estimate of the reduced form parameters. This range characterizes the support of the distribution of IRFs consistent with the sign and zero restrictions, and it is constructed using the Cholesky decomposition of the OLS point estimate of the reduced form parameters and several draws of the conditional uniform distribution of orthogonal matrix $\mathbf{Q}$. The support of the IRFs reported on Panel (b) is in sharp contrast with the single IRFs reported on Panel (a) obtained using the penalty function approach evaluated at the same value of the reduced form parameters. Figure 13 shows that the same happens in the case of the unanticipated government spending shock.

We can summarize our findings in Figure 14. The first column compares the posterior distributions of IRFs at horizon zero and the median IRFs for total government revenue, GDP, consumption, and non-residential investment to an unanticipated government revenue shock using both the

[19] In the case of optimism shocks we show in the same graph the median IRFs from the penalty function approach, and the median and 68 percent confidence intervals from the ARRW methodology to illustrate the bias present in the penalty function approach. In this section we abstract from replicating these graphs to highlight that the penalty function approach neglects the uncertainty associated with each draw of reduced form parameters.

penalty function approach and the ARRW methodology.[20] The posterior distributions are approximated using a kernel smoothing function. Comparing the penalty function approach and the ARRW methodology we reach the following conclusions. First, the posterior distribution of IRFs of total government revenue obtained using the penalty function approach 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 penalty function approach dramatically underestimates the variance of the posterior distribution of IRFs of total government revenue. These two results were expected since the penalty function approach is maximizing the response of total government revenue in order to minimize the loss function. We also observe very narrow and downward biased posterior distributions of the IRFs for GDP, consumption, and non-residential investment. Columns 2 does the same for the IRFs of total government spending, GDP, consumption, and non-residential investment to a unanticipated government spending shock. Again, the posterior distribution of IRFs of total government spending obtained using the penalty function approach is centered around the right-hand tail of the distribution obtained using the ARRW methodology. We also observe very narrow and downward biased posterior distributions of the IRFs for non-residential investment, and there is an upward bias for the response of GDP.

To conclude we show how the compressed and biased posterior distributions of the IRFs manifest themselves in the imposition of sign restrictions on seemingly unrestricted variables. Table 9 compares the posterior probabilities that the IRFs for GDP, consumption, and non-residential investment to an unanticipated government revenue shock are negative in at least one of the first four horizons. The IRF of the three variables are almost always negative when using Mountford and Uhlig (2009) methodology while it is negative only about 25 percent of the time when using the ARRW methodology. Hence, in this case, the penalty function approach imposes additional negative sign restrictions on variables that are seemingly unrestricted.

Table 10 does the same for unanticipated government spending shock. The penalty function approach imposes an additional negative sign restriction on non-residential investment and an additional positive sign restriction on GDP.

[20]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 persist using horizon 0 only.

	The Penalty Function Approach			The ARRW Methodology		
	Mean	Std dev	$\Pr(\cdot < 0)$	Mean	Std dev	$\Pr(\cdot < 0)$
GDP	-0.1370	0.0757	0.9888	0.1563	0.2012	0.2288
Consumption	-0.0983	0.0852	0.8452	0.0972	0.1538	0.2745
Non-res Investment	-0.3554	0.4795	0.7655	0.8013	1.0829	0.2407

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

	The Penalty Function Approach			The ARRW Methodology		
	Mean	Std dev	$\Pr(\cdot < 0)$	Mean	Std dev	$\Pr(\cdot < 0)$
GDP	0.0885	0.0565	0.0558	0.0748	0.2243	0.3842
Consumption	0.0084	0.0531	0.4125	0.0345	0.1766	0.4253
Non-res Investment	-0.2211	0.3851	0.7147	0.1411	1.1947	0.4565

Table 10: Posterior Probabilities of Negative IRFs in at Least One of the First Four Horizons: Unanticipated Government Spending Shock

6.3 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 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 it has been the case with unanticipated fiscal policy shocks, we just report results for GDP, consumption, and non-residential investment. Very similar conclusions can be reached analyzing the responses of the other variables in the SVAR.

6.3.1 IRFs

We begin by the deficit-spending shocks. Panel (a) on Figure 15 replicates the results on Figure 10 in Mountford and Uhlig (2009). Using the penalty function approach one could conclude that deficit-spending shocks produce a drop in GDP (after a few periods of a small increase), consumption, and non-residential investment (although the drop is only statistically significant for non-residential investment). However, once we use our methodology these results disappear. The median responses are similar, but there is a very wide range of IRFs that are consistent with these shocks. For this reason, it is 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 penalty function approach. This means that, once we combine the unanticipated fiscal policy shocks, the confidence intervals get compounded and become even wider than before. For this reason, although there is no apparent bias when analyzing this scenario, it still hard to reach any conclusion about the effects of deficit-spending shocks. Mountford and Uhlig (2009) conclusions are based on artificially narrow confidence intervals.

Next, we study deficit-financed tax cuts shocks. We find a picture similar to the previous case. Panel (a) on Figure 16 replicates the results on Figure 11 in Mountford and Uhlig (2009). We can see that the median IRFs of GDP, consumption, real, and non-residential investment are positive and the tight 68 percent confidence intervals do not contain zero. Mountford and Uhlig (2009) use this result to claim that deficit-financed tax cuts 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, enormous confidence intervals make the interpretation of the median IRFs very hard. Therefore, we find an upward bias when using the penalty function approach to study the deficit-financed tax cuts shocks. This upward bias and the artificially narrow confidence intervals obtained using the penalty function approach are behind Mountford and Uhlig (2009)'s conclusions. Once the ARRW methodology corrects for these two problems it is impossible to support Mountford and Uhlig (2009) claims. In most cases the confidence intervals reported using the ARRW methodology are one order of magnitude bigger than the confidence intervals reported using the penalty function approach.

Finally, Mountford and Uhlig (2009) study a balanced-budget spending scenario. Panel (a) on Figure 17 replicates the results reported on Figure 12 in Mountford and Uhlig (2009). As it can be seen, the median IRFs of GDP, consumption, and non-residential investment are (almost always) negative and the narrow 68 percent confidence intervals do not contain zero. Again, once we consider our methodology there is no evidence to support these results. Although the median responses are positive for the first few periods, confidence intervals are so wide that it is hard to conclude anything at all. As before, downward bias and too narrow confidence intervals are behind any conclusions implied by Mountford and Uhlig (2009) for this scenario.

Thus, we have shown that using our methodology the analysis of the three fiscal policy scenarios paints a completely different picture than the one reported in Mountford and Uhlig (2009) using the penalty function approach. The sign of some median IRFs change when we use the ARRW methodology. The biases that we find are very hard to interpret because these shocks are linear combinations of shocks that are already biased. However, what it is more important is that our methodology obtains 68 percent confidence intervals that are extremely wide. In most cases the difference between confidence intervals is of one order of magnitude. Because of this, it is impossible to compare the three fiscal policy scenarios. In contrast, Mountford and Uhlig (2009) can compare the three scenarios because the penalty function approach produces artificially 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 cuts 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 cuts shock depending on the value of $s \in \{\text{DS,DTC}\}$, and i denotes the average 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) on Figure 18 replicates the results reported on 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 work best to improve GDP. Unfortunately, once we use the 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 cuts shocks. Thus, the signs of the median cumulative discounted IRFs change. As before, these biases are very hard to interpret because the deficit-spending and the deficit-financed tax cuts shocks are linear combinations of shocks that are already biased. But, more importantly, the correctly computed 68 percent confidence intervals contain zero and are at least five times larger than the ones reported using the penalty function approach. They are just incredibly large and make the economic interpretation of the results impossible. They also reiterate that it is not possible to compare the fiscal scenarios and that any conclusion derived from Mountford and Uhlig (2009)'s results relies on the artificially narrow confidence intervals induced by the penalty function approach.

6.3.2 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 cuts shocks. Specifically, they compute the present value multipliers at horizon τ of 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 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 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 cuts 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 11 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 cuts shocks. The upper block of this panel replicates the results reported on Table 2 in Mountford and Uhlig (2009). The bottom block of this panel reports the results obtained using our methodology. Using the penalty function approach, the median multipliers associated with deficit-financed tax cut shocks are positive for, at least 12 quarters. In contrast, these median multipliers

	1 qrt	4 qrts	8 qrts	12 qrts	20 qrts	Max
(a) Present Value Multipliers						
The Penalty Function Approach						
DTC	0.28	0.52	1.61	4.48	-4.52	4.48
	[0.11 , 0.50]	[0.43 , 0.63]	[1.11 , 2.27]	[0.97 , 15.15]	[-9.57 , 4.54]	(qrt 12)
DS	0.65	0.44	0.05	-0.28	-1.25	0.65
	[0.29 , 0.98]	[0.32 , 0.57]	[-0.30 , 0.38]	[-1.04 , 0.41]	[-4.43 , 0.42]	(qrt 1)
The ARRW Methodology						
DTC	-0.81	-0.77	-1.25	-1.76	-1.86	-0.77
	[-3.31 , 0.87]	[-2.91 , 3.49]	[-3.93 , 1.62]	[-4.40 , 0.84]	[-3.40 , -0.17]	(qrt 4)
DS	1.14	0.94	1.22	1.81	3.40	3.59
	[-3.32 , 5.87]	[-6.05 , 6.52]	[-6.17 , 7.36]	[-5.17 , 9.31]	[-1.97 , 10.02]	(qrt 26)
(b) Impact Multipliers: DTC						
The Penalty Function Approach						
GDP	0.28	0.93	2.02	3.07	2.15	3.21
	[0.11 , 0.50]	[0.74 , 1.13]	[1.30 , 2.82]	[1.81 , 4.45]	[0.37 , 3.98]	(qrt 14)
Total Gov Revenue	-1.00	-1.00	0.04	0.91	0.80	
	[-1.00 , -1.00]	[-1.00 , -1.00]	[-0.34 , 0.39]	[0.29 , 1.57]	[-0.16 , 1.76]	
Total Gov Spending	0.00	0.00	0.29	0.43	0.37	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[0.17 , 0.39]	[0.25 , 0.60]	[0.10 , 0.71]	
The ARRW Methodology						
GDP	-0.81	-0.47	0.80	0.61	0.16	0.81
	[-3.31 , 0.87]	[-3.65 , 3.49]	[-4.07 , 1.62]	[-8.06 , 0.84]	[-9.30 , 18.38]	(qrt 11)
Total Gov Revenue	-1.00	-1.00	0.00	0.10	-0.00	
	[-1.00 , -1.00]	[-1.00 , -1.00]	[-2.03 , 10.58]	[-4.46 , 7.63]	[-5.70 , 8.19]	
Total Gov Spending	0.00	0.00	0.11	0.10	0.02	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[-0.54 , 1.84]	[-0.67 , 3.71]	[-1.21 , 3.89]	
(c) Impact Multipliers: DS						
The Penalty Function Approach						
GDP	0.65	0.25	-0.73	-0.95	-1.46	0.65
	[0.29 , 0.98]	[-0.14 , 0.69]	[-1.66 , 0.18]	[-2.31 , 0.43]	[-3.70 , 0.26]	(qrt 1)
Total Gov Spending	1.00	1.00	0.81	0.35	-0.26	
	[1.00 , 1.00]	[1.00 , 1.00]	[0.63 , 0.97]	[0.09 , 0.60]	[-0.69 , 0.17]	
Total Gov Revenue	0.00	0.00	-0.36	-0.76	-1.53	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[-0.93 , 0.21]	[-1.64 , 0.16]	[-2.95 , -0.35]	
The ARRW Methodology						
GDP	1.14	0.74	-0.37	0.11	-0.38	1.36
	[-3.32 , 5.87]	[-15.04 , 9.28]	[-23.25 , 11.67]	[-30.71 , 19.00]	[-29.38 , 22.11]	(qrt 3)
Total Gov Spending	1.00	1.00	0.65	0.25	-0.08	
	[1.00 , 1.00]	[1.00 , 1.00]	[-4.75 , 2.44]	[-7.72 , 2.93]	[-8.05 , 3.70]	
Total Gov Revenue	0.00	0.00	-0.24	-0.35	-0.83	
	[-0.00 , 0.00]	[-0.00 , 0.00]	[-15.14 , 8.17]	[-18.81 , 12.89]	[-22.60 , 15.07]	

Table 11: Fiscal 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 penalty function approach, they are positive during 20 quarters using the ARRW methodology. Also, the median multiplier associated with the deficit-spending shocks are larger. At their maximum value, the deficit-spending shocks multipliers are five times larger than the ones reported using the penalty function approach. Hence, the biases reported when describing the IRFs for the unanticipated fiscal policy shocks also affect the median present value multipliers of the fiscal policies. In any case, these new biases are hard to interpret because the multipliers being analyzed correspond to shocks that are linear combinations of shocks that are already biased. Most importantly, the ARRW methodology reports confidence intervals that are huge relative to the ones obtained using the penalty function approach. They are so wide that it is almost impossible to reach any conclusion about the sign of the reported multipliers. The width of the confidence intervals highlights, once more, the effects of compounding confidence intervals once these have been rightly computed.

