Efficient Inference in Non-Gaussian Structural VARs

JONATHAN H. WRIGHT

October 18, 2017

Abstract

This paper proposes the use of nonparametric rank based estimation of structural vector autroregressions. Identification can be by a Cholesky ordering, or by the method of external instruments. In cases where the structural shocks are non-Gaussian, the proposed estimator brings efficiency gains relative to the conventional estimation method.

PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CIRCULATE.

1 Introduction

This paper considers rank based approaches to estimation of a structural vector autoregression (SVAR). Structural VARs are very widely used tools in empirical macroeconomics, going back to the seminal work of Sims (1980). They are multivariate linear models in which identifying assumptions are used to disentangle the underlying sources of disturbances. The original identification scheme proposed by Sims (1980) involved a Cholesky ordering, but other identification schemes have been proposed, including external instruments. See Kilian and Lütkepohl (2017) for a comprehensive recent overview of structural VARs.

Nonparametric rank approaches to inference have remarkable efficiency properties. In the case of a simple linear regression with iid errors with symmetric density f, there exists a rank estimator that has identical asymptotic distribution to the pseudo-Gaussian MLE if the density, f, is in fact Gaussian, but is more efficient for any other symmetric density f. This important result is due to Chernoff and Savage (1958). There is a very large statistics literature on the potential efficiency gains from nonparametric rank approaches to inference including Hájek and Šidák (1967), Jaeckel (1972), Hettmansperger (1984) and Hettmansperger and McKean (1988). Nonetheless, the potential efficiency gains from rank estimation are not generally a focus of the econometrics literature[1].

This paper demonstrates how nonparametric rank approaches can be applied to a structural VAR identified by either the Cholesky ordering or external instruments. The proposed approach has the potential for efficiency gains if the shocks are non-Gaussian. The efficiency gains are small, unless the shocks are very fat-tailed. But shocks may have very fat-tailed distributions, and in any case structural VARs are sufficiently widely used that modest efficiency gains might still be worthwhile. The proposed approach has the drawback

[1] Andrews and Marmer (2008) and Hasan and Koenker (1997) are however papers giving econometric applications of rank-based methods, with emphasis on their efficiency gains under non-normality.

that it requires the structural shocks to be independent, not just uncorrelated. However, if we are thinking of structural shocks as representing entirely separate underlying economic disturbances, such as fiscal and monetary policy, that does not seem to be an unreasonable assumption.

A number of papers, including Lanne and Lütkepohl (2010) and Lanne et al. (2017), assume that structural shocks have different distributions and use this information for identification. That's a potentially valuable approach to identification, but it is different from what I am envisioning here. I am sticking with the conventional Cholesky or external instruments identifications and making no assumption that structural shocks have different distributions. The proposed procedure will be consistent and asymptotically efficient if all the structural shocks have Gaussian distributions. However, if the shocks are not all Gaussian, then the proposed procedure may bring efficiency gains.

Also, a number of papers, including Hallin and Paindaveine (2004), have considered multivariate rank tests for independence that are invariant to rotations of the variables. For many purposes rotation invariance is a desirable characteristic. However, for identification of a structural VAR where specific structural shocks are identified using economic restrictions, this is not suitable as we want to pin down a specific rotation of the variables.

The plan for the remainder of this paper is as follows. Section 2 describes rank based estimation of a structural VAR. Section 3 contains Monte-Carlo simulation results. Section 4 gives an empirical application and section 5 concludes.

2 Structural VAR estimation

Consider the reduced form VAR for an nx1 vector Y_t:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} ... + A_p Y_{t-p} + \varepsilon_t \tag{2.1}$$

where ε_t is an iid vector of reduced form shocks with mean zero and covariance matrix Σ. Assume that there is a vector of orthonormal structural shocks, η_t, such that $\varepsilon_t = R\eta_t$. The identification problem involves solving for the rotation matrix R.

2.1 External Instruments

One approach to the identification problem is to assume that there exists a variable Z_t with mean zero and variance σ_z^2 such that $E(Z_t\eta_{1t}) = \alpha \neq 0$ but $E(Z_t\eta_{jt}) = 0$ for $j = 2, ...n$. This variable Z_t is called an *external instrument* (Stock and Watson, 2012; Mertens and Ravn, 2013). In this case, the approach to inference is straightforward. The VAR is estimated by OLS, giving $A(L)$. And then the residuals from the VAR are regressed onto Z_t, which identifies the first column of R, R_1, up to scale. Hence the impulse responses to a shock in η_{1t} are obtained—the effects of other structural shocks are not identified.

