Inference in Proxy SVARs*

Carsten Jentsch†
University of Mannheim

Kurt G. Lunsford‡
Federal Reserve Bank of Cleveland

January 2016

Abstract

Proxy structural vector autoregressions (SVARs) provide a method for identifying the impact of structural shocks in vector autoregressions by using external proxy variables that are correlated with the structural shocks of interest but uncorrelated with the other structural shocks. To produce confidence intervals for the impulse response functions (IRFs) from proxy SVARs, Mertens and Ravn (2013) set the precedent of using a recursive-design wild bootstrap. However, this method is not appropriate for structural IRFs, and simulations show that it produces coverage rates that are much too low. Instead of the wild bootstrap, we propose a modified residual-based moving block bootstrap and show that it produces more appropriate coverage rates in simulations. We use this moving block bootstrap to re-estimate confidence intervals for the IRFs from the tax shocks in Mertens and Ravn (2013) and find that inferences about the dynamic effects of tax changes on output, labor, or nonresidential investment cannot be made.

Keywords: Fiscal Policy, Residual-Based Moving Block Bootstrap, Structural Vector Autoregression, Tax Shocks, Wild Bootstrap

JEL Codes: C32, E23, E62, H24, H25, H31, H32

*The views expressed herein are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

†cjentsch@mail.uni-mannheim.de

‡kurt.lunsford@clev.frb.org

1 Introduction

Since the seminal work of Sims (1980), estimating the dynamic effects of structural shocks in vector autoregressions (VARs) has been important for research in macroeconomics. In recent contributions, Stock and Watson (2008, 2012), Montiel Olea, Stock, and Watson (2012), and Mertens and Ravn (2013) have developed a method for estimating structural vector autoregressions (SVARs) that use variables external from the VAR as proxies for the structural shocks of interest.

This proxy SVAR approach has proven to be very useful.[1] Mertens and Ravn (2013) use it to merge the SVAR literature on tax shocks (Blanchard and Perotti, 2002; Mountford and Uhlig, 2009) with the narrative approach taken by Romer and Romer (2010), Gertler and Karadi (2015) and Lunsford (2015a) use it to study the effects monetary policy shocks, Carriero et al. (2015) use it to study the effects of uncertainty shocks, and Stock and Watson (2012) use it to study the effects of a large number of economic shocks, including oil shocks, productivity shocks, uncertainty shocks, and financial shocks. In addition, Mumtaz, Pinter, and Theodoridis (2015) show that it does better in matching the effects credit supply shocks from a dynamic stochastic general equilibrium model than a Choleski decomposition, and Drautzburg (2015) uses it to estimate a Bayesian VAR and a dynamic stochastic general equilibrium model. Mertens and Ravn (2014) show that it can be used to reconcile the differences between structural VAR and narrative estimates of tax multipliers; however, Kliem and Kriwoluzky (2013) argue that it is not able to reconcile structural VAR and narrative estimates of monetary policy shocks. Finally, this proxy SVAR approach has been included as a standard method for identifying macroeconomic shocks in a recent handbook chapter (Ramey, 2015). Despite all of this research, little attention has been paid to estimating standard errors and confidence intervals when using the proxy SVAR method. In this paper, we study the bootstrap algorithms used to produce confidence intervals for the impulse response functions (IRFs) from proxy SVARs.

To produce the confidence intervals for the IRFs, Mertens and Ravn (2013) (MR) use a recursive-design wild bootstrap, and they advertise three appealing features of this method. First, Gonçalves and Kilian (2004) show that this bootstrap design is robust against conditional heteroskedasticity of unknown form. Second, it accounts for uncertainty in estimating the effects of structural shocks with proxy variables. Third, because MR's narrative proxy variables have many observations that are censored to zero, a bootstrap based on re-sampling would have a positive probability of drawing all zeros for a series of bootstrapped proxy variables. However, the wild bootstrap does not allow this event. Because of these benefits, MR set the precedent for producing confidence intervals, and their wild bootstrap method was also used by Mertens and Ravn (2014), Gertler and Karadi (2015), Lunsford (2015a) and Nakamura and Steinsson (2015).

[1] We follow the terminology of Mertens and Ravn (2013, 2014) by referring to the external variables as proxy variables. Because the covariance assumptions on these variables parallels those from the instrumental variables literature, they may also be called external instruments.

In contrast to the claims of MR, the first two features of the wild bootstrap are not true as advertised. With regard to conditional heteroskedasticity of unknown form, Gonçalves and Kilian (2004) only applies to the VAR coefficients and not to the covariance matrix of the VAR errors. Brüggemann, Jentsch, and Trenkler (2016) show that wild bootstrap methods cannot be used for inference on structural IRFs because the wild bootstrap does not correctly replicate the relevant features of the VAR errors. This is true even when the VAR errors are independent and identically distributed. More importantly, the wild bootstrap does not effectively account for the uncertainty of estimating the effects of structural shocks with proxy variables. This is because MR's wild bootstrap multiplies the VAR errors and the corresponding proxy variables by a bootstrap shock that takes a value of either 1 or -1. Then, when computing the bootstrapped covariance between the VAR errors and the proxy variables, the bootstrap shocks become multiplied against themselves and effectively drop out of the bootstrap algorithm. This causes the wild bootstrap algorithm to underestimate the uncertainty of the estimated structural shock. In a Monte Carlo simulation with a sample size of 250, which is similar to MR's sample size, the wild bootstrap's 68% confidence interval includes the actual initial impulse response in only 6% of the simulations, and the 95% confidence interval includes the actual initial impulse response in only 16% of simulations. Further, these coverage rates become even smaller as the sample size increases. These results suggest that MR's wild bootstrap produces confidence intervals for impulse response functions that much are too small.

To replace MR's wild bootstrap method, we propose a modified residual-based moving block bootstrap. Brüggemann, Jentsch, and Trenkler (2016) show that the residual-based moving block bootstrap is asymptotically valid for inference in SVARs with conditional heteroskedasticity of unknown form. We modify their algorithm so that the proxy variables are also re-sampled in moving blocks. In the same Monte Carlo simulation as the wild bootstrap, the moving block bootstrap's 68% confidence interval includes the actual initial impulse response in 63% of the simulations, and the 95% confidence interval includes the actual initial impulse response in 92% of simulations. Although these coverage rates are slightly too low, they improve as the sample size increases, and they are much better than the coverage rates from the wild bootstrap. Based on these simulation results and the theory in Brüggemann, Jentsch, and Trenkler (2016), we conjecture that this modified residual-based moving block bootstrap is asymptotically valid for inference in proxy SVARs. We intend to include a proof to support this conjecture in a future draft of this paper.

As an application of our modified residual-based moving block bootstrap method, we recreate a number of MR's figures, displaying the IRFs from tax cuts where the confidence intervals are produced with the moving block bootstrap. The primary result is that confidence intervals in all of MR's figures become much larger. With the wild bootstrap, MR find that average personal income tax rate (APITR) and average corporate income tax rate (ACITR) cuts have statistically significant impacts on many economic variables at 90% and 95% confidence levels. However, with the moving block bootstrap, many of MR's results are no longer inferable, even at 68% confidence levels. Most importantly, cuts to neither the

APITR nor the ACITR have statistically significant effects on output. Consistent with this, their are also no statistically significant effects on labor market variables or on nonresidential investment. As in MR, a cut to the APITR yields a statistically significant increase in durable goods purchases, and a cut the ACITR yields a statistically significant increase in nonresidential investment. However, the results are only significant at the 68% level – not the 90% or 95% level. Further, as discussed below, because the moving block bootstrap leads to moderately undersized coverage rates at smaller sample sizes, the statistical significance of these results should be taken with a grain of salt. The only result from MR that appears to be robust is that personal income tax revenues fall in response to a cut in APITRs.

To our knowledge, the only paper that addresses inference in proxy SVARs is Montiel Olea, Stock, and Watson (2015). They develop confidence intervals that are robust when a proxy variable is weakly correlated with the structural shock of interest in large samples, similar to the problem of a weak instrumental variable (Staiger and Stock, 1997). However, their confidence intervals only apply to the case where one proxy variable is used to identify one structural shock. In contrast, our moving block bootstrap can be used to produce confidence intervals when multiple proxy variables are used to identify multiple structural shocks, as is the case in Mertens and Ravn (2013) and Drautzburg (2015).

The rest of the paper proceeds as follows. Section 2 describes the proxy SVAR methodology. Section 3 describes both the wild and moving block bootstrap algorithms, and it evaluates their coverage rates with Monte Carlo simulations. Section 4 recreates MR's results using the moving block bootstrap, and Section 5 concludes.

2 Proxy Structural Vector Autoregressions

The proxy SVAR methodology begins with a standard SVAR set up. There is an $N \times 1$ vector of observable time-series variables that follows

$$Y_t = \delta_0 + \delta_1 Y_{t-1} + \cdots + \delta_p Y_{t-p} + u_t, \tag{1}$$

where u_t is an $N \times 1$ vector of VAR innovations. Let ϵ_t be the $N \times 1$ vector of structural shocks in the economy that have the following properties: $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{E}(\epsilon_t \epsilon_t') = I$, and $\mathbb{E}(\epsilon_t \epsilon_\tau') = 0$ for $\tau \neq t$. Then, the VAR innovations and the structural shocks are related according to

$$u_t = B\epsilon_t, \tag{2}$$

where B is an invertible $N \times N$ matrix. Thus, it is the case that

$$\mathbb{E}(u_t u_t') = BB'. \tag{3}$$

The objective here is to identify the effects of k of the structural shocks where $k < N$. To be precise, partition the structural shocks into $\epsilon_t = [\epsilon_{1,t}', \epsilon_{2,t}']'$ where $\epsilon_{1,t}$ is the $k \times 1$ vector that contains the structural shocks of interest, and $\epsilon_{2,t}$ is the $N-k \times 1$ vector of other structural shocks. Next, partition B into $B = [B_1, B_2]$ where B_1 is the $N \times k$ matrix of coefficients that correspond to the structural shocks of interest and B_2 is the $N \times N-k$ matrix of coefficients that correspond to the other shocks. Then, the objective here is to estimate B_1.