Panel (b) presents the impact multipliers associated with deficit-financed tax cut shocks. The upper block of this panel replicates the results reported on Table 3 in Mountford and Uhlig (2009). The bottom block of the panel reports the results obtained using the ARRW methodology. Clearly, there are differences between the penalty function approach and our 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 sign, the median multipliers associated with our methodology are much smaller than the median multipliers implied by the penalty function approach. In any case, as before, the confidence intervals computed using the ARRW methodology are so large that it 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 on 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 quarter) for the median multiplier. However, the magnitudes are different. The absolute value of the GDP median multipliers resulting from the ARRW methodology are approximately twice as large as the ones resulting from the penalty function approach. But it is also the case that the confidence intervals computed using the ARRW methodology are so wide that it 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. We have shown that once we use our methodology it is very hard to support such a claim. Some median multipliers change sign, nevertheless the most important result is that the confidence intervals are so wide that it is very hard to reach any conclusion from a statistical point of view.

6.4 Computational Time

Our methodology is also faster than the penalty function approach in the case of unanticipated fiscal policy shocks. In this case, the penalty function approach is approximately 1.6 times slower than our methodology. The results are closer, but the ARRW methodology is still faster.

	The Penalty Function Approach	The ARRW Methodology	Ratio
Fiscal Policy Shocks	7162.5	4598.9	1.6

Table 12: Computational Time in Seconds

7 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 et al. (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 et al. (2011) and Mountford and Uhlig (2009), respectively, using the penalty function approach. These authors have provided very definitive answers. Unfortunately, we have shown that these sharp conclusions are due to shortcomings in the penalty function approach. In particular, we have shown that the penalty function approach (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 penalty function approach does not correctly draw from the posterior distribution of structural parameters conditional on the sign and zero restrictions. This problem is common to all the existing methods. Our algorithm is also faster than the current ones.

8 Appendices

8.1 Appendix A. Estimation and Inference: Optimism Shocks

Following Beaudry et al. (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 parameters 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 dataset created by Beaudry et al. (2011). This dataset 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 Poors 500 composite index divided by CPI of all items from the Bureau of Labor Statistics (BLS). Consumption is real consumption spending on non-durable goods and services from Bureau of Economics 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 BLS. Hours worked is hours of all persons in the non-farm business sector from BLS. Investment is real gross private domestic investment from BEA. Output is real output in the non-farm business sector from 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 BLS.

8.2 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 parameters draws from the Normal-Wishart posterior $(\mathbf{B}, \mathbf{\Sigma})$ and from the conditional uniform distribution of $\mathbf{Q}$. We used the same dataset as Mountford and Uhlig (2009). This dataset 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. The details of the dataset can be found in Mountford and Uhlig (2009).

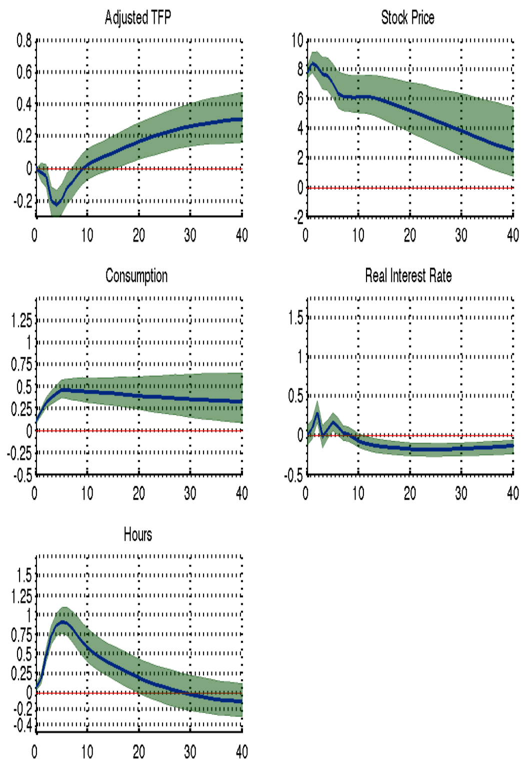
8.3 Appendix C. Tables and Figures

	Identification I						Identification II						Identification III					
The ARRW Methodology																		
	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	0.05	0.09	0.00	0.01	0.02	0.04	0.06	0.11	0.00	0.01	0.02	0.05	0.08	0.16
	[0.00 , 0.00]	[0.00 , 0.04]	[0.01 , 0.05]	[0.01 , 0.09]	[0.02 , 0.14]	[0.02 , 0.23]	[0.00 , 0.00]	[0.00 , 0.04]	[0.01 , 0.06]	[0.01 , 0.10]	[0.02 , 0.15]	[0.03 , 0.26]	[0.00 , 0.00]	[0.00 , 0.03]	[0.01 , 0.05]	[0.02 , 0.10]	[0.03 , 0.18]	[0.05 , 0.31]
Stock Price	0.15	0.15	0.15	0.15	0.15	0.15	0.20	0.23	0.24	0.26	0.26	0.25	0.19	0.25	0.27	0.29	0.29	0.29
	[0.01 , 0.51]	[0.02 , 0.51]	[0.02 , 0.51]	[0.03 , 0.50]	[0.03 , 0.48]	[0.03 , 0.47]	[0.02 , 0.58]	[0.04 , 0.61]	[0.04 , 0.60]	[0.05 , 0.60]	[0.05 , 0.58]	[0.06 , 0.56]	[0.02 , 0.60]	[0.06 , 0.64]	[0.06 , 0.64]	[0.07 , 0.64]	[0.07 , 0.62]	[0.08 , 0.60]
Consumption	0.17	0.17	0.17	0.16	0.15	0.15	0.20	0.29	0.30	0.29	0.28	0.27	0.22	0.35	0.38	0.41	0.41	0.38
	[0.02 , 0.50]	[0.02 , 0.52]	[0.02 , 0.52]	[0.02 , 0.51]	[0.02 , 0.50]	[0.02 , 0.49]	[0.02 , 0.57]	[0.07 , 0.63]	[0.07 , 0.62]	[0.06 , 0.62]	[0.06 , 0.61]	[0.05 , 0.59]	[0.03 , 0.62]	[0.09 , 0.69]	[0.12 , 0.70]	[0.14 , 0.71]	[0.13 , 0.70]	[0.12 , 0.67]
Real Interest Rate	0.16	0.17	0.17	0.18	0.18	0.19	0.16	0.17	0.17	0.17	0.18	0.20	0.18	0.18	0.19	0.19	0.20	0.22
	[0.01 , 0.57]	[0.04 , 0.51]	[0.05 , 0.49]	[0.06 , 0.48]	[0.07 , 0.46]	[0.08 , 0.43]	[0.02 , 0.56]	[0.04 , 0.51]	[0.04 , 0.49]	[0.05 , 0.48]	[0.06 , 0.46]	[0.08 , 0.44]	[0.02 , 0.58]	[0.04 , 0.54]	[0.04 , 0.54]	[0.05 , 0.52]	[0.07 , 0.49]	[0.08 , 0.47]
Hours Worked	0.16	0.17	0.16	0.17	0.16	0.17	0.15	0.25	0.28	0.27	0.26	0.24	0.16	0.26	0.30	0.31	0.31	0.30
	[0.01 , 0.51]	[0.04 , 0.48]	[0.04 , 0.49]	[0.03 , 0.50]	[0.04 , 0.48]	[0.04 , 0.46]	[0.02 , 0.50]	[0.06 , 0.61]	[0.06 , 0.61]	[0.06 , 0.60]	[0.06 , 0.58]	[0.07 , 0.56]	[0.02 , 0.56]	[0.06 , 0.61]	[0.07 , 0.64]	[0.07 , 0.64]	[0.07 , 0.62]	[0.08 , 0.59]
The Penalty Function Approach																		
	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	0.04	0.04	0.05	0.08	0.18	0.00	0.02	0.03	0.05	0.08	0.23	0.00	0.01	0.02	0.05	0.11	0.28
	[0.00 , 0.00]	[0.01 , 0.07]	[0.02 , 0.09]	[0.03 , 0.09]	[0.04 , 0.15]	[0.09 , 0.31]	[0.00 , 0.00]	[0.01 , 0.05]	[0.01 , 0.07]	[0.02 , 0.09]	[0.04 , 0.17]	[0.10 , 0.40]	[0.00 , 0.00]	[0.00 , 0.03]	[0.01 , 0.05]	[0.02 , 0.11]	[0.04 , 0.21]	[0.12 , 0.44]
Stock Price	1.00	0.93	0.91	0.86	0.80	0.73	0.64	0.71	0.71	0.73	0.74	0.72	0.42	0.51	0.53	0.57	0.58	0.58
	[0.98 , 1.00]	[0.89 , 0.96]	[0.85 , 0.95]	[0.77 , 0.92]	[0.68 , 0.89]	[0.56 , 0.86]	[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 , 0.59]	[0.43 , 0.62]	[0.44 , 0.68]	[0.44 , 0.70]	[0.42 , 0.71]
Consumption	0.07	0.27	0.30	0.30	0.29	0.26	0.59	0.81	0.83	0.79	0.76	0.71	0.50	0.71	0.77	0.80	0.79	0.76
	[0.04 , 0.11]	[0.19 , 0.34]	[0.21 , 0.40]	[0.19 , 0.42]	[0.17 , 0.42]	[0.14 , 0.42]	[0.54 , 0.63]	[0.76 , 0.86]	[0.75 , 0.89]	[0.70 , 0.87]	[0.64 , 0.86]	[0.55 , 0.84]	[0.45 , 0.55]	[0.63 , 0.77]	[0.67 , 0.84]	[0.68 , 0.88]	[0.65 , 0.88]	[0.58 , 0.87]
Real Interest Rate	0.00	0.03	0.04	0.06	0.09	0.14	0.01	0.03	0.04	0.06	0.08	0.13	0.37	0.34	0.36	0.35	0.34	0.36
	[0.00 , 0.01]	[0.02 , 0.06]	[0.02 , 0.07]	[0.03 , 0.10]	[0.05 , 0.15]	[0.08 , 0.23]	[0.00 , 0.03]	[0.01 , 0.06]	[0.02 , 0.09]	[0.03 , 0.10]	[0.04 , 0.14]	[0.07 , 0.22]	[0.32 , 0.42]	[0.28 , 0.41]	[0.29 , 0.43]	[0.27 , 0.41]	[0.27 , 0.42]	[0.28 , 0.44]
Hours Worked	0.02	0.29	0.37	0.34	0.32	0.31	0.05	0.56	0.70	0.69	0.66	0.63	0.04	0.39	0.52	0.54	0.52	0.50
	[0.00 , 0.03]	[0.22 , 0.37]	[0.27 , 0.46]	[0.23 , 0.46]	[0.21 , 0.44]	[0.20 , 0.43]	[0.03 , 0.09]	[0.48 , 0.63]	[0.62 , 0.77]	[0.59 , 0.77]	[0.54 , 0.76]	[0.49 , 0.74]	[0.02 , 0.06]	[0.32 , 0.46]	[0.42 , 0.61]	[0.42 , 0.65]	[0.38 , 0.65]	[0.35 , 0.63]

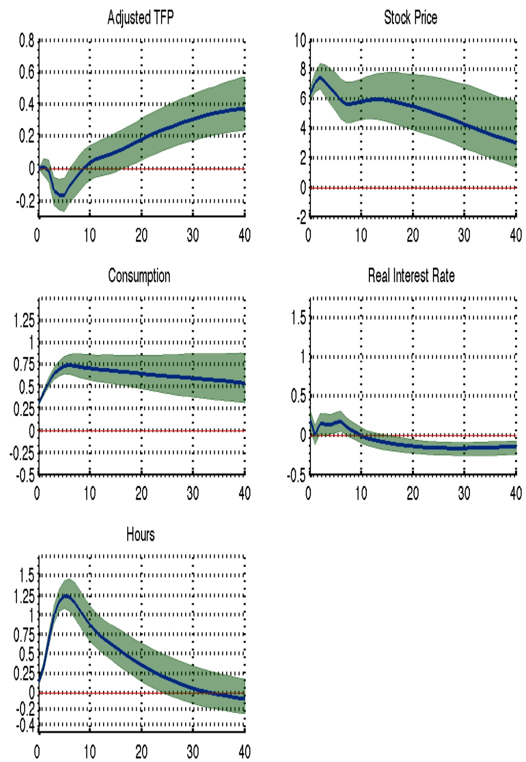
Table 13: Share of FEV Attributable to Optimism Shocks. Five Variable SVAR