This approach can be recast as a GMM estimation problem with the moment conditions:

$$E(\tilde{Y}_{t-1} \otimes \varepsilon_t) = 0 \tag{2.2}$$

$$E(Z_t v_t) = 0 \tag{2.3}$$

where $\tilde{Y}_t = (Y'_{t-1}, Y'_{t-2}...Y'_{t-p})'$, $v_t = \varepsilon_t - \theta Z_t$ and $\theta = \alpha R_1$. The parameters to be estimated are $a = vec([A_1 A_2 ... A_p])$ and θ. Although estimation of the structural VAR is not typically written as this GMM estimation problem, it is numerically equivalent to the conventional least squares estimation. Letting $\hat{a}$ and $\hat{\theta}$ denote the estimators of these parameters from the moment conditions in equations (2.2) and (2.3), it is well known that:

$$T^{1/2} \begin{pmatrix} \hat{a} - a \\ \hat{\theta} - \theta \end{pmatrix} \rightarrow_d N(0, \begin{pmatrix} M^{-1} \otimes \Sigma & 0 \\ 0 & \sigma_z^{-2}\Sigma_v \end{pmatrix}) \tag{2.4}$$

where $\Sigma_v = Var(v_t)$. In this paper, I propose instead strengthening the assumptions of the VAR to include the requirement that the elements of η_t are mutually independent, not just uncorrelated, and then instead using the following moment conditions:

$$E(Y_{t-j} \otimes a(\frac{R(\varepsilon_t)}{T+1})) = 0 \tag{2.5}$$

$$E(Z_t a(\frac{R(v_t)}{T+1})) = 0 \tag{2.6}$$

where $R(.)$ denotes the element-by-element rank operator and $a(.)$ is a non-decreasing score function defined on the unit interval, such that $\int_0^1 a(u)du = 0$ and $\int_0^1 a(u)^2 du = 1$. Two examples are the Wilcoxon scores $a(u) = \sqrt{12}(u - 0.5)$ and the normal or van der Waerden scores $a(u) = \Phi^{-1}(u)$ where $\Phi(.)$ denotes the standard normal cdf. Define $F_{\varepsilon i}(.)$, $F_{vi}(.)$ and $F_{\eta_i}(.)$ as the cdfs of ε_{it}, v_{it} and η_{it}, and let $f_{\varepsilon i}(.)$, $f_{vi}(.)$ and $f_{\eta_i}(.)$ be the corresponding pdfs. Letting $\tilde{a}$ and $\tilde{\theta}$ denote the estimators of the parameters from the moment conditions in equations (2.5) and (2.6), under the assumption that the structural shocks are mutually independent:

$$T^{1/2}\begin{pmatrix} \tilde{a} - a \\ \tilde{\theta} - \theta \end{pmatrix} \to_d N(0, \begin{pmatrix} M^{-1} \otimes \bar{\Sigma} & 0 \\ 0 & \sigma_z^{-2}\bar{\Sigma}_v \end{pmatrix}) \tag{2.7}$$

where the ijth element of $\bar{\Sigma}$ is:

$$\frac{E(a(F_{\varepsilon i}(\varepsilon_{it}))a(F_{\varepsilon j}(\varepsilon_{jt})))}{\int a'(F_{\varepsilon i}(x))f_{\varepsilon i}(x)^2 dx \int a'(F_{\varepsilon j}(x))f_{\varepsilon j}(x)^2 dx} \tag{2.8}$$

and the ijth element of $\bar{\Sigma}_v$ is:

$$\frac{E(a(F_{vi}(v_{it}))a(F_{vj}(v_{jt})))}{\int a'(F_{vi}(x))f_{vi}(x)^2 dx \int a'(F_{vj}(x))f_{vj}(x)^2 dx} \tag{2.9}$$

Note that $E(a(F_{\varepsilon i}(\varepsilon_{it}))^2) = 1$. This means that for each equation in the VAR, the asymptotic distribution of the rank estimates of the slope coefficients has the usual form (Jaeckel, 1972; Hettmansperger and McKean, 1988) but the expressions derived here also reflect the dependence between the errors from the different equations in the VAR.