The difficulty in estimating B_1 is that $\epsilon_{1,t}$ is unobserved and Equation (3) only provides $(N+1)N/2$ moment restrictions for the N^2 elements of B. To provide additional moment restrictions, Stock and Watson (2008, 2012), Montiel Olea, Stock, and Watson (2012), and Mertens and Ravn (2013) introduce the proxy variable approach. They assume that there exists a $k \times 1$ vector of proxy variables, denoted by m_t, taken from outside of the VAR. These proxy variables are mean zero, $\mathbb{E}_t(m_t) = 0$, they are relevant for identification

$$\mathbb{E}(m_t \epsilon_{1,t}') = \Phi, \tag{4}$$

where Φ is an invertible $k \times k$ matrix, and they are exogenous from the other structural shocks

$$\mathbb{E}(m_t \epsilon_{2,t}') = 0. \tag{5}$$

It does not need to be the case that m_t is uncorrelated with $X_t = [1, Y_{t-1}', \ldots, Y_{t-p}']'$. When m_t and X_t are correlated, MR propose simply projecting m_t onto X_t and using the residuals as the proxy variables. Because this projection is always available, MR assume that in practice we can treat m_t as being uncorrelated with X_t.

When applying the proxy SVAR approach, m_t can come from a wide variety of sources. For example, MR follow the narrative approach of Romer and Romer (2009) to construct proxy variables for tax shocks, Gertler and Karadi (2015) follow the high frequency approach of Gürkaynak, Sack, and Swanson (2005) to construct proxy variables for monetary policy shocks, and Carriero et al. (2015) use Bloom's (2009) measure of uncertainty as a proxy for uncertainty shocks. Thus, the assumptions in Equation (4) and (5) allow for the merging of many different methods of identifying economic shocks with SVARs.

To implement the proxy SVAR approach, partition u_t and further partition B so that Equation (2) can be re-written as

$$\begin{bmatrix} u_{1,t} \\ (k \times 1) \\ u_{2,t} \\ (N-k \times 1) \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ (k \times k) & (k \times N-k) \\ B_{21} & B_{22} \\ (N-k \times k) & (N-k \times N-k) \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ (k \times 1) \\ \epsilon_{2,t} \\ (N-k \times 1) \end{bmatrix}. \tag{6}$$

Then, the objective is to estimate $B_1 = [B_{11}', B_{21}']'$. To do this, first notice that Equations (4) and (5) along with the partition in (6) imply

$$\mathbb{E}(m_t u_{1,t}') = \Phi B_{11}' \tag{7}$$

and

$$\mathbb{E}(m_t u'_{2,t}) = \Phi B'_{21}. \tag{8}$$

Jointly, Equations (7) and (8) yield

$$B_{21} B_{11}^{-1} = \left([\mathbb{E}(m_t u'_{1,t})]^{-1} \mathbb{E}(m_t u'_{2,t}) \right)'. \tag{9}$$

Because the right-hand side of Equation (9) can be estimated from the data, it provides additional moment restrictions on parameters of the model in addition to those in Equation (3) that help estimate B_{11} and B_{21}. Specifically, given an estimate of $B_{21}B_{11}^{-1}$ from Equation (9), Equation (3) and the partitions in Equation (6) are sufficient to estimate $B_{12}B'_{12}$, $B_{11}B'_{11}$, $B_{22}B'_{22}$ and $B_{12}B_{22}^{-1}$. Details of these estimations are provided in Appendix A.

When $k = 1$, the estimates following from Equations (3), (6) and (9) are sufficient to estimate B_{11} and B_{21} up to a sign convention. To see this, note that when $k = 1$, it is the case that B_{11} is a scalar. Thus, $B_{11} = \pm\sqrt{B_{11}B'_{11}}$ and Equation (9) can then be used to estimate B_{21}. When $k > 1$, as is the case in Mertens and Ravn (2013), identification becomes more complicated. When Φ is diagonal, then we can use each element of m_t in isolation as if we have multiple $k = 1$ proxy variables and no additional restrictions are needed. However, when Φ is not diagonal, we need additional restrictions. To get these restrictions, MR re-write the system in Equation (6) as

$$u_{1,t} = \eta u_{2,t} + S_1 \epsilon_{1,t} \tag{10}$$

and

$$u_{2,t} = \zeta u_{1,t} + S_2 \epsilon_{2,t}, \tag{11}$$

where $\eta = B_{12}B_{22}^{-1}$, $\zeta = B_{21}B_{11}^{-1}$, and $S_1 = (I - B_{12}B_{22}^{-1}B_{21}B_{11}^{-1})B_{11}$. Inversely, it is the case that

$$\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} I + \eta(I - \eta\zeta)^{-1}\zeta \\ (I - \eta\zeta)^{-1}\zeta \end{bmatrix} S_1 \tag{12}$$

so that B_{11} and B_{21} are functions of η, ζ and S_1. Here ζ can be estimated from Equation (9) and η can be estimated from Equations (3), (6) and (9), but S_1 cannot be estimated without additional restrictions. To get these restrictions, note that

$$S_1 S'_1 = (I - B_{12}B_{22}^{-1}B_{21}B_{11}^{-1})B_{11}B'_{11}(I - B_{12}B_{22}^{-1}B_{21}B_{11}^{-1})' \tag{13}$$

and

$$B_{11} = (I - B_{12}B_{22}^{-1}B_{21}B_{11}^{-1})^{-1} S_1. \tag{14}$$

Equation (13) provides an estimate of $S_1S'_1$. Given this, S_1 is the Choleski decomposition of $S_1S'_1$, imposing that S_1 is lower triangular. Then given S_1, B_{11} and B_{21} can be estimated from Equations (14) and (9). It is this Choleski decomposition that provides the additional restrictions needed to estimate the model by restricting how the shocks of interest can interact. For example, MR estimate the effects of changes to both average personal income tax rates (APITRs) and average corporate income tax rates (ACITRs), and the Choleski decomposition restricts how an APITR shock can effect the ACITR and vice versa. Suppose that APITR is ordered before the ACITR in Y_t. Then, the lower triangular Choleski decomposition implies that an APITR shock impacts the ACITR directly through $\epsilon_{1,t}$ and indirectly through $u_{2,t}$, while an ACITR shock only impacts the APITR indirectly though $u_{2,t}$.

3 Coverage of Bootstrapped Confidence Intervals

3.1 Description of the Bootstrap Algorithms

Once SVARs are estimated, it is common to compute the structural IRFs with respect to the structural shocks of interest. In this paper, we study two bootstrap methods for producing the confidence intervals for these IRFs. First, we study the recursive-design wild bootstrap used by MR. Second, we study the residual-based moving block bootstrap in Brüggemann, Jentsch, and Trenkler (2016), which we modify to include moving blocks of the proxy variables.

For both bootstrap algorithms, the VAR has p lags, Y_t is observed for $t = -p + 1, \ldots, 0, 1, \ldots, T$, and m_t is observed for $t = 1, \ldots, T$. Then, the VAR coefficients are estimated by least squares, and these coefficients and the VAR errors are denoted by $\hat{\delta}$ and $\hat{u}_t$ for $t = 1, \ldots, T$.

Given this notation, the algorithm for the recursive-design wild bootstrap is as follows:

1. Generate a random variable, e_t for $t = 1, \ldots, T$, that takes the value 1 with probability 1/2 and a value -1 with probability 1/2.

2. Use e_t for $t = 1, \ldots, T$ to produce $u_t^* = \hat{u}_t e_t$ for $t = 1, \ldots T$.

3. Choose an initial condition $Y_{-p+1}^*, \ldots, Y_0^*$ by randomly selecting a block of p observations of Y_t as in Berkowitz and Kilian (2000).

4. Use the initial condition from the previous step along with $\hat{\delta}$ and u_t^* to recursively produce Y_t^* for $t = 1, \ldots, T$.

5. Use $Y_{-p+1}^*, \ldots, Y_T^*$ to estimate bootstrapped VAR coefficients and innovations, $\hat{\delta}^*$ and $\hat{u}_t^*$ for $t = 1, \ldots, T$.

6. Use the random variable e_t to produce $m_t^* = m_t e_t$ for $t = 1, \ldots, T$.

7. Use $\hat{u}_t^*$ and m_t^* for $t = 1, \ldots, T$ to estimate $\hat{B}_1^*$.

8. Use $\hat{\delta}^*$ and $\hat{B}_1^*$ to produce the bootstrapped impulse response functions.

Repeat the algorithm $\mathcal{B}$ times and compute the $\alpha\%$ confidence interval to be the corresponding $\alpha\%$ percentile intervals.

To initialize the modified residual-based moving block bootstrap, the researcher much first choose a block length L and compute $J = [T/L]$, where $[\cdot]$ rounds up to the nearest integer so that $JL \geq T$. Next, collect the $N \times L$ blocks $\mathcal{U}_i = [\hat{u}_i, \ldots, \hat{u}_{i+L-1}]$ for $i = 1, \ldots T - L + 1$ and the $k \times L$ blocks $\mathcal{M}_i = [m_i, \ldots, m_{i+L-1}]$ for $i = 1, \ldots T - L + 1$. Then, the algorithm for the modified residual-based moving block bootstrap is as follows:

1. Randomly draw J integers from $i = 1, \ldots, T - L + 1$, putting equal probability on drawing any given i. Denote these draws as $i_1, \ldots, i_J$.
2. Collect the blocks $\mathcal{U}_{i_1}, \ldots, \mathcal{U}_{i_J}$ and drop the last $JL - T$ elements to produce $\tilde{u}_1^*, \ldots, \tilde{u}_T^*$.
3. Center $\tilde{u}_1^*, \ldots, \tilde{u}_T^*$ according to

$$u_{jL+s}^* = \tilde{u}_{jL+s}^* - \frac{1}{T - L + 1} \sum_{r=1}^{T-L} \hat{u}_{s+r-1} \tag{15}$$

   for $s = 1, \ldots, L$ and $j = 0, 1, \ldots, J - 1$ in order to produce $u_1^*, \ldots, u_T^*$.
4. Collect the blocks $\mathcal{M}_{i_1}, \ldots, \mathcal{M}_{i_J}$ and drop the last $JL - T$ elements to produce $\tilde{m}_1^*, \ldots, \tilde{m}_T^*$.
5. Center $\tilde{m}_1^*, \ldots, \tilde{m}_T^*$ similarly to the VAR errors in Equation (15) in order to produce $m_1^*, \ldots, m_T^*$.[2]
6. Set the initial condition $Y_{-p+1}^*, \ldots, Y_0^*$ equal to zero.
7. Use the initial condition from the previous step along with $\hat{\delta}$ and u_t^* to recursively produce Y_t^* for $t = 1, \ldots, T$.
8. Use $Y_{-p+1}^*, \ldots, Y_T^*$ to estimate bootstrapped VAR coefficients and innovations, $\hat{\delta}^*$ and $\hat{u}_t^*$ for $t = 1, \ldots, T$.
9. Use $\hat{u}_t^*$ and m_t^* for $t = 1, \ldots, T$ to estimate $\hat{B}_1^*$.
10. Use $\hat{\delta}^*$ and $\hat{B}_1^*$ to produce the bootstrapped impulse response functions.