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	0.05	0.08	0.13
	[0.00 , 0.00]	[0.01 , 0.06]	[0.01 , 0.08]	[0.02 , 0.13]	[0.02 , 0.16]	[0.03 , 0.21]	[0.00 , 0.00]	[0.01 , 0.06]	[0.01 , 0.08]	[0.02 , 0.13]	[0.03 , 0.18]	[0.03 , 0.24]	[0.00 , 0.00]	[0.01 , 0.06]	[0.01 , 0.08]	[0.02 , 0.14]	[0.03 , 0.20]	[0.04 , 0.28]
Stock Price	0.09	0.09	0.10	0.10	0.11	0.11	0.11	0.15	0.16	0.18	0.18	0.17	0.10	0.17	0.19	0.22	0.22	0.22
	[0.01 , 0.37]	[0.02 , 0.34]	[0.02 , 0.33]	[0.02 , 0.32]	[0.02 , 0.32]	[0.03 , 0.31]	[0.01 , 0.40]	[0.03 , 0.42]	[0.04 , 0.42]	[0.04 , 0.41]	[0.04 , 0.40]	[0.04 , 0.39]	[0.01 , 0.39]	[0.05 , 0.45]	[0.06 , 0.45]	[0.07 , 0.46]	[0.07 , 0.46]	[0.06 , 0.43]
Consumption	0.08	0.09	0.09	0.09	0.09	0.10	0.11	0.16	0.17	0.17	0.17	0.15	0.12	0.19	0.22	0.24	0.23	0.22
	[0.01 , 0.31]	[0.01 , 0.33]	[0.01 , 0.34]	[0.01 , 0.35]	[0.01 , 0.35]	[0.02 , 0.35]	[0.01 , 0.35]	[0.04 , 0.41]	[0.04 , 0.43]	[0.03 , 0.43]	[0.03 , 0.42]	[0.03 , 0.40]	[0.01 , 0.39]	[0.05 , 0.47]	[0.06 , 0.49]	[0.06 , 0.49]	[0.06 , 0.49]	[0.04 , 0.48]
Real Interest Rate	0.10	0.11	0.11	0.12	0.13	0.13	0.10	0.10	0.10	0.11	0.11	0.13	0.11	0.11	0.12	0.12	0.13	0.14
	[0.01 , 0.36]	[0.03 , 0.33]	[0.04 , 0.32]	[0.04 , 0.31]	[0.05 , 0.30]	[0.05 , 0.28]	[0.01 , 0.35]	[0.03 , 0.32]	[0.03 , 0.32]	[0.04 , 0.31]	[0.04 , 0.30]	[0.05 , 0.27]	[0.01 , 0.40]	[0.03 , 0.35]	[0.04 , 0.35]	[0.04 , 0.34]	[0.05 , 0.33]	[0.06 , 0.30]
Hours Worked	0.09	0.09	0.10	0.10	0.10	0.10	0.10	0.14	0.16	0.17	0.16	0.14	0.10	0.14	0.17	0.18	0.17	0.15
	[0.01 , 0.36]	[0.02 , 0.32]	[0.02 , 0.33]	[0.02 , 0.33]	[0.03 , 0.33]	[0.03 , 0.32]	[0.01 , 0.36]	[0.03 , 0.42]	[0.04 , 0.43]	[0.04 , 0.42]	[0.04 , 0.39]	[0.04 , 0.36]	[0.01 , 0.35]	[0.03 , 0.41]	[0.04 , 0.42]	[0.03 , 0.41]	[0.03 , 0.39]	[0.04 , 0.36]
Investment	0.07	0.10	0.10	0.11	0.11	0.12	0.07	0.13	0.16	0.16	0.16	0.16	0.08	0.11	0.15	0.17	0.17	0.17
	[0.01 , 0.25]	[0.03 , 0.28]	[0.03 , 0.28]	[0.04 , 0.28]	[0.04 , 0.29]	[0.04 , 0.29]	[0.01 , 0.25]	[0.04 , 0.34]	[0.05 , 0.38]	[0.05 , 0.37]	[0.06 , 0.37]	[0.06 , 0.35]	[0.01 , 0.25]	[0.04 , 0.31]	[0.05 , 0.37]	[0.05 , 0.37]	[0.06 , 0.37]	[0.06 , 0.36]
Output	0.03	0.08	0.09	0.10	0.10	0.10	0.04	0.14	0.17	0.18	0.18	0.17	0.04	0.14	0.19	0.21	0.23	0.23
	[0.00 , 0.12]	[0.02 , 0.27]	[0.02 , 0.28]	[0.02 , 0.30]	[0.02 , 0.31]	[0.02 , 0.32]	[0.00 , 0.13]	[0.03 , 0.36]	[0.05 , 0.39]	[0.05 , 0.39]	[0.05 , 0.39]	[0.04 , 0.39]	[0.00 , 0.13]	[0.03 , 0.38]	[0.05 , 0.42]	[0.06 , 0.45]	[0.07 , 0.44]	[0.06 , 0.44]
Labor Productivity	0.02	0.06	0.07	0.08	0.09	0.09	0.02	0.07	0.07	0.09	0.10	0.13	0.02	0.07	0.08	0.11	0.14	0.19
	[0.00 , 0.06]	[0.01 , 0.20]	[0.02 , 0.22]	[0.02 , 0.26]	[0.02 , 0.28]	[0.02 , 0.29]	[0.00 , 0.06]	[0.01 , 0.21]	[0.02 , 0.23]	[0.02 , 0.27]	[0.03 , 0.30]	[0.03 , 0.34]	[0.00 , 0.06]	[0.01 , 0.23]	[0.02 , 0.25]	[0.03 , 0.31]	[0.04 , 0.37]	[0.06 , 0.43]
The Penalty Function Approach																		
	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	0.16	0.00	0.02	0.03	0.05	0.09	0.22	0.00	0.01	0.02	0.06	0.13	0.30
	[0.00 , 0.00]	[0.01 , 0.03]	[0.01 , 0.05]	[0.03 , 0.09]	[0.04 , 0.17]	[0.07 , 0.28]	[0.00 , 0.00]	[0.01 , 0.03]	[0.02 , 0.05]	[0.02 , 0.09]	[0.04 , 0.18]	[0.09 , 0.38]	[0.00 , 0.00]	[0.01 , 0.03]	[0.01 , 0.05]	[0.03 , 0.12]	[0.05 , 0.24]	[0.14 , 0.46]
Stock Price	1.00	0.87	0.81	0.73	0.65	0.52	0.63	0.72	0.72	0.72	0.69	0.62	0.40	0.57	0.60	0.63	0.63	0.57
	[0.99 , 1.00]	[0.81 , 0.91]	[0.75 , 0.88]	[0.63 , 0.82]	[0.52 , 0.77]	[0.36 , 0.70]	[0.58 , 0.67]	[0.66 , 0.78]	[0.63 , 0.79]	[0.61 , 0.80]	[0.57 , 0.78]	[0.48 , 0.74]	[0.35 , 0.44]	[0.49 , 0.64]	[0.51 , 0.68]	[0.53 , 0.73]	[0.50 , 0.73]	[0.43 , 0.69]
Consumption	0.04	0.19	0.23	0.22	0.20	0.14	0.53	0.72	0.74	0.69	0.65	0.55	0.45	0.63	0.69	0.71	0.69	0.61
	[0.02 , 0.07]	[0.12 , 0.27]	[0.14 , 0.32]	[0.12 , 0.32]	[0.10 , 0.30]	[0.06 , 0.28]	[0.49 , 0.58]	[0.65 , 0.78]	[0.66 , 0.81]	[0.61 , 0.78]	[0.53 , 0.75]	[0.40 , 0.70]	[0.40 , 0.50]	[0.55 , 0.70]	[0.59 , 0.77]	[0.59 , 0.81]	[0.54 , 0.80]	[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	0.29	0.30	0.28	0.27	0.26
	[0.00 , 0.01]	[0.02 , 0.05]	[0.02 , 0.06]	[0.03 , 0.09]	[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 , 0.36]	[0.24 , 0.36]	[0.22 , 0.35]	[0.21 , 0.34]	[0.20 , 0.34]
Hours Worked	0.01	0.23	0.29	0.24	0.20	0.16	0.04	0.47	0.61	0.57	0.50	0.40	0.02	0.30	0.43	0.41	0.38	0.31
	[0.00 , 0.03]	[0.17 , 0.31]	[0.21 , 0.38]	[0.16 , 0.35]	[0.12 , 0.31]	[0.09 , 0.28]	[0.02 , 0.07]	[0.40 , 0.54]	[0.53 , 0.69]	[0.47 , 0.66]	[0.40 , 0.62]	[0.27 , 0.54]	[0.01 , 0.04]	[0.22 , 0.37]	[0.33 , 0.53]	[0.29 , 0.55]	[0.25 , 0.52]	[0.18 , 0.45]
Investment	0.00	0.27	0.34	0.33	0.31	0.26	0.01	0.38	0.50	0.50	0.49	0.46	0.01	0.22	0.35	0.37	0.40	0.40
	[0.00 , 0.01]	[0.21 , 0.34]	[0.26 , 0.43]	[0.24 , 0.42]	[0.21 , 0.40]	[0.16 , 0.38]	[0.00 , 0.02]	[0.31 , 0.45]	[0.42 , 0.59]	[0.41 , 0.59]	[0.39 , 0.59]	[0.35 , 0.58]	[0.00 , 0.02]	[0.15 , 0.28]	[0.24 , 0.44]	[0.26 , 0.48]	[0.27 , 0.51]	[0.26 , 0.54]
Output	0.01	0.27	0.34	0.34	0.31	0.24	0.03	0.56	0.67	0.68	0.66	0.60	0.02	0.41	0.54	0.61	0.63	0.60
	[0.01 , 0.02]	[0.20 , 0.35]	[0.25 , 0.43]	[0.24 , 0.44]	[0.20 , 0.42]	[0.13 , 0.38]	[0.02 , 0.05]	[0.49 , 0.63]	[0.60 , 0.74]	[0.61 , 0.76]	[0.57 , 0.75]	[0.46 , 0.72]	[0.01 , 0.03]	[0.33 , 0.48]	[0.45 , 0.63]	[0.49 , 0.71]	[0.49 , 0.73]	[0.45 , 0.73]
Labor Productivity	0.00	0.05	0.05	0.10	0.15	0.24	0.00	0.10	0.09	0.12	0.20	0.37	0.00	0.10	0.11	0.18	0.30	0.49
	[0.00 , 0.00]	[0.02 , 0.10]	[0.02 , 0.10]	[0.04 , 0.18]	[0.07 , 0.28]	[0.11 , 0.38]	[0.00 , 0.01]	[0.05 , 0.17]	[0.04 , 0.16]	[0.05 , 0.22]	[0.08 , 0.34]	[0.17 , 0.56]	[0.00 , 0.00]	[0.06 , 0.17]	[0.05 , 0.18]	[0.09 , 0.30]	[0.15 , 0.45]	[0.28 , 0.65]

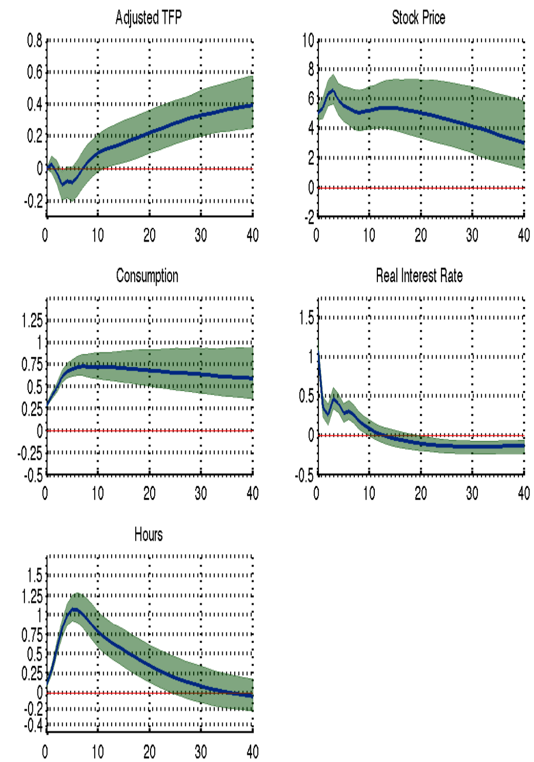
Table 14: Share of FEV Attributable to Optimism Shocks. Seven Variable SVAR