The expressions in equation (2.8) simplifies for normal or Wilcoxon scores. For Wilcoxon scores, the ijth element of $\bar{\Sigma}$ is:

$$\frac{E((F_{\varepsilon i}(\varepsilon_{it})-0.5)(F_{\varepsilon j}(\varepsilon_{jt})-0.5))}{\int f_{\varepsilon i}(x)^2dx\int f_{\varepsilon j}(x)^2dx} \tag{2.10}$$

For normal scores, the ijth element of $\bar{\Sigma}$ is:

$$\frac{E(\Phi^{-1}(F_{\varepsilon i}(\varepsilon_{it}))\Phi^{-1}(F_{\varepsilon j}(\varepsilon_{jt})))}{\int \frac{f_{\varepsilon i}(x)^2}{\phi(\Phi^{-1}(F_{\varepsilon i}(x)))}dx\int \frac{f_{\varepsilon j}(x)^2}{\phi(\Phi^{-1}(F_{\varepsilon j}(x)))}dx} \tag{2.11}$$

where $\phi(.)$ denotes the standard normal pdf. Equation (2.9) can be simplified similarly for normal or Wilcoxon scores.

For normal scores, if the shocks are all Gaussian with mean zero such that $Var(\varepsilon_{it})=\sigma_i^2$, then $\Phi^{-1}(F_{\varepsilon i}(\varepsilon_{it}))=\frac{\varepsilon_{it}}{\sigma_i}$ and $E(\Phi^{-1}(F_{\varepsilon i}(\varepsilon_{it}))\Phi^{-1}(F_{\varepsilon j}(\varepsilon_{jt})))$ is just the correlation between ε_{it} and ε_{jt}. Meanwhile, $f_{\varepsilon i}(x)=\phi(\frac{x}{\sigma_i})\frac{1}{\sigma_i}$ and so $\int\frac{f_{\varepsilon i}(x)^2}{\phi(\Phi^{-1}(F_{\varepsilon i}(x)))}dx=\frac{1}{\sigma_i}$. Hence, $\bar{\Sigma}=\Sigma$ and similarly $\bar{\Sigma}_v=\Sigma_v$. Thus, with Gaussian shocks, the normal scores rank estimator is asymptotically equivalent to least squares, which is a familiar result in the simple linear regression model.

2.2 Cholesky Identification

An alternative, and much older, approach to identification, is to posit that the matrix R is lower triangular. This restriction is sufficient to solve uniquely for R from the variance-covariance matrix of the reduced form shocks, and consequently to identify all of the structural shocks.

The Cholesky identification can be implemented as a just identified GMM estimator with the moment conditions:

$$E(\tilde{Y}_{t-1}\otimes\varepsilon_t)=0 \tag{2.12}$$

$$E(vech(\eta_t\eta_t')-I)=0 \tag{2.13}$$

The parameters to be estimated are $a = vec([A_1 A_2 ... A_p])$ and $r = vech(R)$. Although estimation of the structural VAR is not typically written as the GMM estimation problem in equations (2.12) and (2.13), this is numerically equivalent to the conventional implementation. Letting $\hat{a}$ and $\hat{r}$ denote the estimators of these parameters from the moment conditions in equations (2.12) and (2.13), it is well known that:

$$T^{1/2} \begin{pmatrix} \hat{a} - a \\ \hat{r} - r \end{pmatrix} \rightarrow_d N(0, \begin{pmatrix} M^{-1} \otimes \Sigma & 0 \\ 0 & J^{-1} E J^{-1\prime} \end{pmatrix}) \tag{2.14}$$

where $J = -L(I_{n^2} + K)(R' \otimes I_n)L'L(R^{-1\prime} \otimes R^{-1})L'$, $E = Var(vech(\eta_t \eta_t') - I)$ and L and K are the elimination and commutation matrices, respectively (Lütkepohl, 1990).