As with the wild bootstrap, repeat the algorithm $\mathcal{B}$ times and compute the $\alpha\%$ confidence interval to be the corresponding $\alpha\%$ percentile intervals.

Brüggemann, Jentsch, and Trenkler (2016) show that this residual-based moving block bootstrap method is asymptotically valid for inference for both the VAR coefficients, δ,

[2] Because of the censoring in the MR proxy variables, I only apply the centering to the non-censored observations and leave the censored proxy variables with a value of zero.

and the covariance matrix of the VAR errors, $\mathbb{E}(u_t u_t')$, in the presence of conditional heteroskedasticity of unknown form. In contrast, the wild bootstrap is only asymptotically valid for inference for the VAR coefficients, but not the covariance matrix of the VAR errors. Further, Brüggemann, Jentsch, and Trenkler (2016) show that the wild bootstrap is not even valid for inference on the covariance matrix even when the VAR errors are i.i.d.. Because the structural IRFs are functions of both the VAR coefficients and the covariance matrix of the VAR errors, the wild bootstrap is unable to provide the correct inference for the IRFs. However, the moving block bootstrap is.

In order to accommodate the proxy variable framework, we have added several steps to Brüggemann, Jentsch, and Trenkler's (2016) moving block bootstrap. To initialize the algorithm, we create blocks of the proxy variable in manner parallel to the creation of the blocks of VAR errors. In steps 4 and 5 of the algorithm, we bootstrap the proxy variables by re-sampling and centering them so that they correspond to the re-sampled and centered blocks of VAR errors. Finally, in step 9, we use the bootstrapped proxy variables and estimated bootstrapped VAR errors to bootstrap the structural identification.

Despite its asymptotic validity, the moving block bootstrap does have one important drawback when applied to proxy SVARs. When a large number of the observations of m_t are censored to zero, as is the case in the MR, and when the block size L is sufficiently small, it may be the case that some of the blocks $\mathcal{M}_i$ contain only zeros. Then, there is a possibility that the bootstrapped proxy variables $m_1^*, \ldots, m_t^*$ contain only zeros. In contrast, it will never be the case that $m_1^*, \ldots, m_t^*$ contains only zeros in the wild bootstrap method. However, as we will discuss below, this does not seem to be an important issue in practice.

3.2 Monte Carlo Simulations

To study these two bootstrap algorithms, we use Monte Carlo simulations with two different data generating processes (DGPs): one with i.i.d. VAR errors and one with VAR errors that follow GARCH(1,1) processes. We assume that Y_t is 2×1 and follows the VAR(1) process

$$Y_t = \begin{bmatrix} 0.2 & 0 \\ 0.5 & 0.5 \end{bmatrix} Y_{t-1} + u_t, \quad \mathbb{E}(u_t u_t') = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix},$$

which closely follows the DGP used in Kilian (1998). To produce the simulated u_t, we assume that u_t follows Equation (2) where ϵ_t is 2×1 and

$$B = \begin{bmatrix} -0.592 & 0.806 \\ 0.592 & 0.806 \end{bmatrix}.$$

In the i.i.d. simulation, we assume that each element of ϵ_t is an independent standard normal random variable. In the GARCH(1,1) simulation, we assume that each element of ϵ_t is independent and follows

$$\epsilon_{i,t} = h_{i,t} w_{i,t},$$

and

$$h_{i,t}^2 = \beta_0 + \beta_1 \epsilon_{i,t-1}^2 + \beta_2 h_{i,t-1}^2,$$

for $i = 1, 2$. Here, $w_{i,t}$ for $i = 1, 2$ are independent standard normal random variables, $\beta_1 = 0.5$, $\beta_2 = 0.45$ and $\beta_0 = 1 - \beta_1 - \beta_2$.

We assume that $k = 1$ so that the proxy variable is a scalar. To simulate the proxy variable, we follow Equation (8) from Mertens and Ravn (2013), which is

$$m_t = D_t(\Gamma \epsilon_{1,t} + v_t),$$

where D_t is a dummy variable that takes a value of either 1 or 0 to account for censoring,[3] and v_t has the properties $\mathbb{E}(v_t) = 0$, $\mathbb{E}(v_t v_t') = \Sigma_{vv'}$, $\mathbb{E}(v_t v_\tau') = 0$ for $\tau \neq t$, and $\mathbb{E}(v_t \epsilon_{1,t}') = 0$. For simplicity, we assume that D_t is independent of both $\epsilon_{1,t}$ and v_t. For the simulations, we assume that $D_t = 0$ with probability 0.8. In addition, we assume that $\Phi = 0.5$, giving an implied value of $\Gamma = 2.5$. Finally, we assume that v_t is a standard normal random variable.

We assume that $\epsilon_{1,t}$ is the structural shock of interest so that the first column of B is used to produce the structural IRFs. In addition, we follow the approach taken in Mertens and Ravn (2013) and normalize the size of the shock so that the initial response of the first element of Y_t is -1 to produce the true IRFs and each of the bootstrapped IRFs.

For each bootstrap method, we run the Monte Carlo simulations with effective samples sizes 100, 250, 500 and 1000. For each simulation, we draw ϵ_t and compute u_t for $t = 1, \ldots, 1000 + T$, where T is the relevant effective sample size. In the i.i.d. DGP, we then use $Y_0 = 0$ and u_t for $t = 1, \ldots, 1000 + T$ to recursively generate Y_t for $t = 1, \ldots, 1000 + T$. In the GARCH(1,1) DGP, we use $Y_0 = 0$, $h_{i,0}^2 = 1$ for $i = 1, 2$ and $\epsilon_{i,0}^2 = 1$ for $i = 1, 2$ to recursively generate ϵ_t and Y_t for $t = 1, \ldots, 1000 + T$. We then throw away the first 999 observations Y_t and use $Y_{1000}, \ldots, Y_{1000+T}$ to estimate $\hat{\delta}$ and $\hat{u}_t$ before running the bootstrap.

For the residual-based moving block bootstrap, we use block lengths of 16, 20, 24, and 28 for the sample sizes 100, 250, 500, and 1000, respectively. These block lengths yield $J = [7, 13, 21, 36]$. From Brüggemann, Jentsch, and Trenkler (2016), these block lengths must satisfy $L \to \infty$ and $L^3/T \to 0$ as $T \to \infty$. Because of this, we follow the rule $L = \kappa T^{1/4}$ where κ is normalized so that a sample size of $T = 250$ corresponds exactly to $L = 20$. This yields $\kappa = 5.03$.

For each Monte Carlo trial, we use 1000 simulations and 2000 bootstrap replications. Then, we compute the coverage rate of a confidence interval to be the fraction of simulations where the true parameter of interest lies within the confidence interval. Table 1 shows the coverage rates for the 68% confidence intervals for each bootstrap method and all four sample sizes. Table 2 shows the coverage rates for the 95% confidence intervals for each bootstrap method and all four sample sizes. In both tables, we display the coverage rates for the first 6 impulse responses. For the IRFs, we use $Y_{1,t}$ to denote the first element of Y_t and $Y_{2,t}$ to denote the second element of Y_t.

[3] When $k > 1$, MR assume that this variable is either a $k \times k$ matrix of zeros or the identity matrix.

Table 1: Coverage Rates of the 68% Confidence Intervals

Wild Bootstrap (i.i.d DGP):								
Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.12	1.00	0.06	1.00	0.05	1.00	0.04
2	0.38	0.36	0.34	0.35	0.35	0.35	0.35	0.37
3	0.37	0.44	0.35	0.41	0.34	0.41	0.36	0.43
4	0.41	0.47	0.40	0.47	0.40	0.45	0.42	0.47
5	0.46	0.50	0.45	0.50	0.44	0.50	0.47	0.51
6	0.48	0.51	0.49	0.53	0.48	0.52	0.50	0.55

Moving Block Bootstrap (i.i.d. DGP):								
Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.65	1.00	0.63	1.00	0.67	1.00	0.66
2	0.65	0.64	0.61	0.62	0.67	0.66	0.67	0.63
3	0.64	0.63	0.62	0.64	0.67	0.65	0.65	0.62
4	0.61	0.61	0.61	0.63	0.65	0.65	0.64	0.62
5	0.59	0.59	0.61	0.62	0.65	0.64	0.63	0.62
6	0.57	0.57	0.61	0.61	0.65	0.63	0.63	0.64

Wild Bootstrap (GARCH DGP):								
Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.19	1.00	0.15	1.00	0.11	1.00	0.10
2	0.43	0.44	0.42	0.44	0.46	0.43	0.47	0.49
3	0.43	0.51	0.41	0.50	0.42	0.50	0.46	0.55
4	0.49	0.53	0.49	0.53	0.50	0.54	0.55	0.59
5	0.51	0.54	0.53	0.57	0.56	0.57	0.61	0.63
6	0.52	0.53	0.57	0.57	0.59	0.60	0.64	0.62

Moving Block Bootstrap (GARCH DGP):								
Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.68	1.00	0.66	1.00	0.66	1.00	0.64
2	0.70	0.66	0.64	0.64	0.65	0.63	0.64	0.63
3	0.63	0.65	0.64	0.64	0.63	0.62	0.61	0.61
4	0.61	0.62	0.64	0.64	0.62	0.63	0.60	0.61
5	0.59	0.59	0.65	0.63	0.60	0.61	0.61	0.61
6	0.56	0.56	0.62	0.61	0.61	0.62	0.61	0.62

Table 2: Coverage Rates of the 95% Confidence Intervals

Wild Bootstrap (i.i.d DGP):

Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.30	1.00	0.16	1.00	0.11	1.00	0.09
2	0.67	0.67	0.61	0.62	0.62	0.63	0.63	0.64
3	0.68	0.74	0.63	0.72	0.63	0.70	0.63	0.72
4	0.73	0.77	0.70	0.77	0.69	0.77	0.72	0.78
5	0.76	0.79	0.75	0.79	0.76	0.81	0.77	0.82
6	0.77	0.80	0.79	0.82	0.80	0.85	0.82	0.87