(a) Identification 1

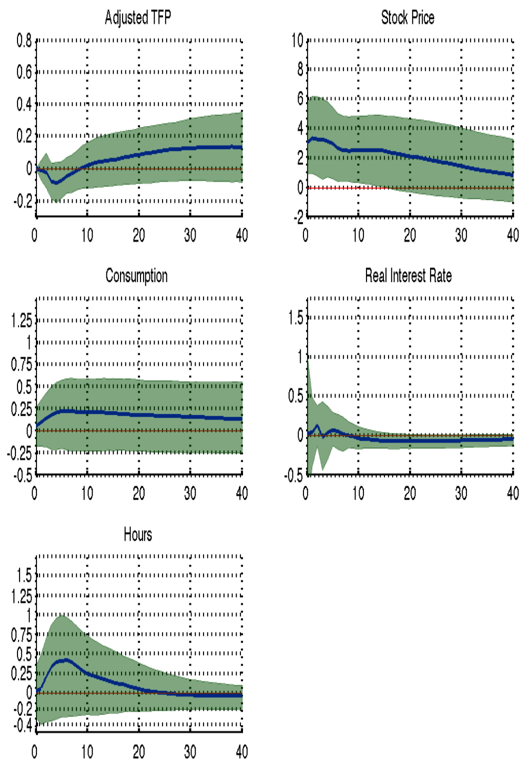


(b) Identification 2

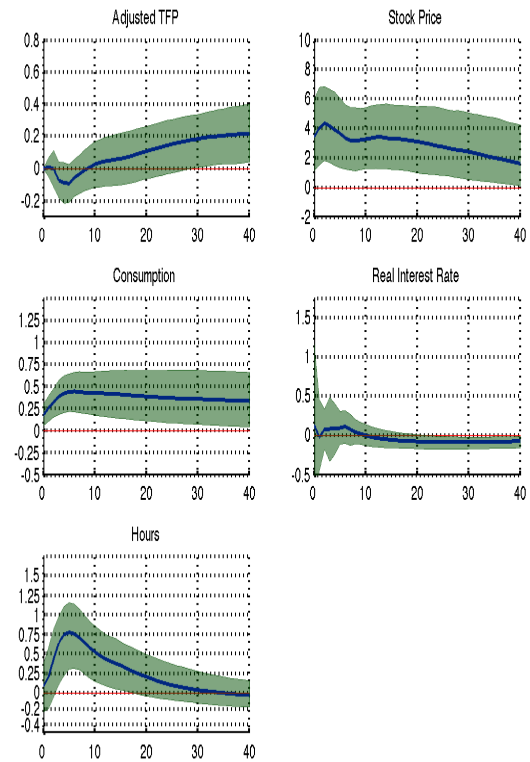


(c) Identification 3

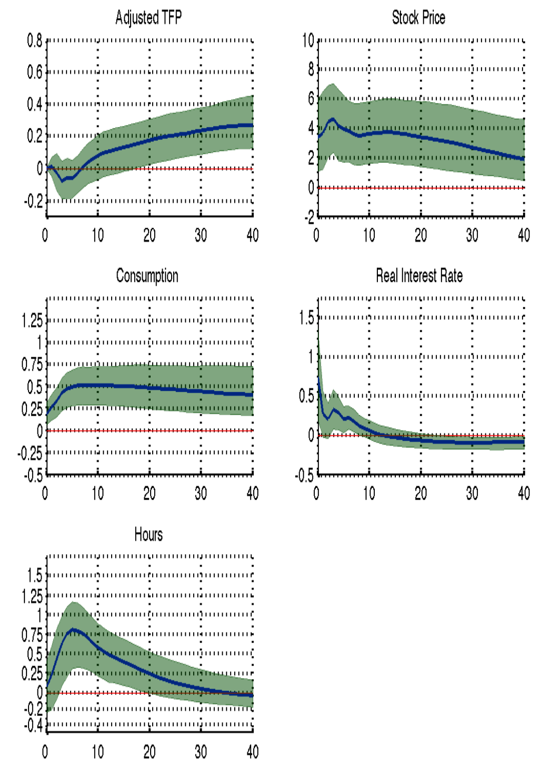
Figure 1: IRFs to an Optimism Shock Using the Penalty Function Approach. Five Variables SVAR



(a) Identification 1

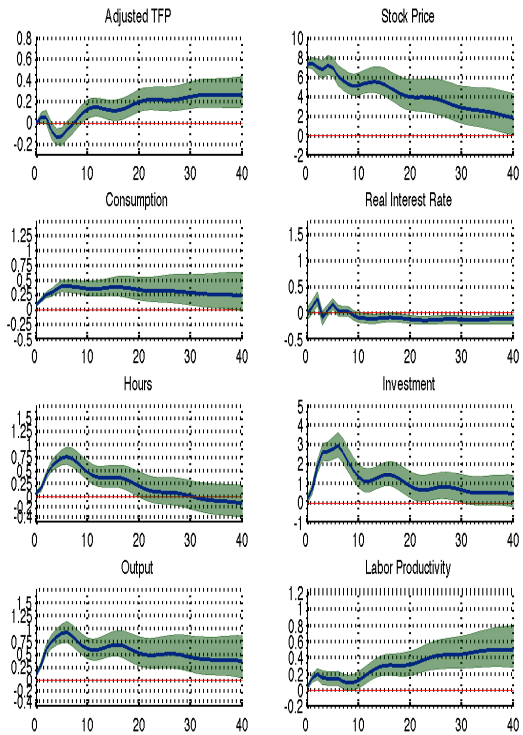


(b) Identification 2

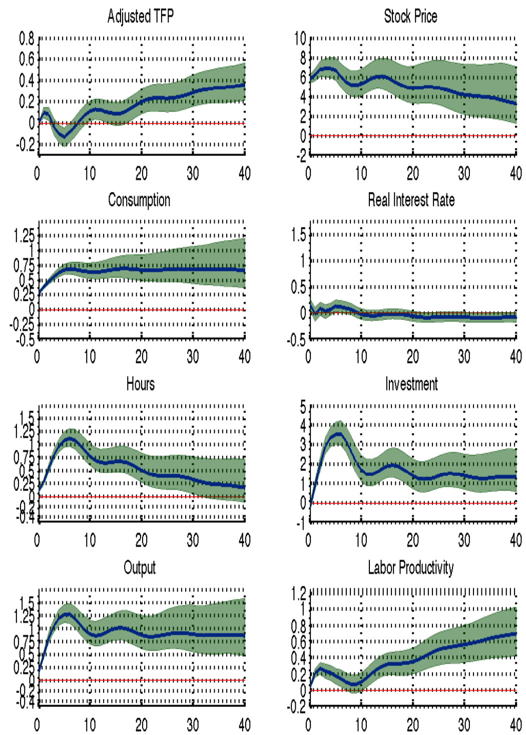


(c) Identification 3

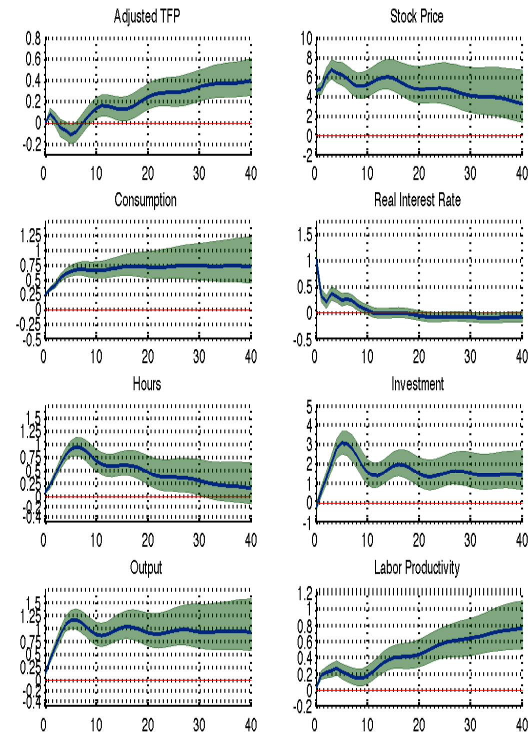
Figure 2: IRFs to an Optimism Shock Using the ARRW Methodology. Five Variables SVAR



(a) Identification 1

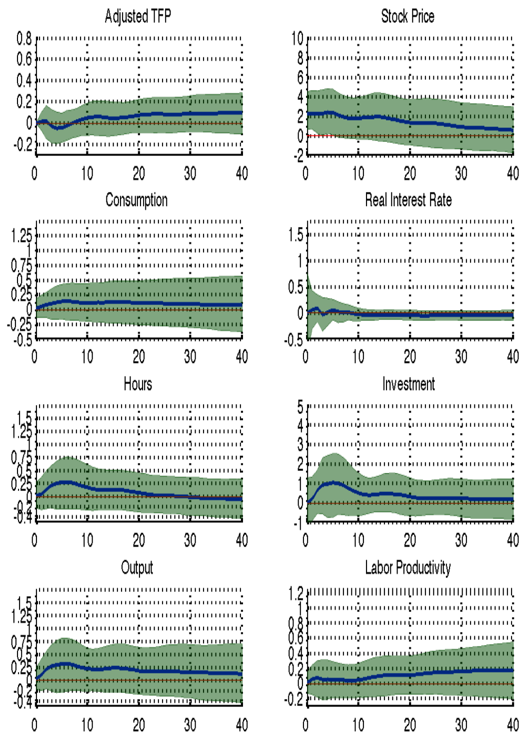


(b) Identification 2

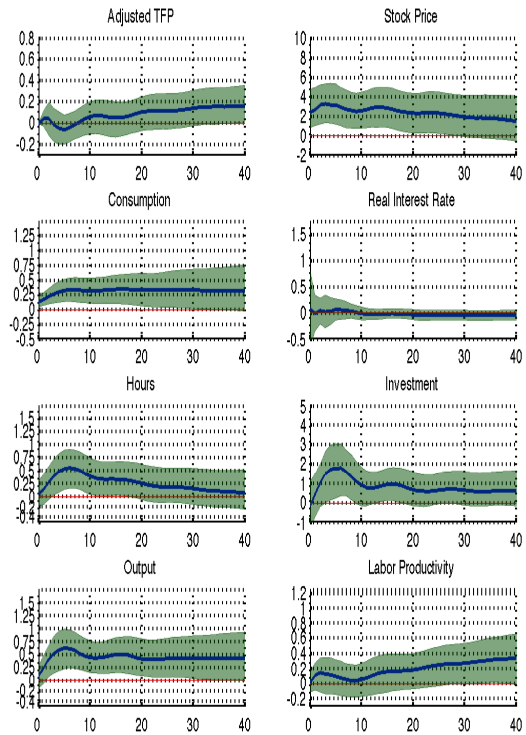


(c) Identification 3

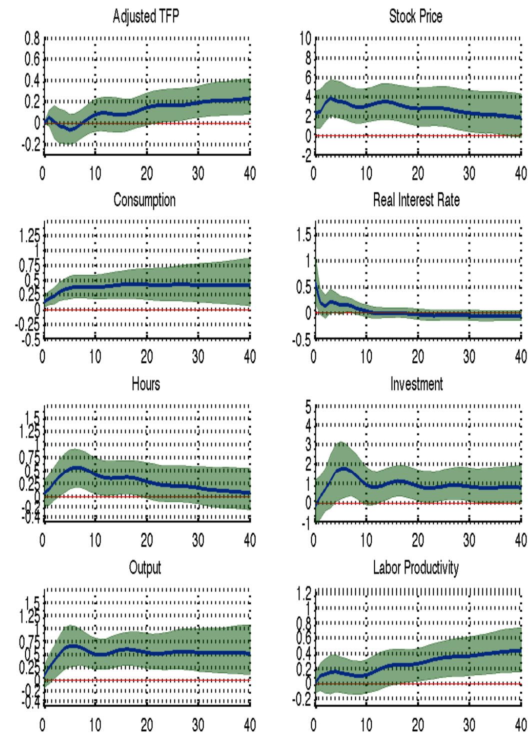
Figure 3: IRFs to an Optimism Shock Using the Penalty Function Approach. Seven Variables SVAR