Alternatively, in this paper I propose adding the requirement that the elements of η_t are mutually independent, not just uncorrelated. I define the nxn matrix $C = [c_{ij}]$ where

$$c_{ii} = \eta_{it}^2 - 1$$
$$c_{ij} = \eta_{it} a(\frac{R(\eta_{jt})}{T+1}), \ i \neq j$$

and then propose instead using the following moment conditions:

$$E(Y_{t-j} \otimes a(\frac{R(\varepsilon_t)}{T+1})) = 0, j = 1, ...p \tag{2.15}$$
$$E(vech(C)) = 0 \tag{2.16}$$

Letting $\tilde{a}$ and $\tilde{r}$ denote the estimators of the parameters from the moment conditions in equations (2.15) and (2.16),under the assumption that the structural shocks are mutually independent:

$$T^{1/2} \begin{pmatrix} \tilde{a} - a \\ \tilde{r} - r \end{pmatrix} \rightarrow_d N(0, \begin{pmatrix} M^{-1} \otimes \bar{\Sigma} & 0 \\ 0 & J^{-1} \bar{E} J^{-1\prime} \end{pmatrix}) \tag{2.17}$$

where $\bar{E} = diag(vech(W))$ and W is an nxn matrix $[w_{ij}]$ such that:

$$w_{ii} = E(\eta_{it}^4)$$
$$w_{ij} = 1/[\int a'(F_{\eta_j}(x))f_{\eta_j}(x)^2 dx]^2, \; i \neq j$$

2.3 Numerical Calculations

I did numerical computations taking a VAR calibrated to the dataset of Gertler and Karadi (2015), described later. The VAR was $Y_t = AY_{t-1} + R\eta_t$ with four variables and:

$$A = \begin{pmatrix} 0.990 & 0.002 & -0.044 & -0.488 \\ 0.009 & 0.990 & 0.012 & -0.075 \\ 0.009 & -0.015 & -0.935 & -0.104 \\ 0.000 & 0.001 & 0.010 & 0.852 \end{pmatrix}$$
$$R = \begin{pmatrix} 0.658 & 0 & 0 & 0 \\ 0.071 & 0.388 & 0 & 0 \\ 0.128 & -0.001 & 0.462 & 0 \\ -0.035 & -0.005 & -0.008 & 0.282 \end{pmatrix} \tag{2.18}$$

The structural shocks are all t-distributed on v degrees of freedom, rescaled to have unit variance. I considered impulse responses to the third structural shock, η_{3t}, which corresponds to the monetary policy shock in the model of Gertler and Karadi. The Cholesky identification is clearly applicable to this VAR. I also assumed that an external instrument exists with a correlation of α with η_{3t}, and with the same distribution as η_{3t}.

I used the asymptotic distributions in equations (2.4), (2.7), (2.14) and (2.17) along with the delta method to work out the ratio of the asymptotic standard errors of impulse response coefficient estimates using rank based methods to their counterparts using conventional inference. These are shown for $v = 5$ and $\alpha = 0.5$ in Figure 1 (normal scores, external instruments), Figure 2 (normal scores, Cholesky identification), Figure 3 (Wilcoxon scores, external instruments) and Figure 4 (Wilcoxon scores, Cholesky identification). Results are shown for impulse responses of all four variables at lags up to 60. In all four Figures, we see that the rank based methods are somewhat more efficient than conventional least squares inference. Asymptotic standard errors are reduced by around 10 percent.

3 Simulations

In this section, I report some Monte-Carlo evidence on the finite-sample properties of rank based estimates of structural VARs. The design is a VAR(1) as in equation (2.18) where the structural shocks are all t-distributed on v degrees of freedom, rescaled to have unit variance. Impulse responses to the third structural shock are considered using an external instrument that has a correlation of α with η_{3t}, and with the same distribution as η_{3t}.

I compute the simulated root mean square errors of the rank-based estimates of impulse responses, relative to their least squares counterparts. I consider (i) sample sizes $T = 200$ and $T = 1000$, (ii) normal and Wilcoxon scores, (iii) t-distributed shocks on 5 degrees of freedom and normal shocks ($v = 5$ or ∞). The results are shown in Figures 5-12 for 8 different permutations of these settings. Asymptotic relative root mean square errors, computed as discussed in the previous section, are also included.

For normal scores, the asymptotic relative root mean square errors (rank relative to least squares) are around 1 for Gaussian shocks and around 0.9 for $t(5)$ shocks. The finite sample relative root mean square errors can be a bit different with the sample size $T = 200$. In most cases, the rank method does better than predicted by asymptotic theory. By the time the sample size is increased to $T = 1000$, the finite sample and asymptotic relative root mean square errors are quite close. In simple regression contexts, the normal scores have equal asymptotic efficiency as the Gaussian PMLE if the errors are in fact Gaussian, while the normal scores are otherwise asymptotically more efficient. The simulation results here are consistent with this idea.