Moving Block Bootstrap (i.i.d. DGP):

Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.93	1.00	0.92	1.00	0.94	1.00	0.95
2	0.94	0.91	0.91	0.91	0.93	0.92	0.94	0.93
3	0.91	0.90	0.91	0.91	0.93	0.92	0.93	0.93
4	0.89	0.87	0.91	0.90	0.92	0.93	0.93	0.94
5	0.88	0.86	0.90	0.90	0.92	0.93	0.92	0.94
6	0.85	0.86	0.90	0.89	0.93	0.93	0.94	0.95

Wild Bootstrap (GARCH DGP):

Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.45	1.00	0.31	1.00	0.24	1.00	0.24
2	0.73	0.75	0.72	0.70	0.73	0.72	0.78	0.78
3	0.74	0.80	0.69	0.79	0.71	0.82	0.75	0.85
4	0.78	0.82	0.77	0.83	0.80	0.85	0.86	0.88
5	0.79	0.83	0.83	0.87	0.83	0.87	0.89	0.90
6	0.82	0.84	0.87	0.88	0.86	0.88	0.91	0.91

Moving Block Bootstrap (GARCH DGP):

Impulse	$T = 100$		$T = 250$		$T = 500$		$T = 1000$	
	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$	$Y_{1,t}$	$Y_{2,t}$
1	1.00	0.93	1.00	0.94	1.00	0.94	1.00	0.94
2	0.94	0.92	0.94	0.93	0.94	0.94	0.93	0.93
3	0.91	0.93	0.92	0.93	0.93	0.93	0.94	0.93
4	0.90	0.90	0.92	0.91	0.93	0.93	0.93	0.93
5	0.88	0.89	0.91	0.90	0.93	0.92	0.92	0.92
6	0.88	0.88	0.91	0.90	0.92	0.92	0.92	0.91

Both Table 1 and Table 2 show that the initial impulse of $Y_{1,t}$ has 100% coverage regardless of the confidence interval percent, bootstrap method or sample size. This is a result of the Mertens and Ravn (2013) normalization that produces $Y_{1,t} = -1$ for all bootstrap replications. Thus, we will ignore this particular entry when discussing the coverage of the different bootstrap methods. Given this caveat, Table 1 and Table 2 show that the wild bootstrap method produces coverage rates that are too small, especially for the initial impulse responses. This is true for both the i.i.d. and the GARCH(1,1) DGPs and both the 68% and 95% confidence intervals. What is particularly striking is how small the coverage rates are. For $T = 250$, which most closely corresponds to MR's effective sample size of 224, the coverage rates for the initial response for the 68% confidence interval are only 6% in the i.i.d. simulation and 15% in the GARCH(1,1) simulation. The 95% confidence intervals are also very small with coverage rates of 16% and 31% at a sample size of 250. In addition, the coverage of the initial impulse responses gets worse as the sample size increases, implying that more data pushes the wild bootstrap further away from the true confidence intervals.

The reason for the poor coverage rates of the wild bootstrap procedure is that it does not effectively account for the uncertainty in estimating the covariances in Equations (3), (7) and (8). Brüggemann, Jentsch, and Trenkler (2016) also make this point with regard to the covariance matrix of the VAR errors in Equation (3). Because these covariances are central to estimating B_1, the wild bootstrap underestimates the variance B_1. Further, because the initial impulse response is just B_1 scaled by a shock, the underestimated variance of B_1 produces the small coverage rates for the initial impulse responses observed in Tables 1 and 2 when using the wild bootstrap.

To see why the wild bootstrap underestimates the variance of B_1, we temporarily consider a simpler specification than the VAR and assume that u_t can be observed directly and does not need to be estimated from the VAR. Then, the wild bootstrap estimate of $\Phi B_{11}'$ from Equation (7) is given by

$$\widehat{\Phi^* B_{11}^{*\prime}} = T^{-1} \sum_{t=1}^{T} m_t^* u_{1,t}^{*\prime}.$$

Because $u_{1,t}^* = u_{1,t} e_t$ and $m_t^* = m_t e_t$ and e_t equals 1 or -1, it is the case that

$$\widehat{\Phi^* B_{11}^{*\prime}} = T^{-1} \sum_{t=1}^{T} m_t u_{1,t}',$$

which is simply the non-bootstrapped sample estimate. That is, when u_t is directly observable, the wild bootstrap yields $\widehat{\Phi^* B_{11}^{*\prime}} = \widehat{\Phi B_{11}'}$ for every bootstrap replication and implies that there is no uncertainty in the estimate of this covariance. The same is also true for the covariances in Equations (3) and (8). Thus, when u_t is directly observable, the wild bootstrap produces $\hat{B}_1^* = \hat{B}_1$ for every bootstrap replication and the bootstrapped variance of B_1 is zero.

Going back to the VAR, it is not the case that u_t is directly observable. Thus, in the bootstrap, we use $\hat{u}_t^*$ rather than u_t^* to estimate the covariances. Because $\hat{u}_t^*$ is different for each bootstrap replication, it will not be the case that $\hat{B}_1^* = \hat{B}_1$ for every bootstrap replication, and the bootstrapped variance of B_1 will not be zero. However, as shown in Tables 1 and 2, these differences in $\hat{u}_t^*$ for each bootstrap replication are not sufficient to properly estimate the variance of B_1, and the coverage rates of the confidence intervals are too small. Further, as the sample size increases, $\hat{u}_t^*$ becomes an increasingly accurate estimate of u_t^*. This causes the bootstrapped variance of B_1 to converge to zero as the sample size increases and explains why the coverage rates of the wild bootstrap shrink as the sample grows. Hence, in contrast to claim of Mertens and Ravn (2013), the wild bootstrap does not effectively account for uncertainty when estimating B_1 even though steps 6 and 7 of the wild bootstrap algorithm subject m_t to the wild bootstrap shock and re-estimate B_1 for every bootstrap replication.

The under-sized variance of B_1 in the wild bootstrap algorithm motivates the use of the residual-based moving block bootstrap. As shown by Brüggemann, Jentsch, and Trenkler (2016), this bootstrap method is asymptotically valid for inference on the covariance of the VAR errors in Equation (3), even in the presence of conditional heteroskedasticity of unknown form. Further, Tables 1 and 2 indicate that this bootstrap has good coverage rates for the initial impulse response for both the i.i.d. and the GARCH(1,1) DGPs and for both the 68% and 95% confidence intervals. This suggests that the moving block bootstrap can also properly account for the uncertainty in the covariances in Equations (7) and (8) when conducting inference on B_1.

After the initial impulse response, the moving block bootstrap produces coverage rates that moderately undersized. However, this problem becomes less severe as the sample size increases. The reason for this is that the moving block structure produces less re-sampling than in a traditional residual-based i.i.d. bootstrap, yielding smaller confidence intervals. Given the asymptotic requirement that $L^3/T \to 0$ as $T \to \infty$, this is a problem at small sample sizes but vanishes asymptotically. In practice, this leads to a trade-off for the choice of block length, L. Blocks that are too small will not be able to effectively account for the conditional heteroskedasticity, but blocks that are too large will not produce enough re-sampling. For example, a block length of 1 will produce the appropriate coverage rates in the i.i.d. data generating process, but will produce coverage rates that are too small in the GARCH(1,1) data generating process. A noted above, we set a block length of 20 for a sample size of 250, striking a balance between getting good coverage in the presence of conditional heteroskedasticity without producing coverage rates that are too low in the i.i.d. case.

Despite this problem of modestly undersized coverage rates, the moving block bootstrap still produces coverage rates that are consistently higher than the wild bootstrap – especially for the initial impulse response. Because of this, the moving block bootstrap is more appropriate for producing confidence intervals for IRFs from proxy SVARs than the wild bootstrap. As an application of the moving block bootstrap, we recreate several of the results from Mertens and Ravn (2013) in the next section.

4 The Effects Tax Changes in the United States

As an application of the residual-based moving block bootstrap, we recreate a number of figures from Mertens and Ravn (2013) using this bootstrap method instead of the wild bootstrap. MR study the dynamic effects of two types of tax changes on the U.S. economy: average personal income income tax rates (APITRs) and average corporate income tax rates (ACITRs). To do this, they construct narrative accounts of both shocks by decomposing Romer and Romer's (2009) narrative account of postwar tax changes. Then, they use these narrative accounts as proxy variables for the tax shocks in a SVAR as described in Section 2. In this section, we use MR's replication files to recreate several of their key figures with the residual-based moving block bootstrap instead of the wild bootstrap.[4]

We begin with MR's baseline specification, which estimates a seven variable VAR that includes the APITR, the ACITR, the logarithm of the personal income tax base (PITB), the logarithm of the corporate income tax base (CITB), the logarithm of government spending, the logarithm of GDP divided by population, and the logarithm of government debt held by the public divided by the GDP deflator and population.[5] The data are available quarterly from 1950:Q1 to 2006:Q4. We include a constant in the VAR, and we estimate the VAR with four lags, giving an effective sample size of 224. Given this sample size, we use a block length of 19 based on the rule used for the Monte Carlo simulations above.

Following MR, we consider two orderings for the variables. In the first ordering, the APITR comes before the ACITR, and in the second ordering, the ACITR comes before the APITR. As discussed at the end of Section 2, this ordering influences how the shocks in the model will impact one another. Figure 1 displays the IRFs of a 1% cut to the APITR for both orderings along with the 68% confidence intervals produced using the moving block bootstrap with 10,000 replications. Blue solid lines are the point estimate when the APITR is ordered first, and the blue dashed lines are the corresponding confidence intervals. Red diamonds are the point estimate when the ACITR is ordered first, and the red dashed lines are the corresponding confidence intervals. This figure parallels Figure 2 in Mertens and Ravn (2013) but with two important differences. First, the confidence intervals here are produced using the moving block bootstrap. Second, the confidence intervals presented by

[4] The replication files that we use are from the American Economic Association's website at `https://www.aeaweb.org/articles.php?doi=10.1257/aer.103.4.1212`.