(a) Identification 1

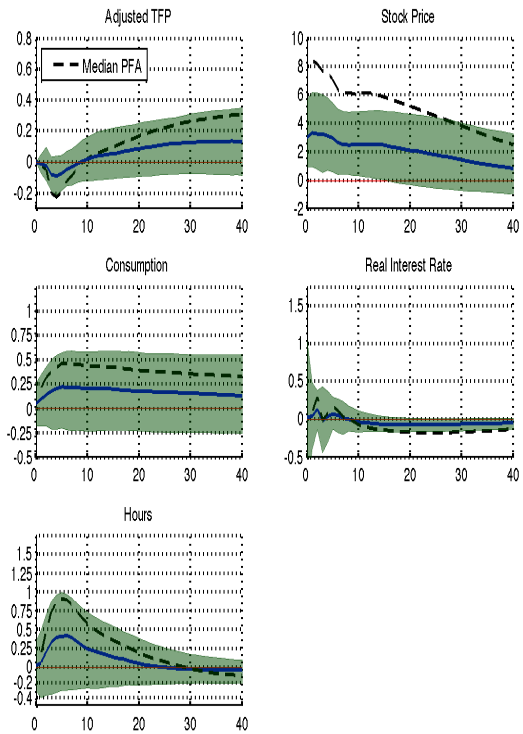


(b) Identification 2

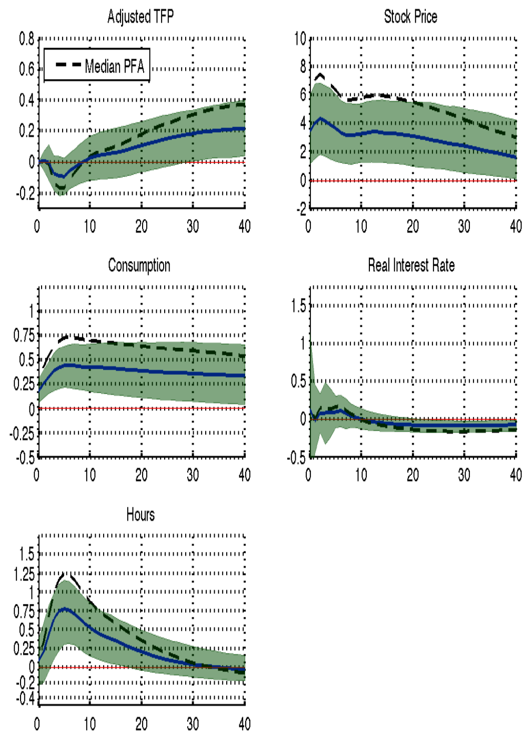


(c) Identification 3

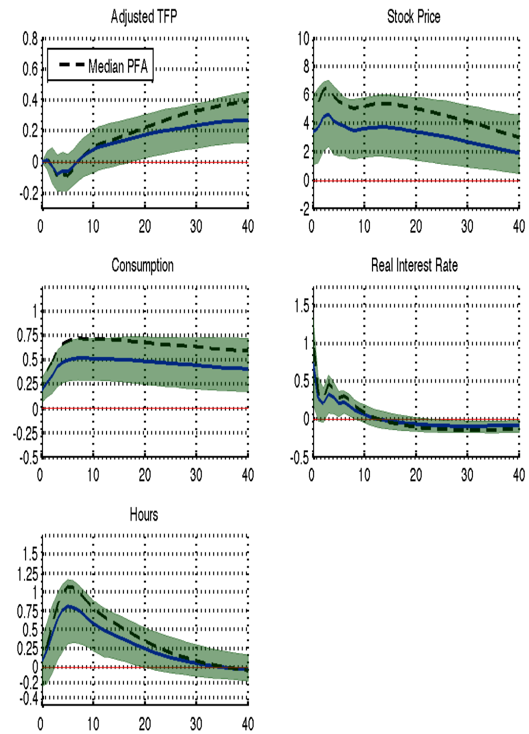
Figure 4: IRFs to an Optimism Shock Using the ARRW Methodology. Seven Variables SVAR



(a) Identification 1



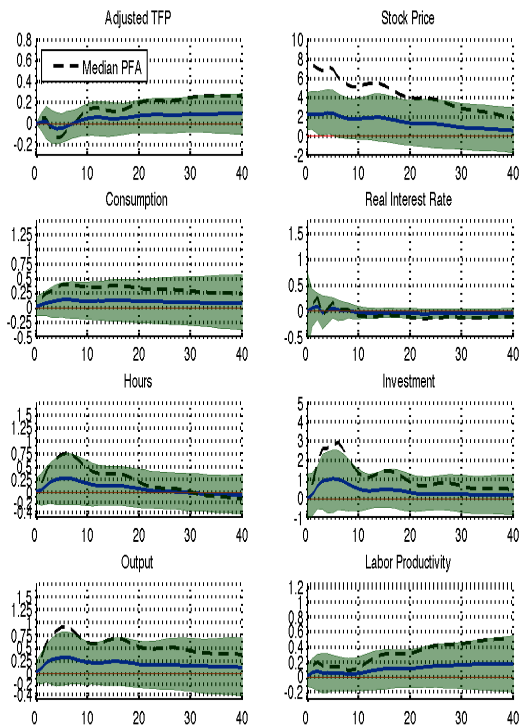
(b) Identification 2



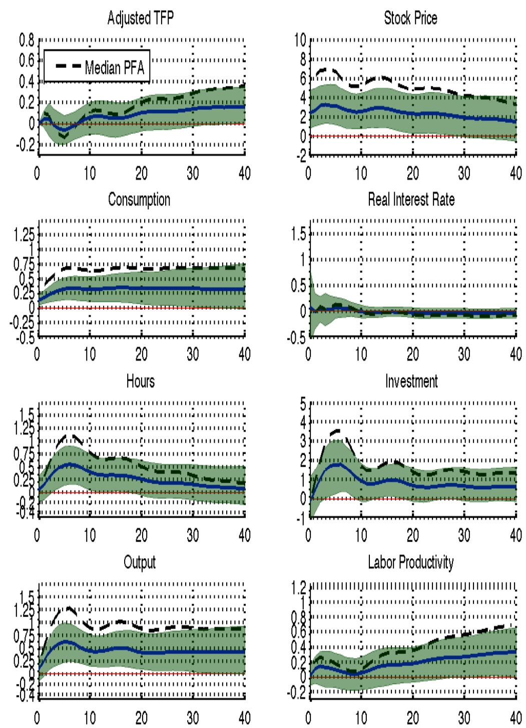
(c) Identification 3

Figure 5: Comparison of IRFs to an Optimism Shock. Five Variables SVAR

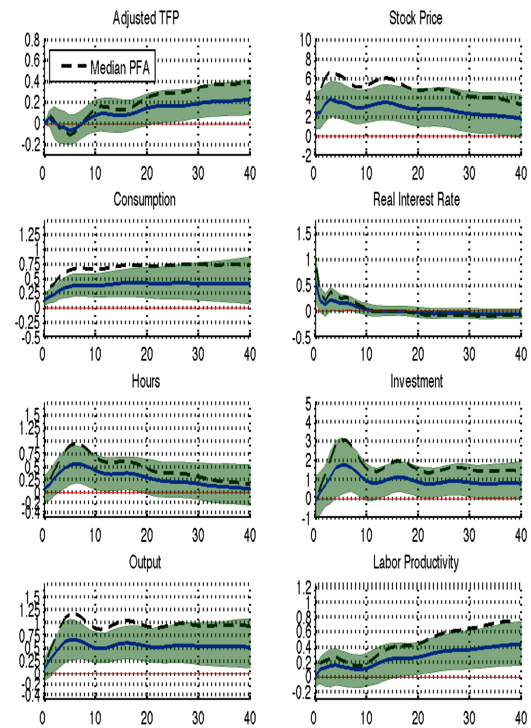
Note: Median PFA refers to the median IRF obtained using the penalty function approach.



(a) Identification 1



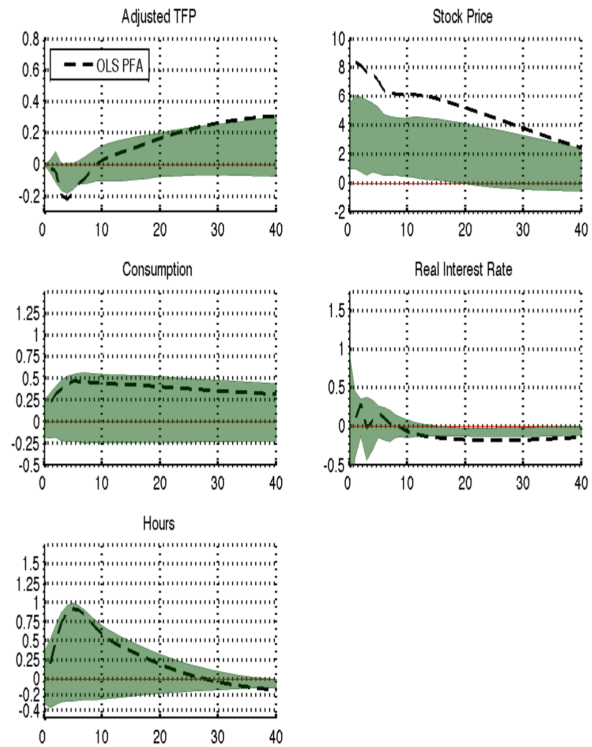
(b) Identification 2



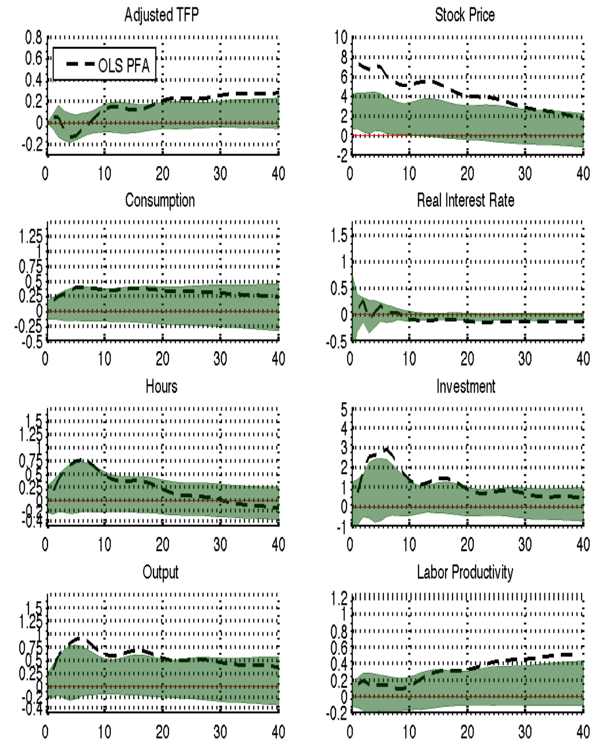
(c) Identification 3

Figure 6: Comparison of IRFs to an Optimism Shock. Seven Variables SVAR

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



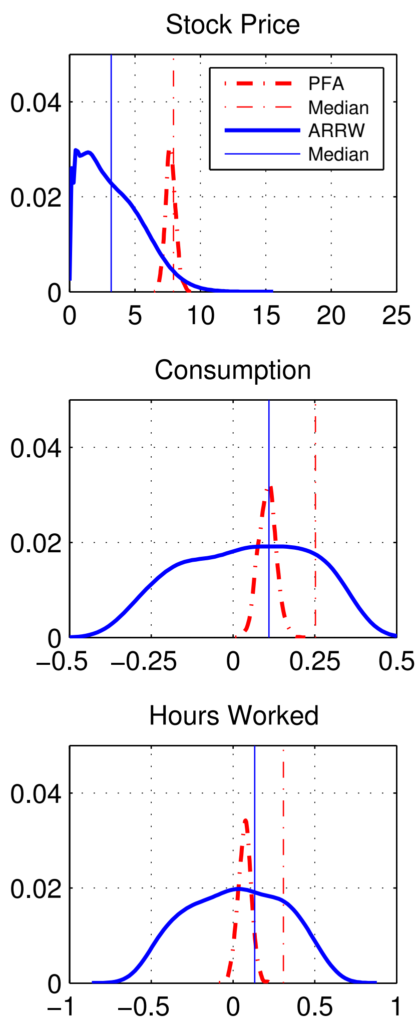
(a) Five Variables SVAR



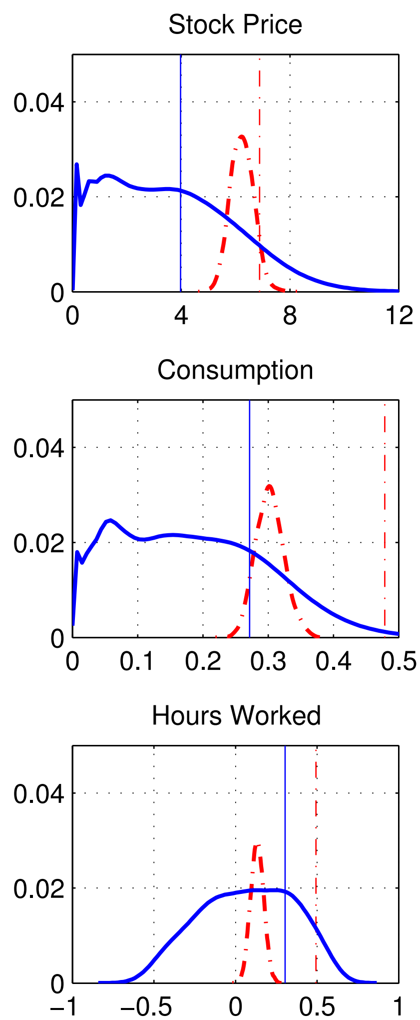
(b) Seven Variables SVAR

Figure 7: Distribution of IRFs with the ARRW methodology vs Single IRFs with the Penalty Function Approach. Identification 1

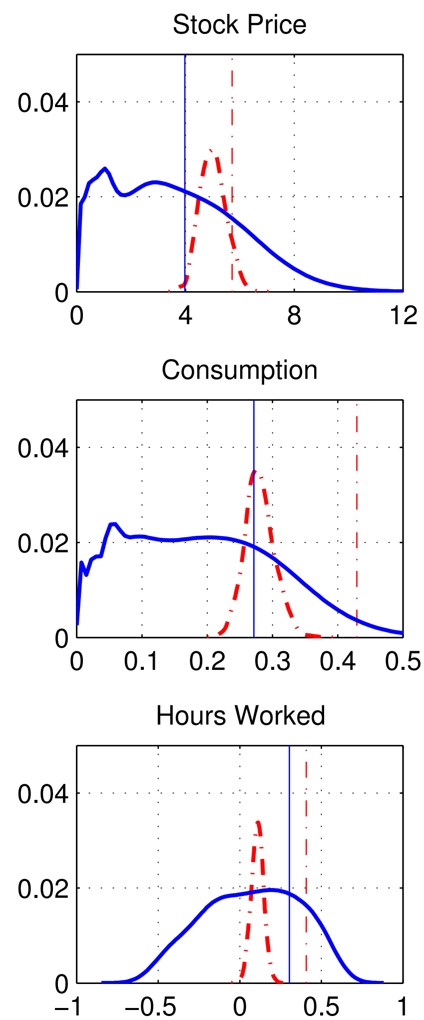
Note: OLS PFA refers to the IRF obtained using the penalty function approach and the OLS reduced form estimates.