For Wilcoxon scores, the asymptotic relative root mean square errors are a bit over 1 for Gaussian shocks and around 0.9 for $t\,(5)$ shocks. Again, the finite sample relative root mean square errors are somewhat different with a sample size of $T = 200$, but the finite sample and asymptotic relative root mean square errors are quite close once the sample size is up to $T = 1000$.

3.1 Confidence Intervals

I also evaluated the effective coverage and average width of confidence intervals for impulse responses in this simulation, where confidence intervals are formed using the bias-adjusted bootstrap of Kilian (1998) with 95% nominal coverage in all cases. When resampling from the reduced form residuals, I also draw the matched value of the instrument. I report results just using Wilcoxon scores where the shocks are $t(5)$ and the sample size is $T = 200$. The rank-based estimator is bias-adjusted in exactly the same way as the least squares estimator. Effective coverage results are shown in Figure 13 and average width results are shown in Figure 14. It can be seen that the effective coverage rates of the two confidence intervals are very close to each other, and are fairly close to the 95% nominal coverage. However, the confidence intervals using the rank-based estimator are a bit shorter, consistent with the efficiency gains from rank-based estimation.

3.2 Common Stochastic Volatility

It seems reasonable to impose independence of the different structural shocks in a VAR, as they are meant to reflect fundamentally different sources of risk. However, a natural way in which independence might fail is common stochastic volatility. Each structural shock might be conditionally Gaussian and independent, but where all the shocks have variance ω_t^2 where $\log(\omega_t^2)$ follows a Gaussian autoregression with coefficient 0.95. This common stochastic volatility violates the assumptions required for the rank-based estimators to have their asymptotic distributions in equations (2.7) and (2.17). Still, conditional volatility causes unconditional fat tails and so it might not be too surprising if the rank based estimators actually worked well in finite samples.

Simulated relative root mean square errors (rank relative to least squares) with this conditional stochastic volatility are reported in Figures 15-18, for sample sizes $T = 200$ and $T = 1000$ and normal and Wilcoxon scores. Use of the rank based estimators can give quite large improvements in estimator accuracy. Note that in these figures I do not report asymptotic relative root mean square errors, because I have no asymptotic distribution theory for the rank based estimators with stochastic volatility.

4 Empirical Application

My empirical application is to the VAR of Gertler and Karadi (2015). This consists of four variables: CPI, industrial production, interest rates (the one-year Treasury yield), and the excess bond premium of Gilchrist and Zakrajšek (2012). The one-year yield is taken as the measure of the stance of monetary policy, to incorporate the period when the federal funds rate was stuck at the zero lower bound and more generally to reflect the forward guidance that the Federal Reserve has given about monetary policy over the last two decades.

The data sources are as in Gertler and Karadi (2015) and the sample period is 1980:07 to 2012:06. Two approaches are considered to identification. The first is the Cholesky identification in which the data are ordered as CPI, IP, interest rates, excess bond premium. The second is external instruments in which the sum of daily changes in federal funds futures rates on days of FOMC announcements is used as the instrument for the monetary policy shock. This uses the fourth federal funds futures contract, which is a bet on the average level of interest rates three months hence. The data on this instrument are only available back to 1991:01. This instrument is assumed to be correlated with the monetary policy shock, but not with any other structural shocks.

I form standard errors from a residual bootstrap with the bias-adjustment of Kilian (1998), but not the wild bootstrap used in Mertens and Ravn (2013) and Gertler and Karadi (2015). This wild bootstrap multiplies the reduced form errors and instruments at any time period by ± 1, each with equal probability. But that means that the projection of the reduced form errors onto the instruments is identical in every bootstrap replication, and so does not reflect any uncertainty about the relationship between reduced form errors and instruments. This point is also made by Jentsch and Lunsford (2016). Instead I use the ordinary bootstrap with bias-adjustment, except that when resampling from the reduced form residuals, I also draw the matched value of the instrument, as discussed in the previous section.