[5] In their replication files, MR define the APITR as federal personal income tax revenues including contributions to government social insurance divided by personal income tax base, the ACITR as federal corporate income tax revenues divided by corporate income tax base, the PITB as personal income less government transfers plus contributions to government social insurance divided by GDP deflator and by population, the CITB as corporate profits less Federal Reserve Bank profits divided by GDP deflator and by population, and government spending as real Federal government consumption and investment expenditures divided by population.

MR are 95% intervals – not the 68% intervals presented here. We present the 68% intervals here because the 95% intervals with moving block bootstrap are huge, and no inferences can be drawn from them. We show the 95% confidence intervals in Appendix B.

When using the wild bootstrap, MR find that a cut to the APITR causes increases in output and in the PITB with a drop in personal income tax revenues that are statistically significant with 95% confidence intervals. With the moving block bootstrap, despite the smaller confidence percent, the confidence intervals here are larger than those in MR. Because of this, Figure 1 indicates that, in contrast to MR, no inference can be made about the effects of an APITR cut on PITB or output, even at a 68% level. However, personal income tax revenues do fall with statistical significance.[6] Further, unlike in MR where the confidence intervals were similar for both orderings, Figure 1 indicates that the confidence intervals can be quite different depending on the ordering. This is particularly apparent for the confidence interval on output. When the ACITR is ordered first the confidence interval is much narrower. At the third step in the IRF, which gives the peak point estimate, this confidence interval is [-0.5%, 2.8%]. In contrast, the confidence interval when the APITR is ordered first is [-3.3%, 4.8%].

Figure 2 displays the IRFs of a 1% cut to the ACITR for both orderings along with the 68% confidence intervals produced using the moving block bootstrap with 10,000 replications. Blue solid lines are the point estimate when the APITR is ordered first, and the blue dashed lines are the corresponding confidence intervals. Red diamonds are the point estimate when the ACITR is ordered first, and the red dashed lines are the corresponding confidence intervals. This figure parallels Figure 3 in Mertens and Ravn (2013) but with the same differences as Figure 1 above. The same figure with 95% confidence intervals can be found in Appendix B.[7]

When using the wild bootstrap, MR find that a cut to the ACITR causes increases in output and the CITB that are statistically significant with 95% confidence intervals. However, with the moving block bootstrap, no inference can be made about the effects of an ACITR cut on output. Further, as with the APITR cut, the confidence intervals surrounding output in Figure 2 are noticeably different depending on the ordering of the variables. With the moving block bootstrap, inference about the effect of an ACITR cut on the CITB is ambiguous. When the APITR is ordered first, it appears that the CITB increases in the first year after the shock by a statistically significant amount. However, this result disappears when the ACITR is ordered first. Further, as shown in Appendix B, this result is insignificant for both orderings at the 95% level. Finally, Figure 2 suggests that a cut to the ACITR will increase government purchases by a statistically significant amount after a year and a half. However, as in MR, this result disappears at the 95% level.

As noted in Section 3 above, one drawback of the moving block bootstrap algorithm

[6] Personal income tax revenues are not a variable included in the VAR. Rather, following MR, we compute them from the other IRFs as APITR/0.1667 + PITB.

[7] Note that corporate income tax revenues are not included in the VAR. Following MR, they are computed from the other IRFs as ACITR/0.2996 + CITB.

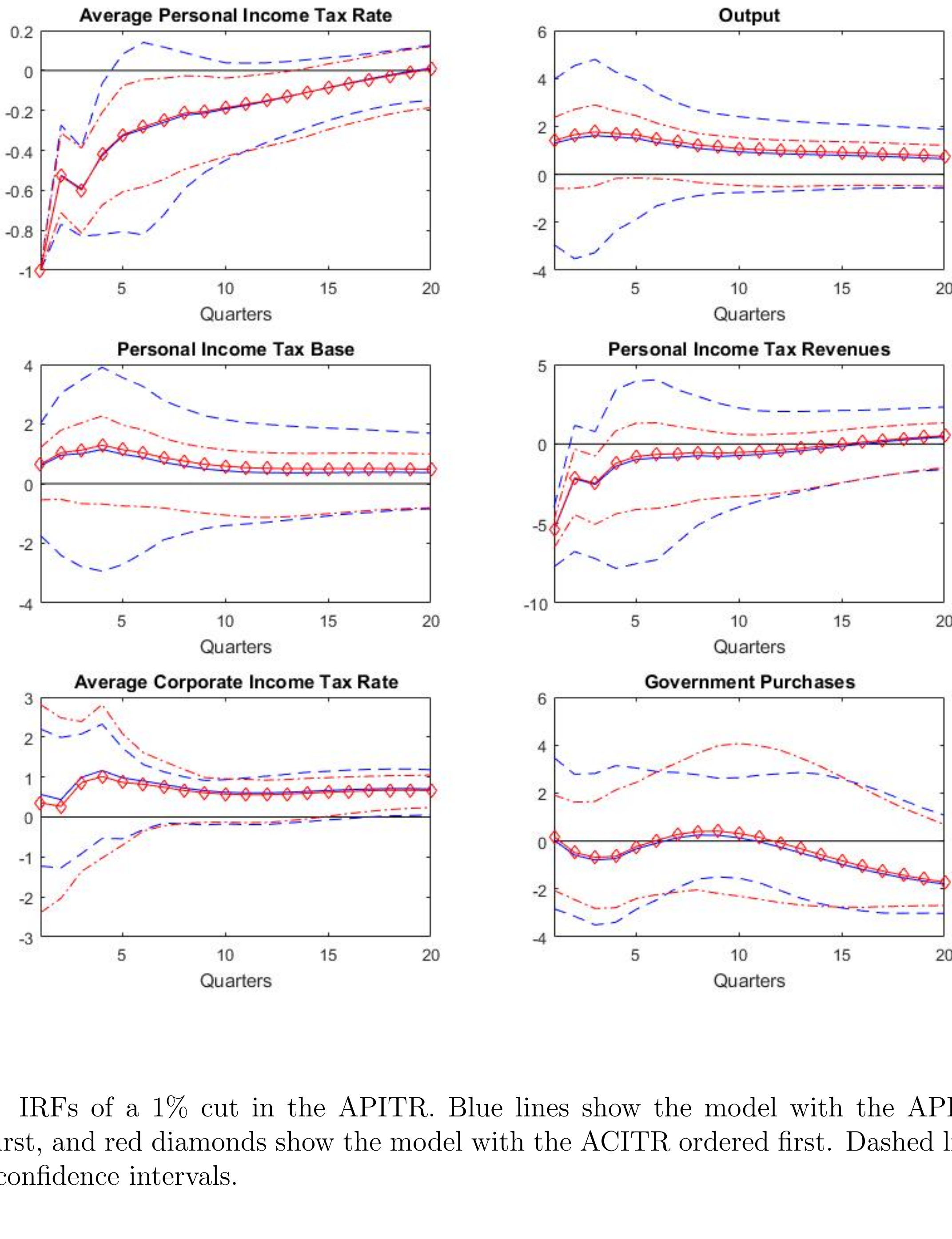

Figure 1: IRFs of a 1% cut in the APITR. Blue lines show the model with the APITR ordered first, and red diamonds show the model with the ACITR ordered first. Dashed lines are 68% confidence intervals.

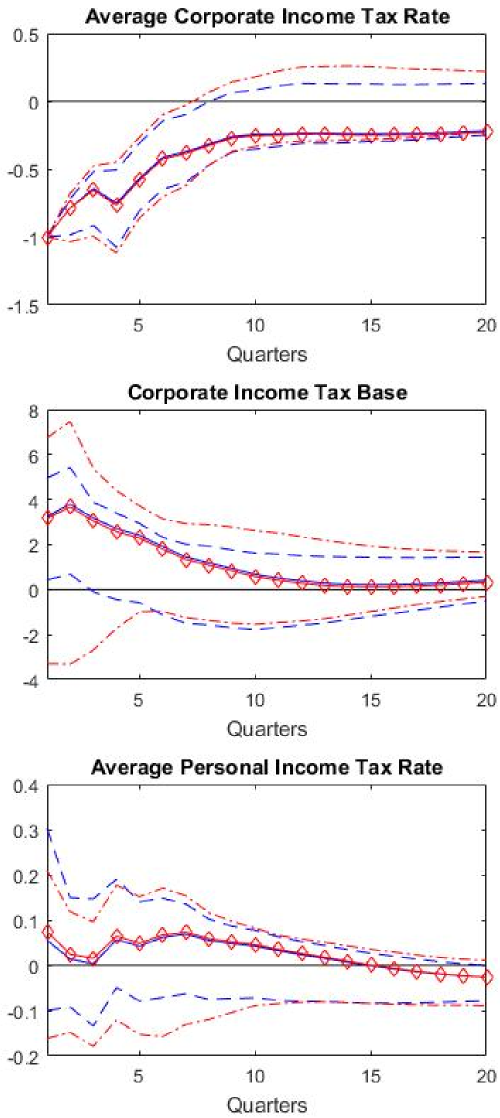

Figure 2: IRFs of a 1% cut in the ACITR. Blue lines show the model with the APITR ordered first, and red diamonds show the model with the ACITR ordered first. Dashed lines are 68% confidence intervals.

is that the possibility exists for all of the bootstrapped proxy variable observations to be censored to zero. This problem arises because many of the observations in MR's narrative account are censored to zero. Of the 224 observed proxy APITR shocks, 211 (94%) are zero. For the ACITR, this number is 208 (93%). Further, there can be many periods between these proxied shocks. For the APITR, the largest gap between shocks is 39 quarters, and for the ACITR, the largest gap is 49 quarters. Because of this, a block length of 19 does not guarantee that a shock will be observed in every block. However, given this block length, the moving block bootstrap algorithm draws 12 block and the probability of drawing all 12 blocks of zeros is very small. For the APITR, 153 of the 206 blocks contain at least one non-censored shock, implying that the probability of drawing a block of zeros is 0.257. Then, the probability of drawing 12 blocks of zeros will be $0.257^{12} = 8 \times 10^{-8}$. For the ACITR, 162 of the 206 blocks contain at least one non-censored shocks, yielding a probability of 9×10^{-9} of drawing 12 blocks of zeros. Hence, even with 10,000 bootstrap replications, the probability of drawing one replication of all zeros is very small, and this is not an important drawback in practice. In producing Figures 1 and 2, all of the computed bootstrap replications contained at least 3 non-zero observations for both the APITR and the ACITR proxies.