(a) Identification 1

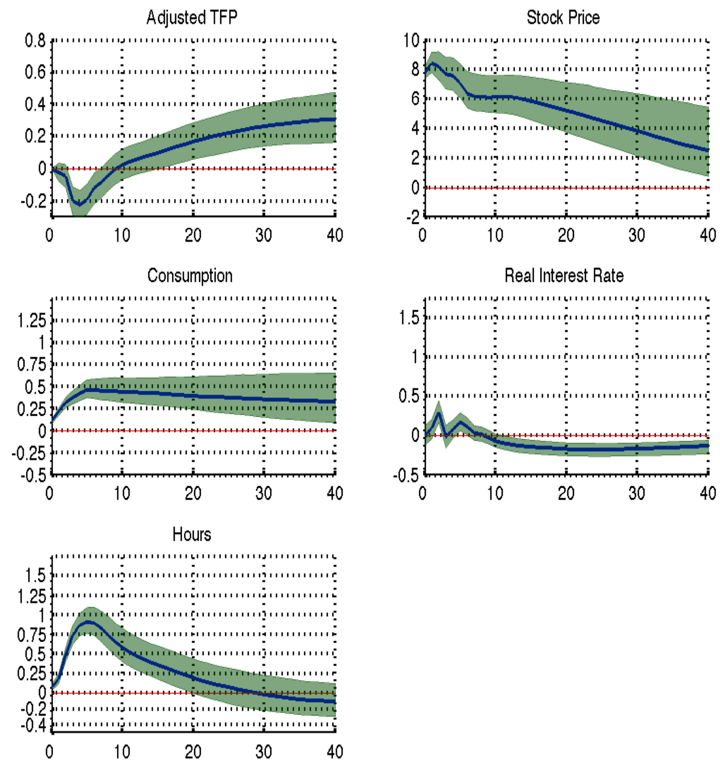


(b) Identification 2

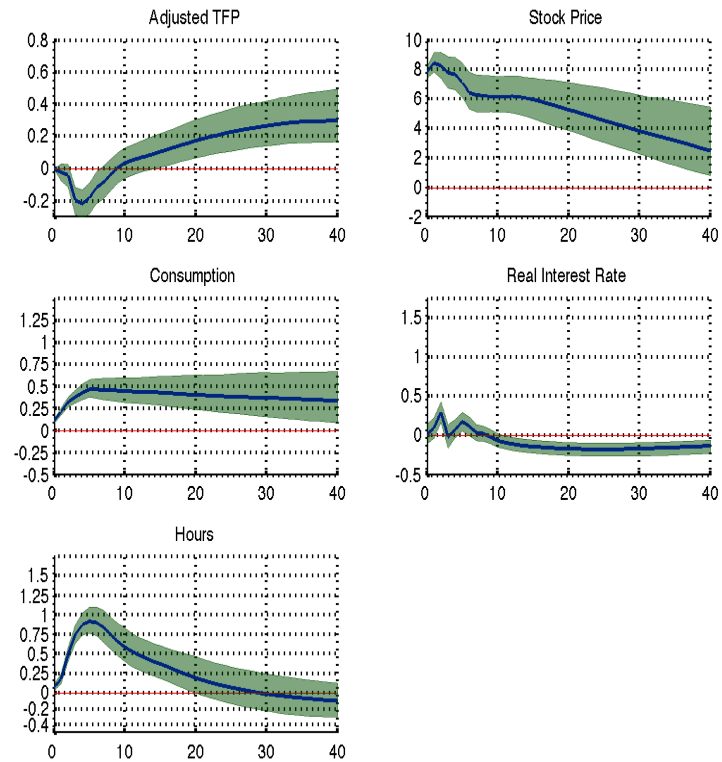


(c) Identification 3

Figure 8: Density of IRFs at Horizon Zero and Median of IRFs from Horizons Zero to Three

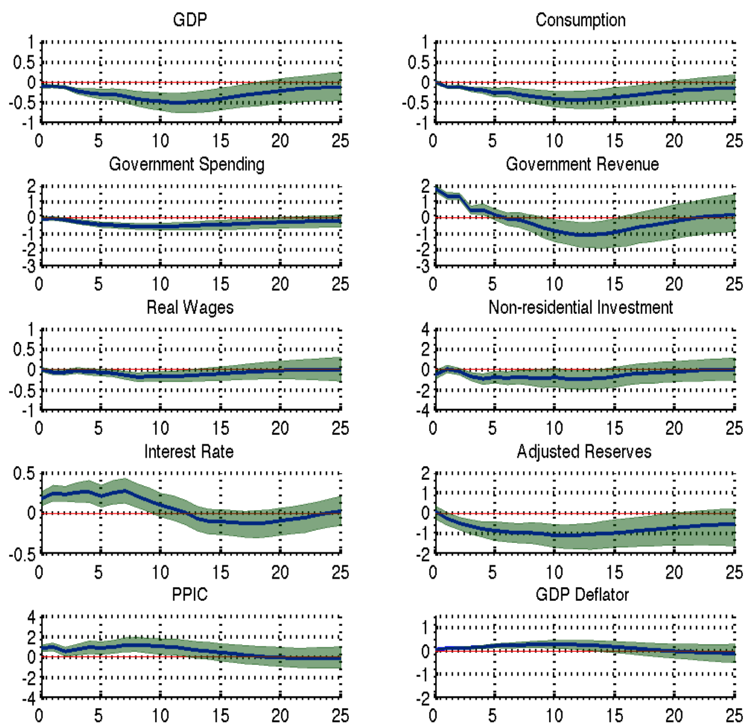


(a) The Penalty Function Approach

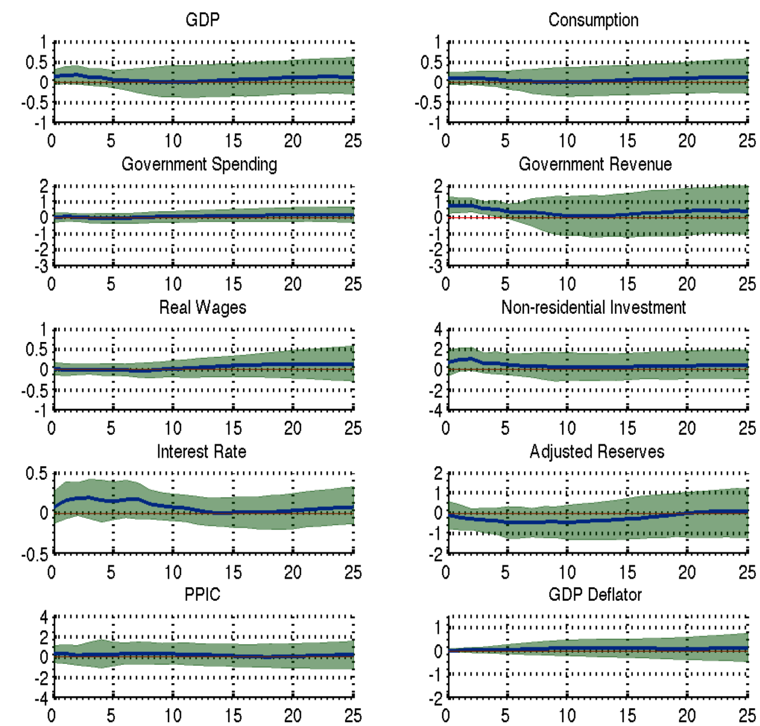


(b) The ARRW Methodology with Additional Restrictions

Figure 9: Replicating the Penalty Function Approach using the ARRW Methodology. Five Variables SVAR