Figure 15 shows the impulse response estimates and confidence intervals for a monetary policy shock using conventional least squares estimation. These results are similar to those reported by Gertler and Karadi (2015), except that the confidence intervals are somewhat wider, owing to the use of a different bootstrap, as discussed above. The impulse response estimates are qualitatively similar using rank-based estimation, but the confidence intervals are a little tighter. To represent this, Figure 16 plots the ratio of the confidence interval width for the rank-based estimate with normal scores to the counterpart using least squares estimation in this application. Likewise, Figure 17 plots the ratio of the confidence interval width for the rank-based estimate with Wilcoxon scores to the counterpart using least squares estimation. In both cases, we see that the rank-based method shrinks the width of confidence interval for almost all variables and horizons, and in many cases by as much as 20 percent. The efficiency gains are a bit bigger with Wilcoxon scores than with normal scores, which is a typical finding in the literature on rank estimation, when the shocks have quite fat tails. Applying the tail index estimator of Hill (1975) to the reduced form VAR residuals, gives estimates between 2 and 4, indicating that the shocks do indeed have quite fat tails, and consistent with the apparent efficiency gains from nonparametric rank estimation.

5 Conclusion

In this paper, I have proposed the use of nonparametric rank based estimation of structural vector autoregressions. Identification can be by a Cholesky ordering, or by the method of external instruments. I have found both asymptotically, and in simulations, that in cases where the shocks are non-Gaussian, the proposed estimator brings efficiency gains relative to the conventional estimation method. The rank based methods should be applicable to other schemes for identifying a structural VAR, but I leave this for future work.

The efficiency gains from rank based estimation are modest, but they are easily available, and structural VARs are very widely used. Researchers who are comfortable with an assumption that the structural shocks are independent, and not just uncorrelated, should use these rank based estimates.

Appendix: Heuristic Derivation of Equation (2.7)

In this appendix, I give a heuristic derivation of equation (2.7). The argument is similar to the derivation of the asymptotic distribution of rank-based estimators in a regression model with multiple dependent variables of Davis and McKean (1993). From the Central Limit Theorem, and since ε_t is independent of $\tilde{Y}_{t-1}$:

$$T^{-1/2}\Sigma(\tilde{Y}_{t-1} \otimes a(\frac{R(\varepsilon_t)}{T+1})) \to_d N(0, M \otimes J_1)$$

where the ijth element of J_1 is $E(a(F_{\varepsilon i}(\varepsilon_{it}))a(F_{\varepsilon j}(\varepsilon_{jt})))$.

Meanwhile,

$$T^{-1}\frac{d}{da}\Sigma(\tilde{Y}_{t-1}\otimes a(\frac{R(\varepsilon_t)}{T+1}))\rightarrow_p M\otimes J_2$$

where J_2 is a diagonal matrix and the iith element is $E(a'(F_{\varepsilon i}(\varepsilon_{it}))f_{\varepsilon i}(\varepsilon_{it}))=\int a'(F_{\varepsilon i}(x))f_{\varepsilon i}(x)^2dx$. Hence, from the usual asymptotic distribution of a just-identified GMM estimator, $T^{1/2}(\tilde{a}-a)\rightarrow_d N(0,M^{-1}\otimes\bar{\Sigma})$ where $\bar{\Sigma}=J_2^{-1}J_1J_2^{-1}$. Similarly, $T^{1/2}(\tilde{\theta}-\theta)\rightarrow_d N(0,\sigma_z^{-2}\bar{\Sigma}_v)$. Lastly, $T^{-1/2}\Sigma(\tilde{Y}_{t-1}\otimes a(\frac{R(\varepsilon_t)}{T+1}))$ and $T^{-1/2}\Sigma Z_t a(\frac{R(v_t)}{T+1})$ are asymptotically independent because Z_t, v_t and ε_t are all independent of $\tilde{Y}_{t-1}$. The derivation of equation (2.17) is similar.

References

Andrews, D. W. and V. Marmer (2008): "Exactly distribution-free inference in instrumental variables regression with possibly weak instruments," *Journal of Econometrics*, 142, 183–200.

Chernoff, H. and I. R. Savage (1958): "Asymptotic normality and efficiency of certain nonparametric test statistics," *Annals of Mathematical Statistics*, 29, 972–994.

Davis, J. B. and J. W. McKean (1993): "Rank-based methods for multivariate linear models," *Journal of the American Statistical Association*, 88, 245–251.

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

Gilchrist, S. and E. Zakrajšek (2012): "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102, 1692–1720.