As in MR, we also consider the effects of APITR and ACITR cuts on labor market variables, consumption and investment. To study the effect on labor market variables, we follow MR and estimate an eight variable VAR that includes the APITR, the ACITR, the logarithm of government spending, the logarithm of GDP divided by population, the logarithm of government debt held by the public divided by the GDP deflator and population, the logarithm of total economy employment divided by population, the logarithm of total economy hours worked divided by total economy employment, and the logarithm of labor force divided by population. As with the previous VARs, the sample is 1950:Q1 to 2006:Q4, and the VAR includes a constant and four lags. To produce confidence intervals, we use the moving block bootstrap with 10,000 replications and block length of 19.

The left hand panels of Figure 3 display the IRFs of the labor market variables to a 1% APITR cut, and the right hand panels of Figure 3 display the corresponding IRFs to a 1% ACITR cut. Solid blue lines are the point estimates, and the dashed lines show the 68% and 90% confidence intervals. This figure corresponds to Figure 9 in Mertens and Ravn (2013).[8] With the wild bootstrap, MR find statistically significant increases in the employment to population ratio and hours per worker along with a statistically significant decrease in the unemployment rate to an APITR cut. However, Figure 3 shows that no such inferences can be made, even at the 68% level. Rather, as with output, cuts to both the APITR and the ACITR have no inferable impact on the labor market in the United States.

To study the effects of APITR and ACITR cuts on consumption, we estimate an eight variable VAR that includes the APITR, the ACITR, the logarithm of the PITB, the logarithm of government spending, the logarithm of GDP divided by population, the logarithm

[8]The APITR is ordered first for all IRFs in this figure. Also, the unemployment rate is not included in the VAR. Following MR, we compute it from the other IRFs as $5.25\{\exp[-0.9475(\text{Employment/Pop} - \text{Labor Force/Pop})/5.25] - 1\}$.

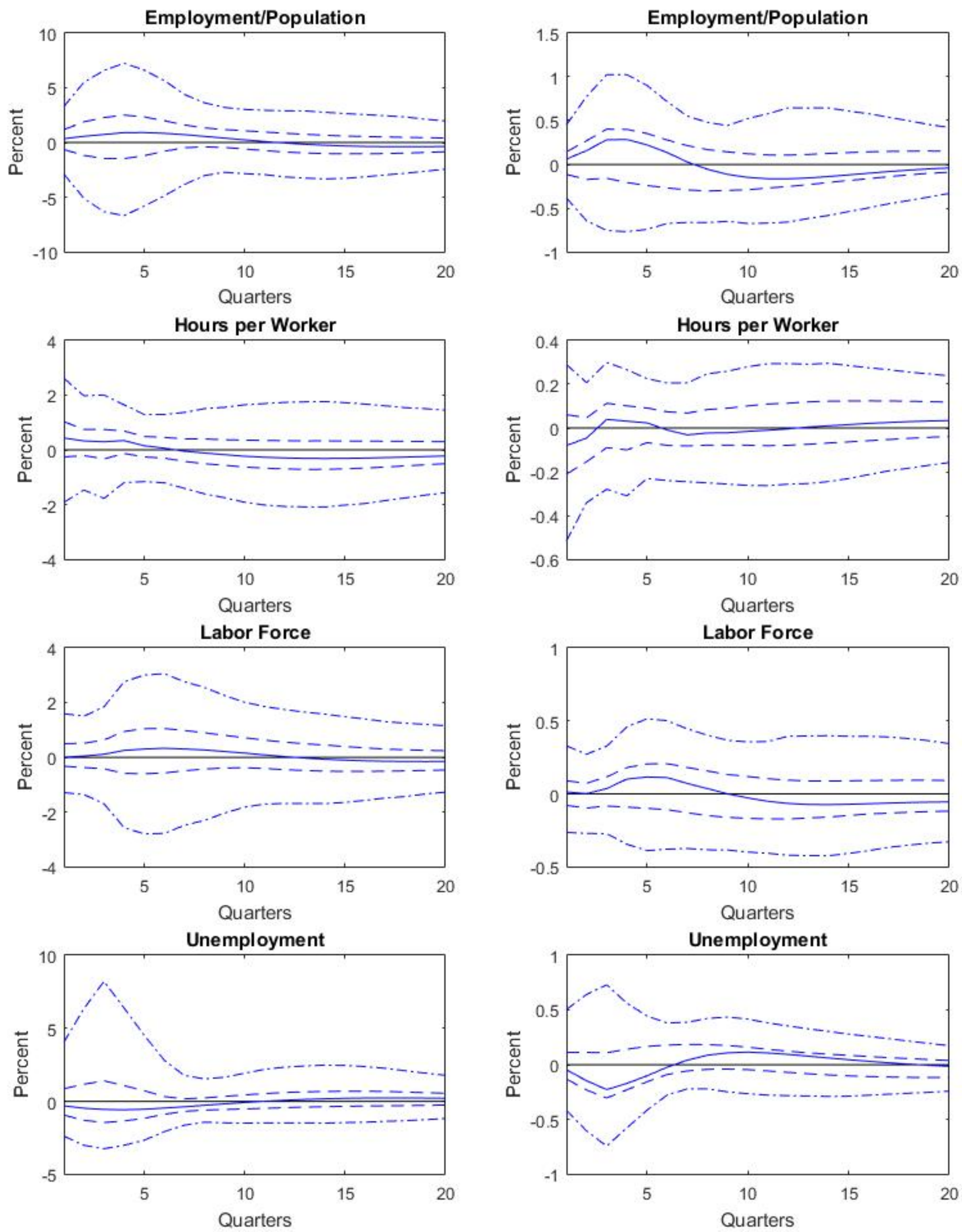

Figure 3: The left hand panels display IRFs of a 1% cut in the APITR. The right hand panels display IRFs of a 1% cut in the APITR. Blue lines show the point estimates and the dashed lines show the 68% and 90% confidence intervals.

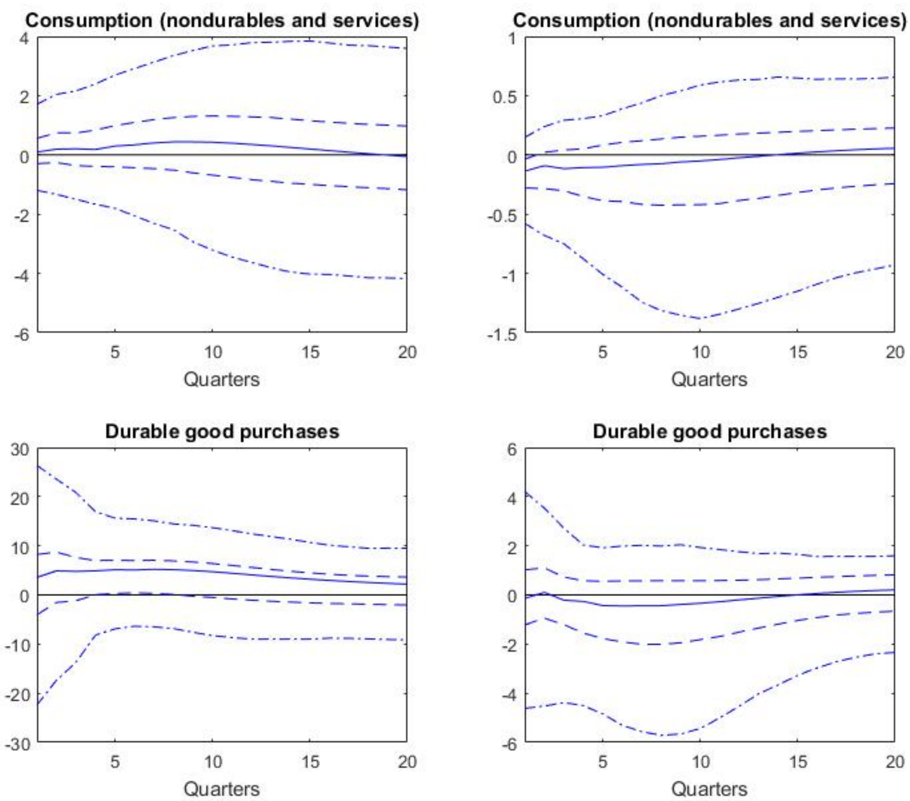

Figure 4: The left hand panels display IRFs of a 1% cut in the APITR. The right hand panels display IRFs of a 1% cut in the APITR. Blue lines show the point estimates and the dashed lines show the 68% and 90% confidence intervals.

of government debt held by the public divided by the GDP deflator and population, the logarithm of chain-aggregated nondurable consumption and service goods consumption divided by population, and logarithm of real durable consumption goods expenditures divided by population. The left hand panels of Figure 4 display the IRFs of consumption and durable good purchases to a 1% APITR cut, and the right hand panels of Figure 4 display the corresponding IRFs to a 1% ACITR cut. Solid blue lines are the point estimates, and the dashed lines show the 68% and 90% confidence intervals. This figure corresponds to the top panels in Figure 10 in Mertens and Ravn (2013).[9] With the wild bootstrap, MR find a statistically significant increase in durable good purchases at both the 90% and 95% levels in response to an APITR cut. However, Figure 3 shows that no such inferences can be made at the 90% level. There is a statistically significant positive response at the 68% level beginning in quarter 4. However, as noted in the Monte Carlo simulations, the coverage rates for the

[9]The APITR is ordered first for all IRFs in this figure.