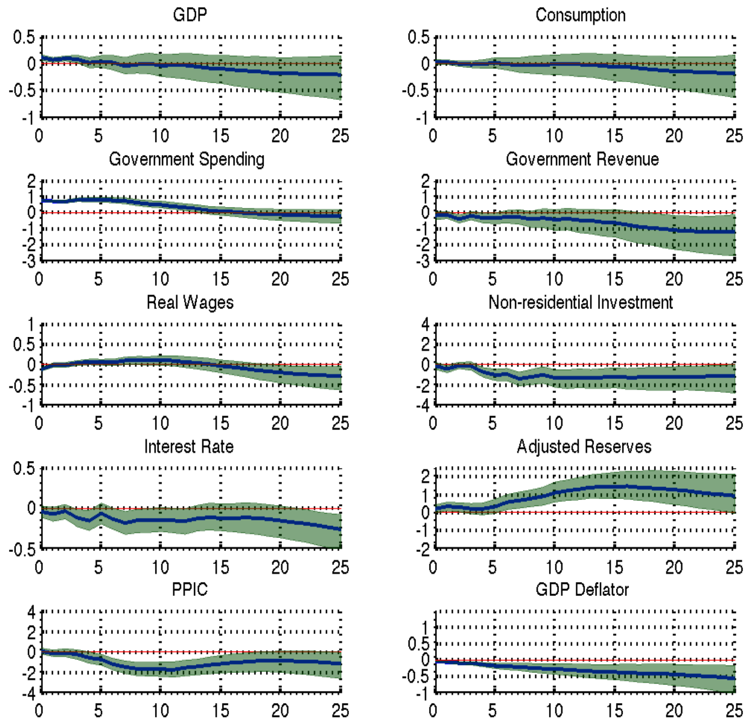


(a) The Penalty Function Approach

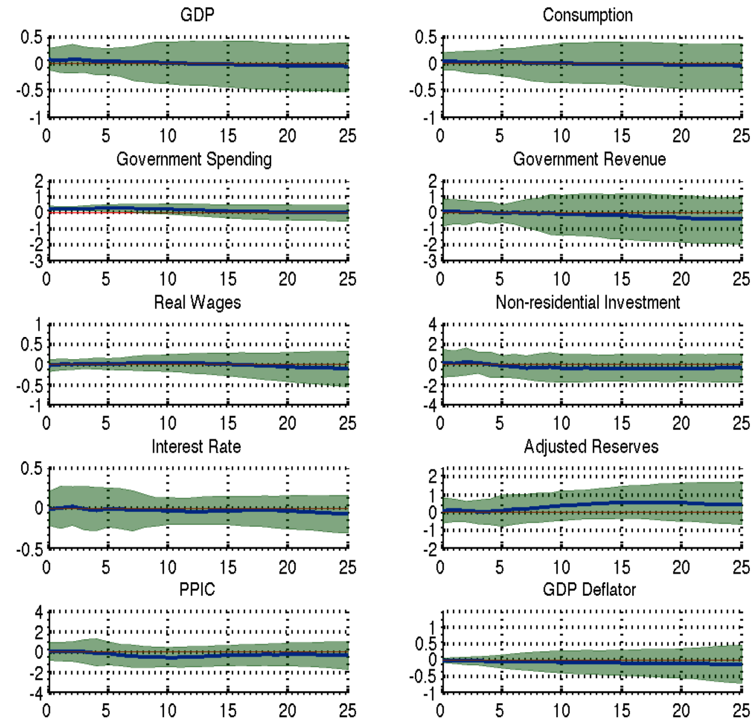


(b) The ARRW Methodology

Figure 10: IRFs to an Unanticipated Government Revenue Shock

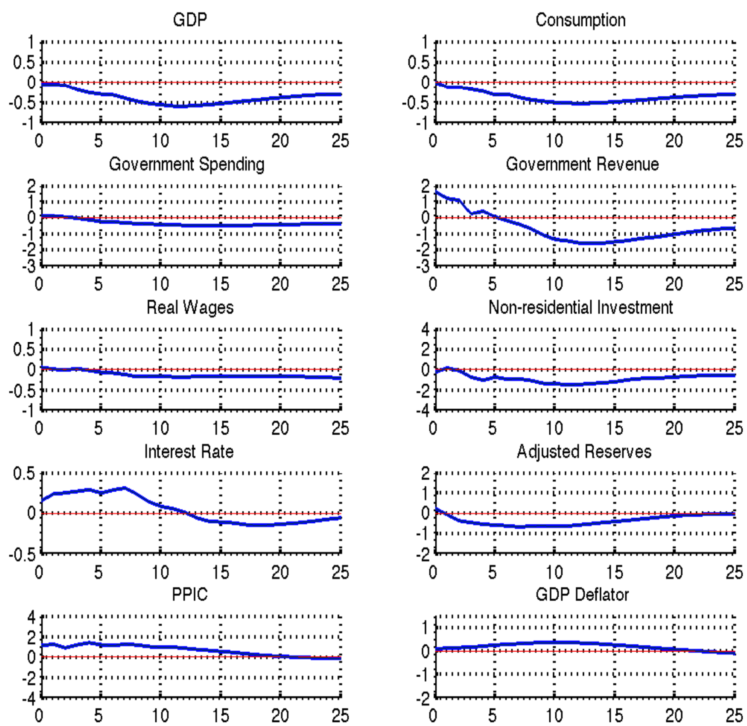


(a) The Penalty Function Approach

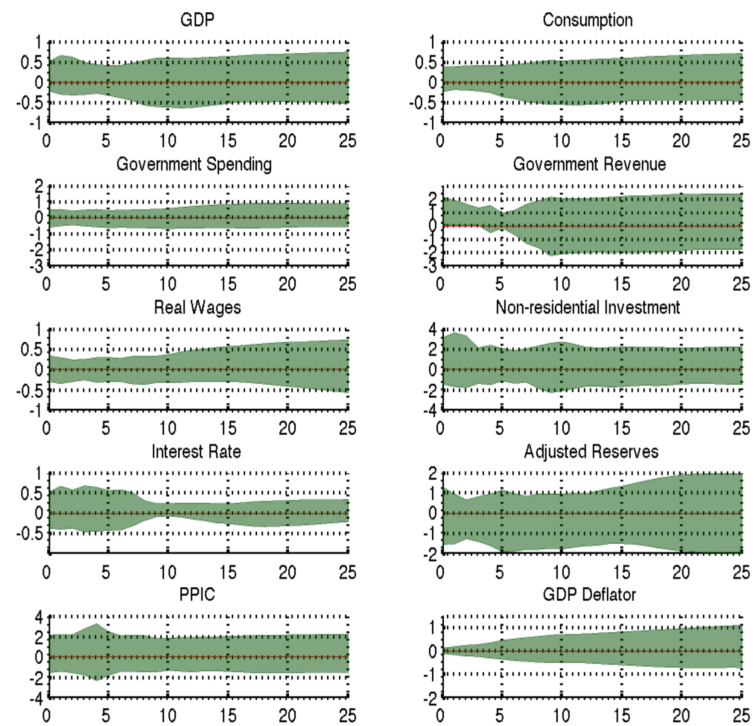


(b) The ARRW Methodolgy

Figure 11: IRFs to an Unanticipated Government Spending Shock

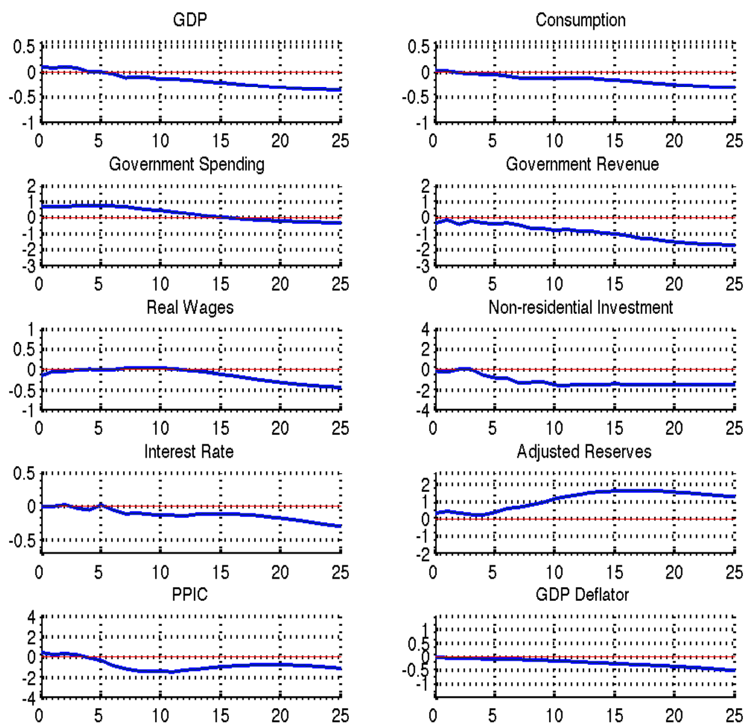


(a) The Penalty Function Approach

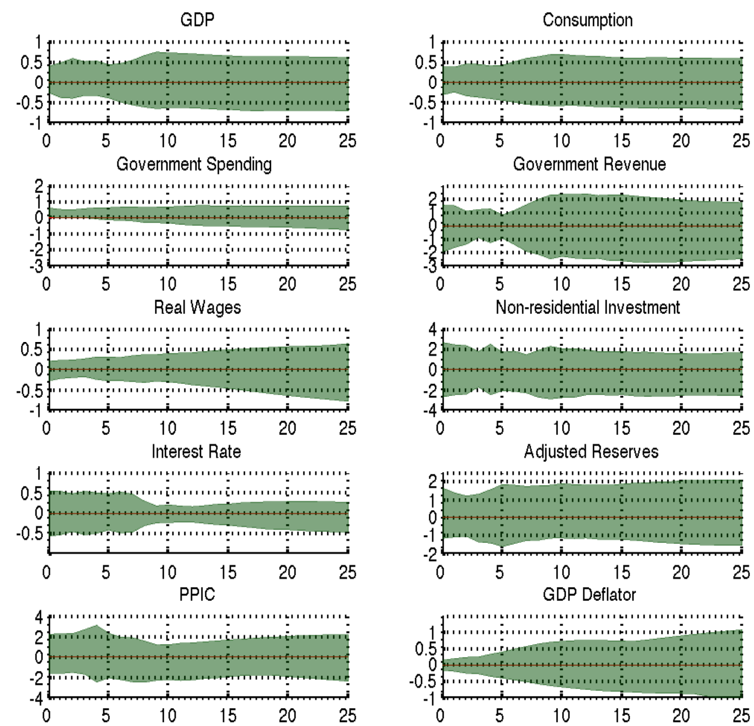


(b) The ARRW Methodolgy

Figure 12: IRFs to an Unanticipated Government Revenue Shock

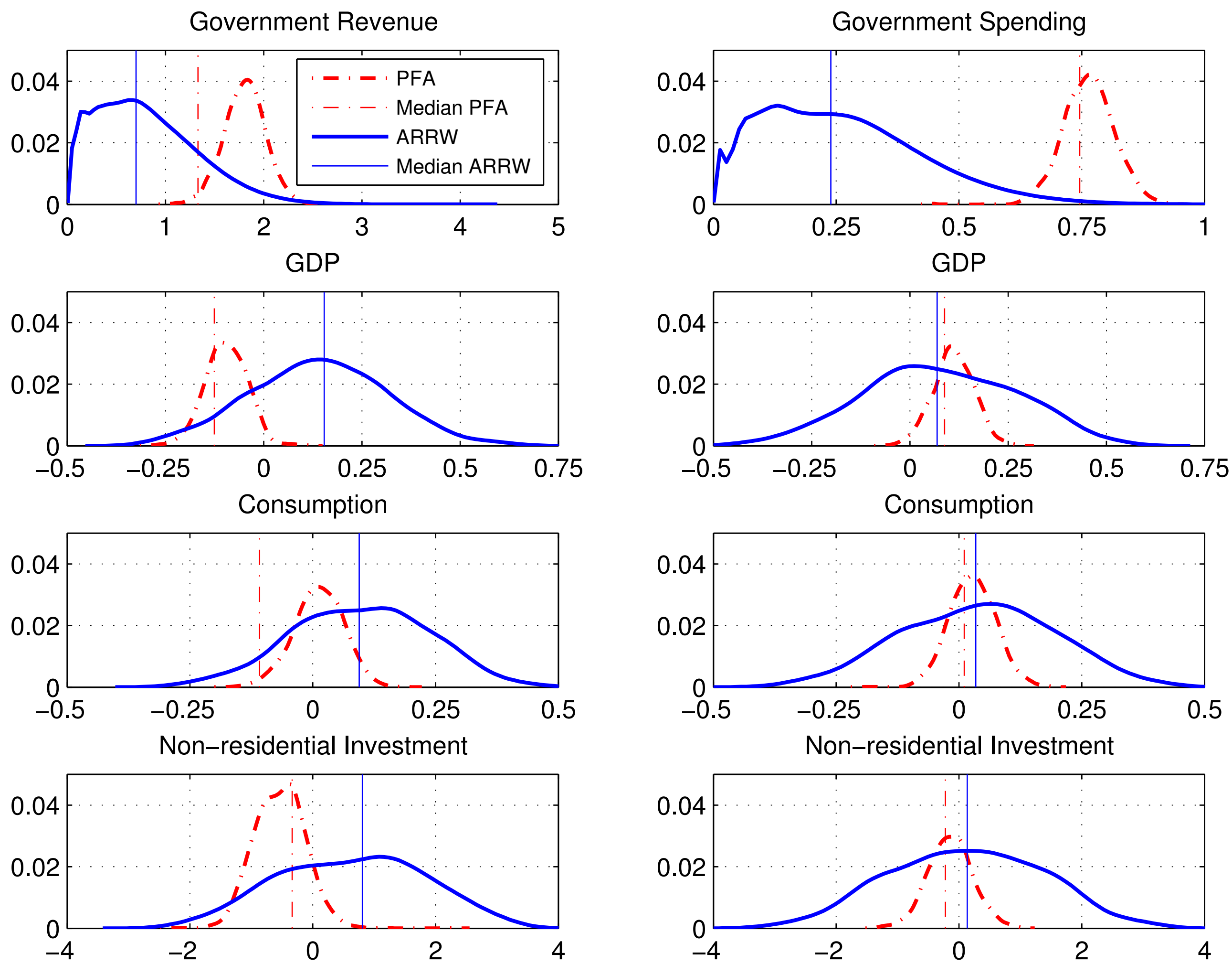


(a) The Penalty Function Approach



(b) The ARRW Methodology

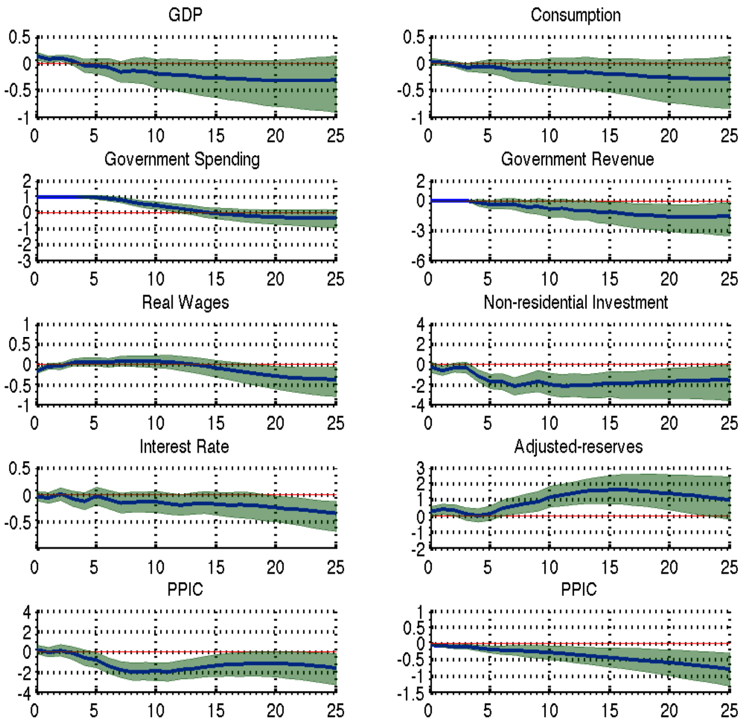
Figure 13: IRFs to an Unanticipated Government Spending Shock