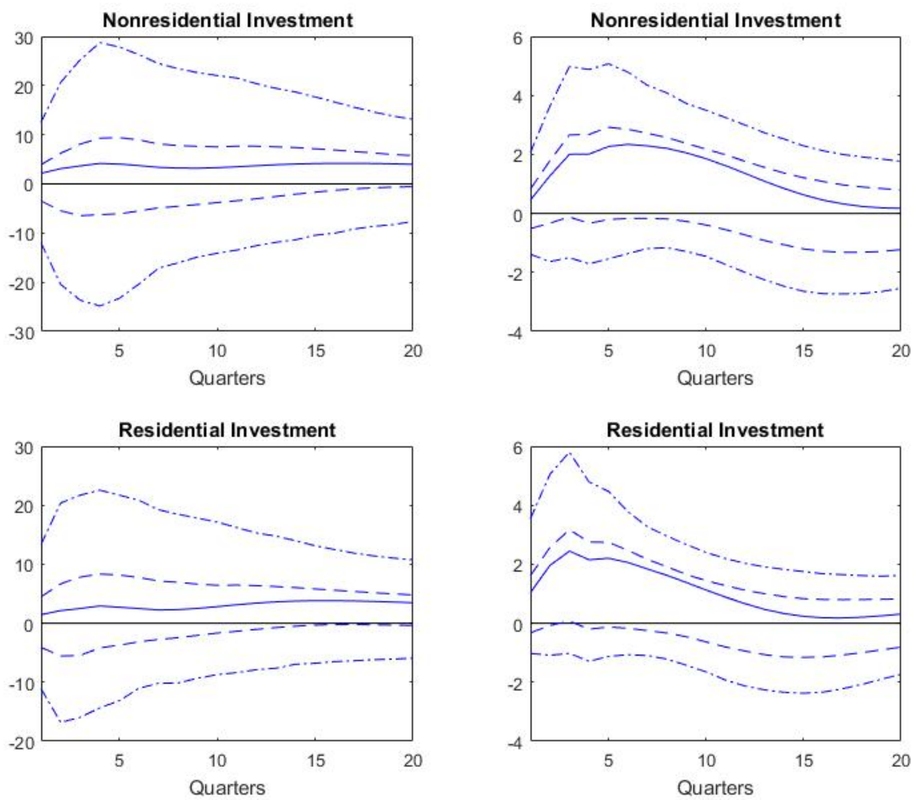

Figure 5: The left hand panels display IRFs of a 1% cut in the APITR. The right hand panels display IRFs of a 1% cut in the APITR. Blue lines show the point estimates and the dashed lines show the 68% and 90% confidence intervals.

moving block bootstrap are modestly undersized at multi-step horizons. When the confidence intervals are re-computed using a 72% level, this statistical significance disappears, suggesting that this statistical significance at the 68% should taken with a grain of salt.

Finally, to study the effects of APITR and ACITR cuts on investment, we estimate a 8 variable VAR that includes the APITR, the ACITR, the logarithm of the CITB, the logarithm of government spending, the logarithm of GDP divided by population, the logarithm of government debt held by the public divided by the GDP deflator and population, the logarithm of real non-residential fixed investment divided by population, and the logarithm of real residential fixed investment divided by population. The left hand panels of Figure 5 display the IRFs of nonresidential and residential investment to a 1% APITR cut, and the right hand panels of Figure 5 display the corresponding IRFs to a 1% ACITR cut. Solid blue lines are the point estimates, and the dashed lines show the 68% and 90% confidence intervals. This figure corresponds to the bottom panels in Figure 10 in Mertens and Ravn

(2013).[10] With the wild bootstrap, MR find a statistically significant increase in nonresidential investment at both the 90% and 95% levels in response to both an APITR and an ACITR cut. Further, they find a statistically significant increase in residential investment at both the 90% and 95% levels in response to both an ACITR cut. In contrast, Figure 5 shows no statistically significant response of nonresidential investment at either the 68% or 90% level to either tax cut. Figure 5 does show that there is a statistically significant response in residential investment at the 68% level to an ACITR cut. However, this result only holds for one period, and the same caveat with regard to durable good purchases applies. When the confidence intervals are re-computed using a 72% level, the statistical significance here also disappears. Thus, given the modestly undersized coverage rates in the Monte Carlo simulation, this statistical significance at the 68% should be taken with a grain of salt.

5 Conclusions

Estimating the dynamic effects of structural shocks from SVARs is important for macroeconomic research. Recently, Stock and Watson (2008, 2012), Montiel Olea, Stock, and Watson (2012), and Mertens and Ravn (2013) developed a method for estimating SVARs that uses an external proxy variable that is correlated with the structural shocks of interest but uncorrelated with the other structural shocks. This paper studies methods for inference when using this proxy SVAR method. We find that the commonly used recursive-design wild bootstrap produces confidence intervals that much too small. Alternatively, we propose a residual-based moving block bootstrap algorithm, and show via simulation that this produces confidence intervals with much better coverage rates than the wild bootstrap. Based on these simulation results and the theory in Brüggemann, Jentsch, and Trenkler (2016), we conjecture that this moving block bootstrap is asymptotically valid for inference in proxy SVARs and intent to prove so in future research.

When the moving block bootstrap is applied to Mertens and Ravn (2013), we find that many of their results are no longer statistically significant. Specifically, cuts to both personal and corporate tax rates have no inferable effect on output or nonresidential investment. Cuts to personal rates have no inferable effect on employment, hours worked per worker or the unemployment rate. These results suggest that the narrative proxy variables used by MR are not informative enough to discern the dynamic effects of tax changes on economic activity in the United States. However, these results do not imply that proxy SVARs will always be uninformative. As an example, Lunsford (2015b) shows that inferences can be made at the 90% level with the moving block bootstrap when using Fernald's (2014) utilization-adjusted total factor productivities as proxy variables. Thus, proxy SVARs and the moving block bootstrap are useful tools for inferring the dynamic effects of structural shocks.

[10]The APITR is ordered first for all IRFs in this figure.

Appendix A: Derivation of the Estimators

For the purposes of notation, define $\mathbb{E}(m_t u'_{1,t}) = \Sigma_{mu'_1}$, $\mathbb{E}(m_t u'_{2,t}) = \Sigma_{mu'_2}$, $\mathbb{E}(u_{1,t} u'_{1,t}) = \Sigma_{11}$, $\mathbb{E}(u_{2,t} u'_{1,t}) = \Sigma_{21}$, and $\mathbb{E}(u_{2,t} u'_{2,t}) = \Sigma_{22}$ to be the moments that can be estimated from the data. Then, Equation (9) can be re-written as

$$(\Sigma_{mu'_1}^{-1} \Sigma_{mu'_2})' = B_{21} B_{11}^{-1}. \tag{A1}$$

Next, Equations (3) and (6) imply

$$\Sigma_{11} = B_{11} B'_{11} + B_{12} B'_{12}, \tag{A2}$$

$$\Sigma_{21} = B_{21} B'_{11} + B_{22} B'_{12}, \tag{A3}$$

and

$$\Sigma_{22} = B_{21} B'_{21} + B_{22} B'_{22}. \tag{A4}$$

Using Equations (A2) through (A4), it is the case that

$$\begin{aligned} \Sigma_{21} - B_{21} B_{11}^{-1} \Sigma_{11} &= B_{21} B'_{11} + B_{22} B'_{12} - B_{21} B_{11}^{-1} (B_{11} B'_{11} + B_{12} B'_{12}) \\ &= (B_{22} - B_{21} B_{11}^{-1} B_{12}) B'_{12}. \end{aligned}$$

Next, define

$$\begin{aligned} Z &= (B_{22} - B_{21} B_{11}^{-1} B_{12})(B_{22} - B_{21} B_{11}^{-1} B_{12})' \\ &= B_{22} B'_{22} - B_{21} B_{11}^{-1} B_{12} B'_{22} - B_{22} B'_{12} (B_{21} B_{11}^{-1})' + B_{21} B_{11}^{-1} B_{12} B'_{12} (B_{21} B_{11}^{-1})' \\ &= B_{21} B'_{21} + B_{22} B'_{22} - B_{21} B'_{21} - B_{21} B_{11}^{-1} B_{12} B'_{22} \\ &\quad - B_{21} B'_{21} - B_{22} B'_{12} (B_{21} B_{11}^{-1})' + B_{21} B'_{21} + B_{21} B_{11}^{-1} B_{12} B'_{12} (B_{21} B_{11}^{-1})' \\ &= B_{21} B'_{21} + B_{22} B'_{22} - B_{21} B_{11}^{-1} B_{11} B'_{21} - B_{21} B_{11}^{-1} B_{12} B'_{22} \\ &\quad - B_{21} B'_{11} (B_{21} B_{11}^{-1})' - B_{22} B'_{12} (B_{21} B_{11}^{-1})' + B_{21} B_{11}^{-1} B_{11} B'_{11} (B_{21} B_{11}^{-1})' + B_{21} B_{11}^{-1} B_{12} B'_{12} (B_{21} B_{11}^{-1})' \\ &= B_{21} B'_{21} + B_{22} B'_{22} - B_{21} B_{11}^{-1} (B_{11} B'_{21} + B_{12} B'_{22}) \\ &\quad - (B_{21} B'_{11} + B_{22} B'_{12})(B_{21} B_{11}^{-1})' + B_{21} B_{11}^{-1} (B_{11} B'_{11} + B_{12} B'_{12})(B_{21} B_{11}^{-1})' \\ &= \Sigma_{22} - B_{21} B_{11}^{-1} \Sigma'_{21} - \Sigma_{21} (B_{21} B_{11}^{-1})' + B_{21} B_{11}^{-1} \Sigma_{11} (B_{21} B_{11}^{-1})'. \end{aligned}$$

Given this list of equations, it is the case that

$$\begin{aligned} B_{12} B'_{12} &= B_{12} [(B_{22} - B_{21} B_{11}^{-1} B_{12})^{-1} (B_{22} - B_{21} B_{11}^{-1} B_{12})]' [(B_{22} - B_{21} B_{11}^{-1} B_{12})(B_{22} - B_{21} B_{11}^{-1} B_{12})^{-1}] B'_{12} \\ &= (\Sigma_{21} - B_{21} B_{11}^{-1} \Sigma_{11})' Z^{-1} (\Sigma_{21} - B_{21} B_{11}^{-1} \Sigma_{11}). \end{aligned}$$

Then, to estimate the model, begin by estimating Equation (A1). Next,

$$Z = \Sigma_{22} - B_{21} B_{11}^{-1} \Sigma'_{21} - \Sigma_{21} (B_{21} B_{11}^{-1})' + B_{21} B_{11}^{-1} \Sigma_{11} (B_{21} B_{11}^{-1})'.$$

Then,

$$B_{12}B_{12}' = (\Sigma_{21} - B_{21}B_{11}^{-1}\Sigma_{11})'Z^{-1}(\Sigma_{21} - B_{21}B_{11}^{-1}\Sigma_{11}),$$
$$B_{11}B_{11}' = \Sigma_{11} - B_{12}B_{12}',$$
$$B_{22}B_{22}' = \Sigma_{22} - B_{21}B_{11}^{-1}B_{11}B_{11}'(B_{21}B_{11}^{-1})',$$

and

$$B_{12}B_{22}^{-1} = [\Sigma_{21}' - B_{11}B_{11}'(B_{21}B_{11}^{-1})'](B_{22}B_{22}')^{-1}.$$

Appendix B: The Baseline Specification with 95% Confidence Intervals

Figures 6 and 7 display the baseline specification used to produce Figures 1 and 2 above, but with 95% confidence intervals from the residual-based moving block bootstrap.

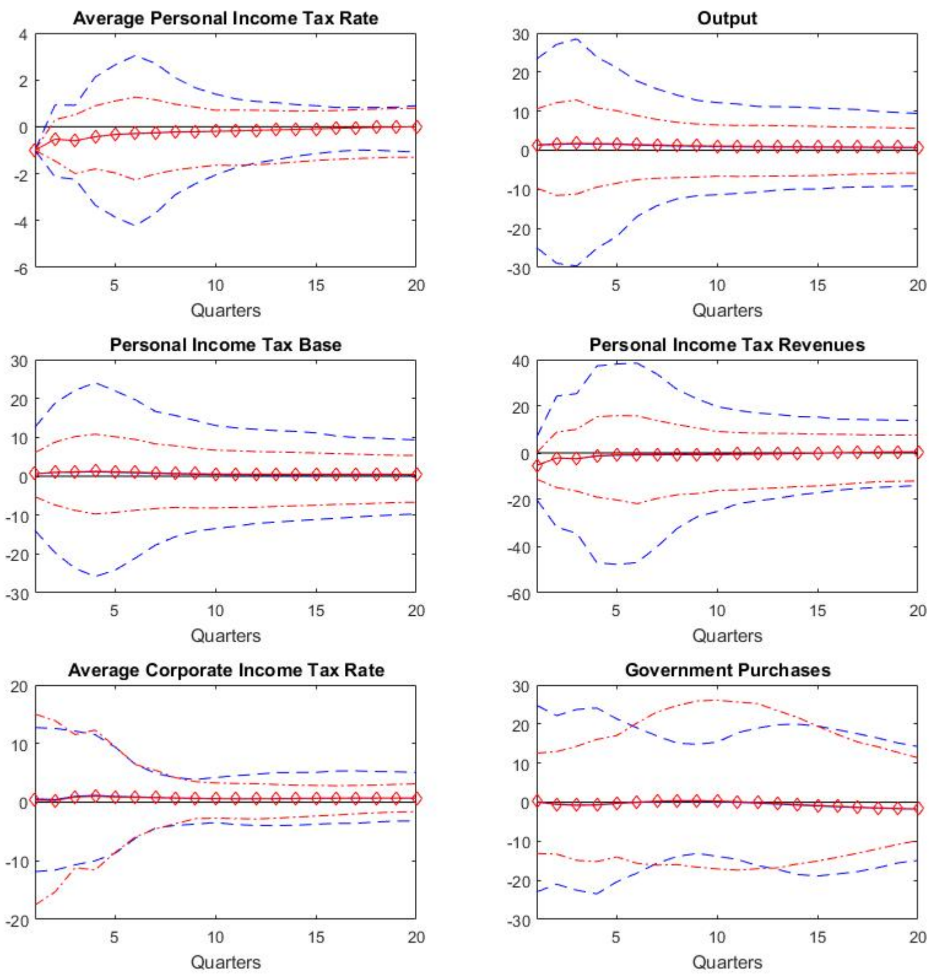

Figure 6: IRFs of a 1% cut in the APITR. Blue lines show the model with the APITR ordered first, and red diamonds show the model with the ACITR ordered first. Dashed lines are 95% confidence intervals.

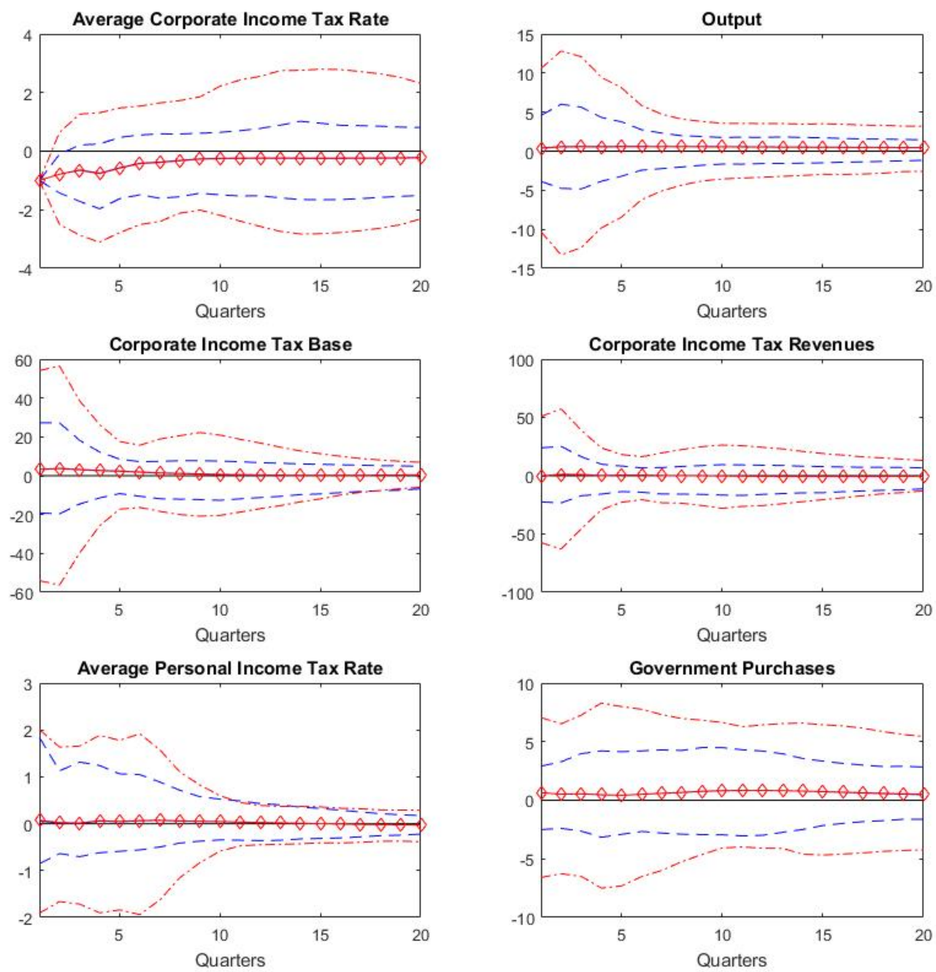

Figure 7: IRFs of a 1% cut in the ACITR. Blue lines show the model with the APITR ordered first, and red diamonds show the model with the ACITR ordered first. Dashed lines are 95% confidence intervals.

References

Berkowitz, Jeremy and Lutz Kilian. 2000. "Recent Developments in Bootstrapping Time Series." *Econometric Reviews* 19 (1):1–48.

Blanchard, Olivier and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *Quarterly Journal of Economics* 117 (4):1329–1368.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3):623–685.

Brüggemann, Ralf, Carsten Jentsch, and Carsten Trenkler. 2016. "Inference in VARs with Conditional Heteroskedasticity of Unknown Form." *Journal of Econometrics* 191 (1):69–85.

Carriero, Andrea, Haroon Mumtaz, Konstantinos Theodoridis, and Angeliki Theophilopoulou. 2015. "The Impact of Uncertainty Shocks under Measurement Error: A Proxy SVAR Approach." *Journal of Money, Credit and Banking* 47 (6):1223–1238.

Drautzburg, Thorsten. 2015. "A Narrative Approach to a Fiscal DSGE Model." Working Paper, Federal Reserve Bank of Philadelphia.

Fernald, John. 2014. "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity." Working Paper, Federal Reserve Bank of San Francisco.

Gertler, Mark and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics* 7 (1):44–76.

Gonçalves, Sílvia and Lutz Kilian. 2004. "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form." *Journal of Econometrics* 123 (1):89–120.

Gürkaynak, Refet S, Brian Sack, and Eric T. Swanson. 2005. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking* 1 (1):55–93.

Kilian, Lutz. 1998. "Small-Sample Confidence Intervals for Impulse Response Functions." *Review of Economics and Statistics* 80 (2):218–230.

Kliem, Martin and Alexander Kriwoluzky. 2013. "Reconciling Narrative Monetary Policy Disturbances with Structural VAR Model Shocks?" *Economics Letters* 121 (2):247–251.

Lunsford, Kurt G. 2015a. "Housing Search and Fluctuations in Residential Construction." Job Market Paper, University of Wisconsin – Madison.

———. 2015b. "Identifying Structural VARs with a Proxy Variable and a Test for a Weak Proxy." Federal Reserve Bank of Cleveland Working Paper no. 15-28.

Mertens, Karel and Morten O. Ravn. 2013. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." *American Economic Review* 103 (4):1212–1247.

———. 2014. "A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers." *Journal of Monetary Economics* 68 (Supplement):S1–S19.

Montiel Olea, José Luis, James H. Stock, and Mark W. Watson. 2012. "Inference in Structural VARs with External Instruments." Presentation Slides, Harvard University.

———. 2015. "Uniform Inference in SVARs with External Instruments." Working Paper, New York University.

Mountford, Andrew and Harald Uhlig. 2009. "What Are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics* 24 (6):960–992.

Mumtaz, Haroon, Gabor Pinter, and Konstantinos Theodoridis. 2015. "What Do VARs Tell Us About the Impact of a Credit Supply Shock?" Working Paper No. 739 Queen Mary University of London.

Nakamura, Emi and Jón Steinsson. 2015. "High Frequency Identification of Monetary Non-Neutrality." Working Paper, Columbia University.

Ramey, Valerie A. 2015. "Macroeconomic Shocks and Their Propagation." Working Paper, University of California, San Diego.

Romer, Christina D. and David H. Romer. 2009. "A Narrative Analysis of Postwar Tax Changes." Unpublished manuscript, University of California, Berkeley.

———. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review* 100 (3):763–801.

Sims, Christopher A. 1980. "Macroeconomics and Reality." *Econometrica* 48 (1):1–48.

Staiger, Douglas and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65 (3):557–586.

Stock, James H. and Mark W. Watson. 2008. "Lecture 7 - Recent Developments in Structural VAR Modeling." Presentd at the NBER Summer Institute Econometrics Lectures: What's New in Economics - Time Series, July 15.

———. 2012. "Disentangling the Channels of the 2007-09 Recession." *Brookings Papers on Economic Activity* Spring:81–135.