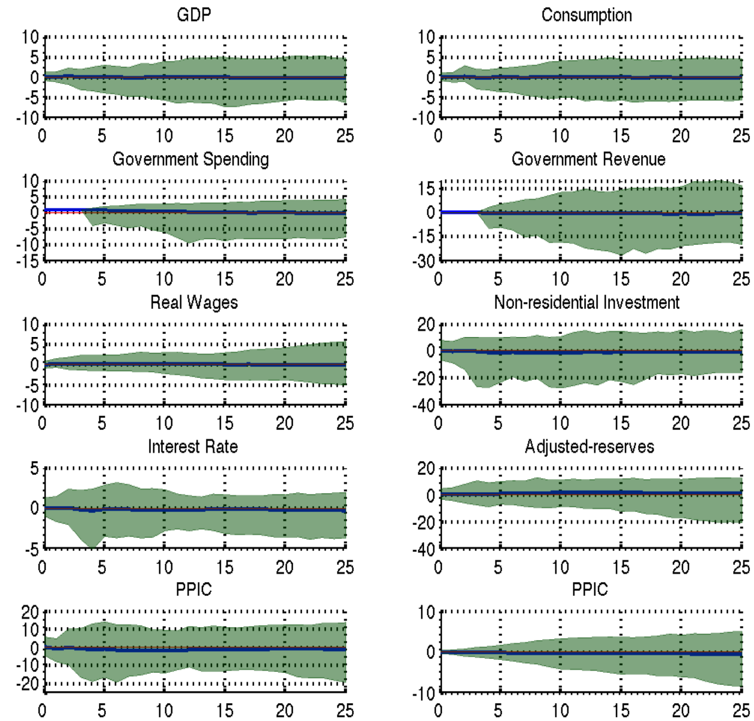
(a) Gov Revenue Shock

(b) Gov Spending Shock

Figure 14: Density of IRFs at Horizon Zero and Median of IRFs from Horizons Zero to Three

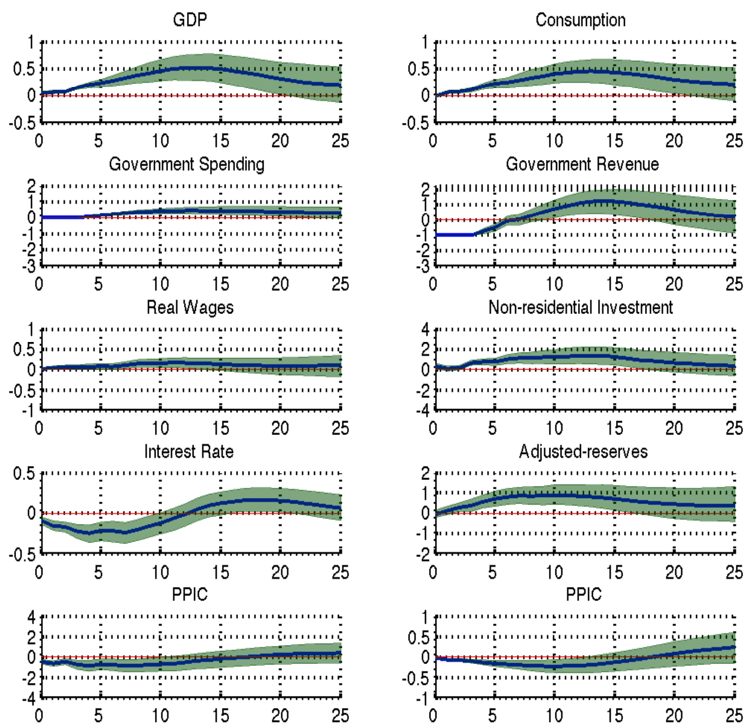


(a) The Penalty Function Approach

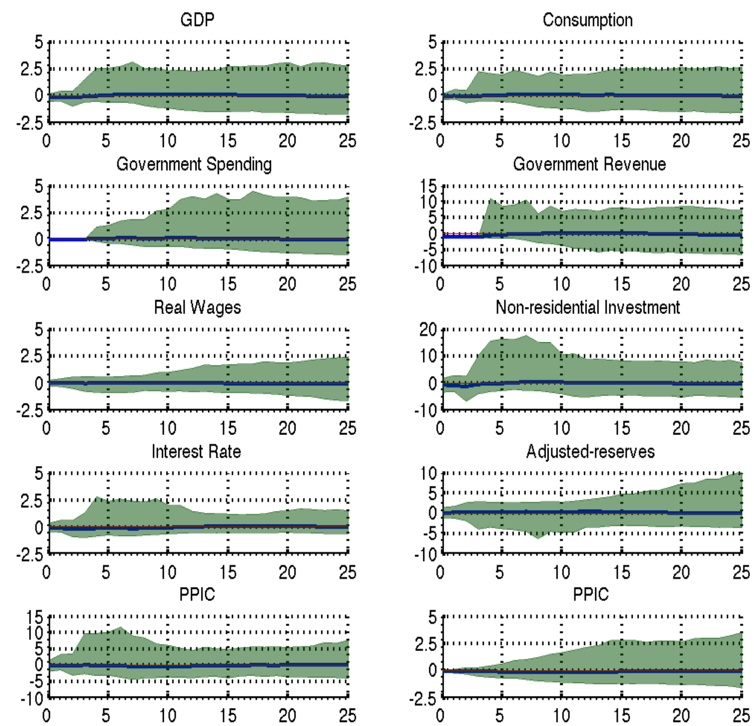


(b) The ARRW Methodology

Figure 15: IRFs to a Deficit-Spending Policy Shock

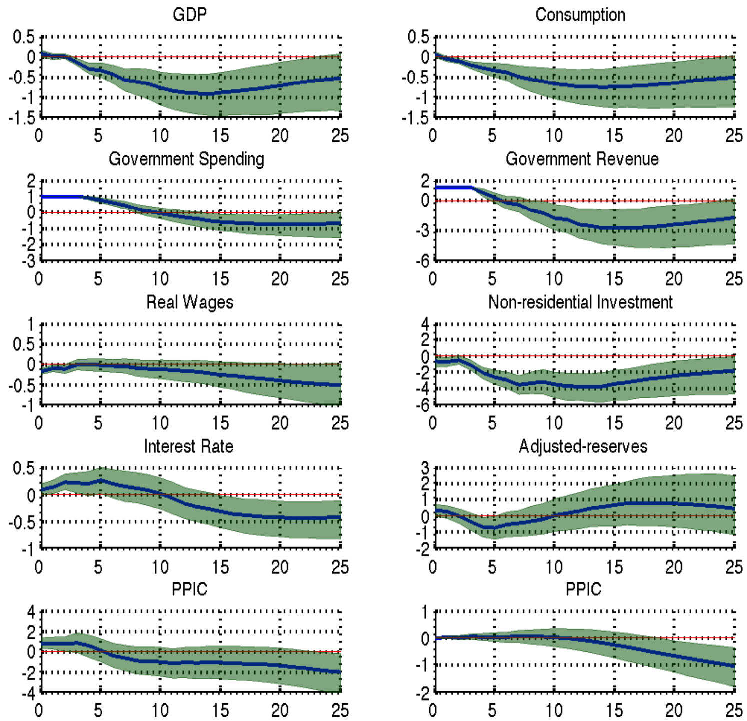


(a) The Penalty Function Approach

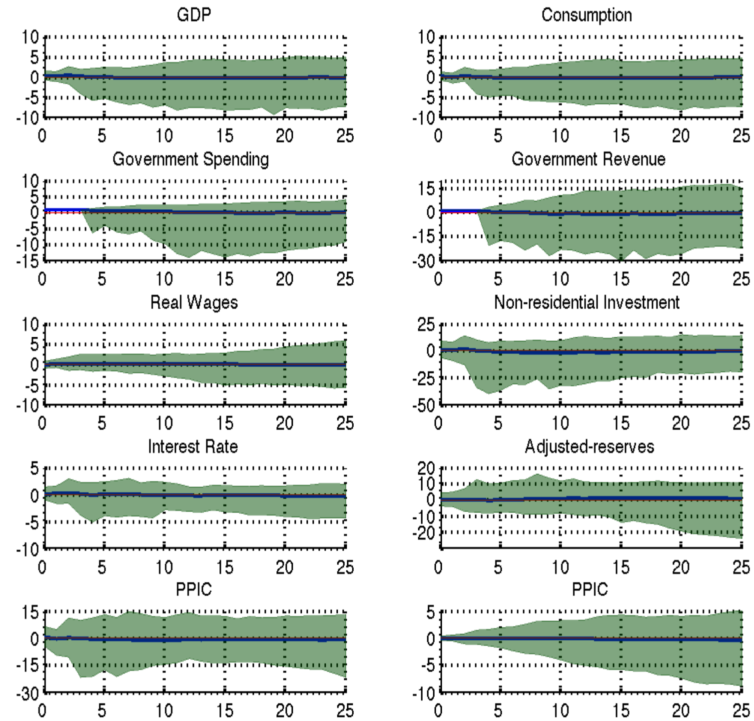


(b) The ARRW Methodology

Figure 16: IRFs to a Deficit-Financed Tax Cut Shock

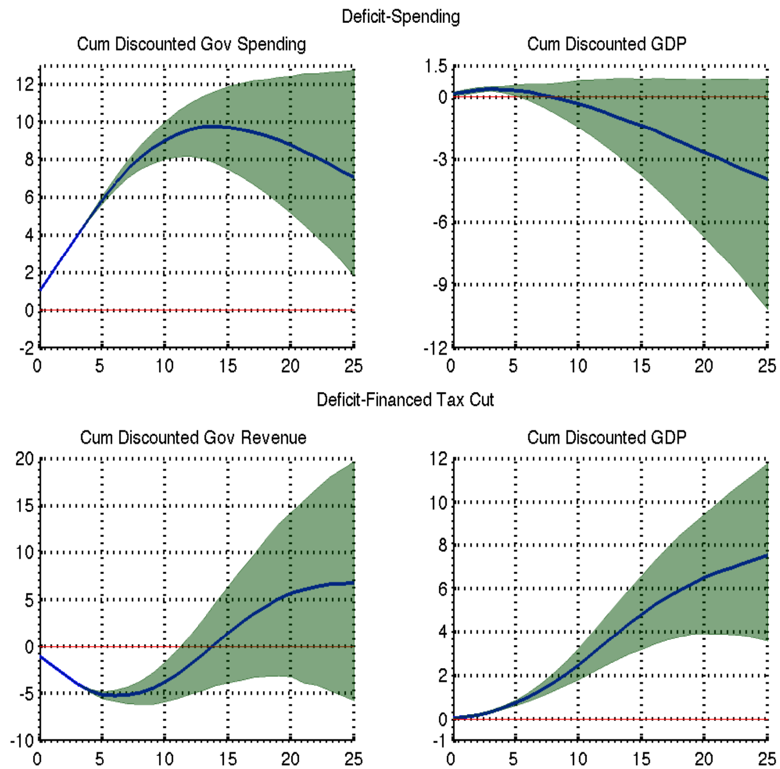


(a) The Penalty Function Approach

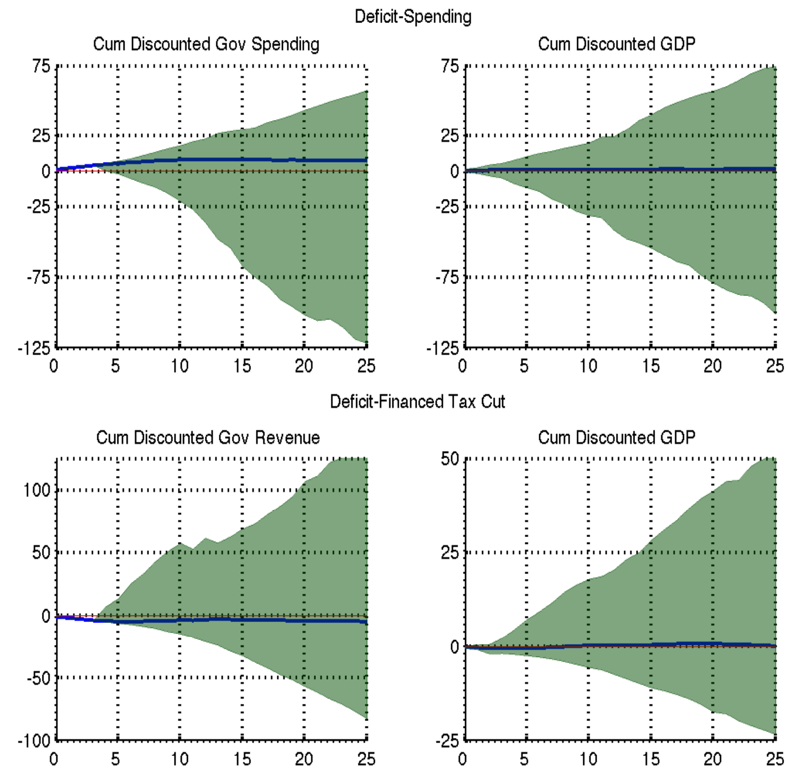


(b) The ARRW Methodolgoy

Figure 17: IRFs to a Balanced-Budget Shock



(a) The Penalty Function Approach



(b) The ARRW Methodology

Figure 18: Cumulative IRFs to Deficit-Spending and Deficit-FinancedTax Cut Shocks

References

Baumeister, C. and L. Benati (2010). Unconventional monetary policy and the great recession. *European Central Bank WP* (1258).

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?

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

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

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.

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.

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. *The 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? *The American Economic Review*, 54–81.

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 (2003a). A gibbs sampler for structural vector autoregressions. *Journal of Economic Dynamics and Control 28*(2), 349–366.

Waggoner, D. F. and T. Zha (2003b). 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.