Hájek, J. and Z. Šidák (1967): *Theory of rank tests*, Academic Press.

Hallin, M. and D. Paindaveine (2004): "Multivariate signed-rank tests in vector autoregressive order identification," 19, 697–711.

Hasan, M. N. and R. Koenker (1997): "Robust rank tests of the unit root hypothesis," *Econometrica*, 65, 133–161.

Hettmansperger, T. P. (1984): *Statistical Inference Based on Ranks*, Wiley.

Hettmansperger, T. P. and J. W. McKean (1988): *Robust nonparametric statistical methods.*

Hill, B. M. (1975): "A simple general approach to inference about the tail of a distribution," *Annals of Statistics*, 3, 1163–1174.

Jaeckel, L. A. (1972): "Estimating regression coefficients by minimizing the dispersion of the residuals," *Annals of Mathematical Statistics*, 43, 14491458.

Jentsch, C. and K. G. Lunsford (2016): "Proxy SVARs: Asymptotic Theory, Bootstrap Inference, and the Effects of Income Tax Changes in the United States," Working Paper.

Kilian, L. (1998): "Small-sample confidence intervals for impulse responses," *Review of Economics and Statistics*, 80, 218–230.

Kilian, L. and H. Lütkepohl (2017): *Structural Vector Autoregressive Analysis*, Cambridge University Press.

Lanne, M. and H. Lütkepohl (2010): "Structural vector autoregressions with nonnormal residuals," *Journal of Business and Economic Statistics*, 28, 159–168.

Lanne, M., M. Meitz, and P. Saikkonen (2017): "Identification and estimation of non-Gaussian structural vector autoregressions," *Journal of Econometrics*, 196, 288–304.

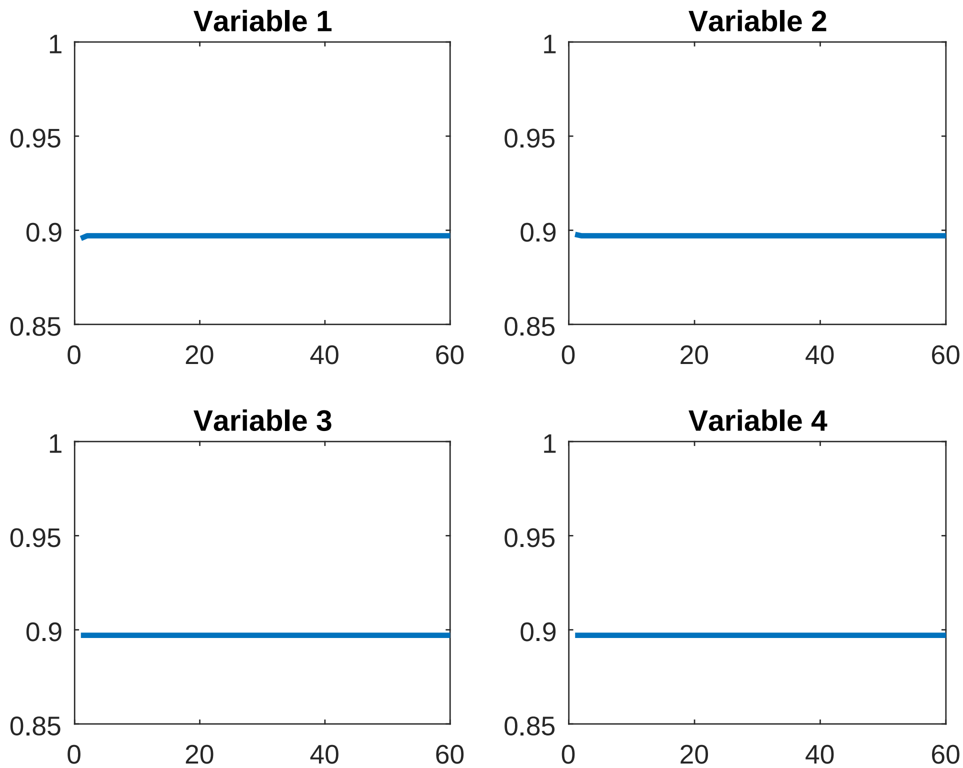
Lütkepohl, H. (1990): "Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models," *Review of Economics and Statistics*, 72, 116–125.

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

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

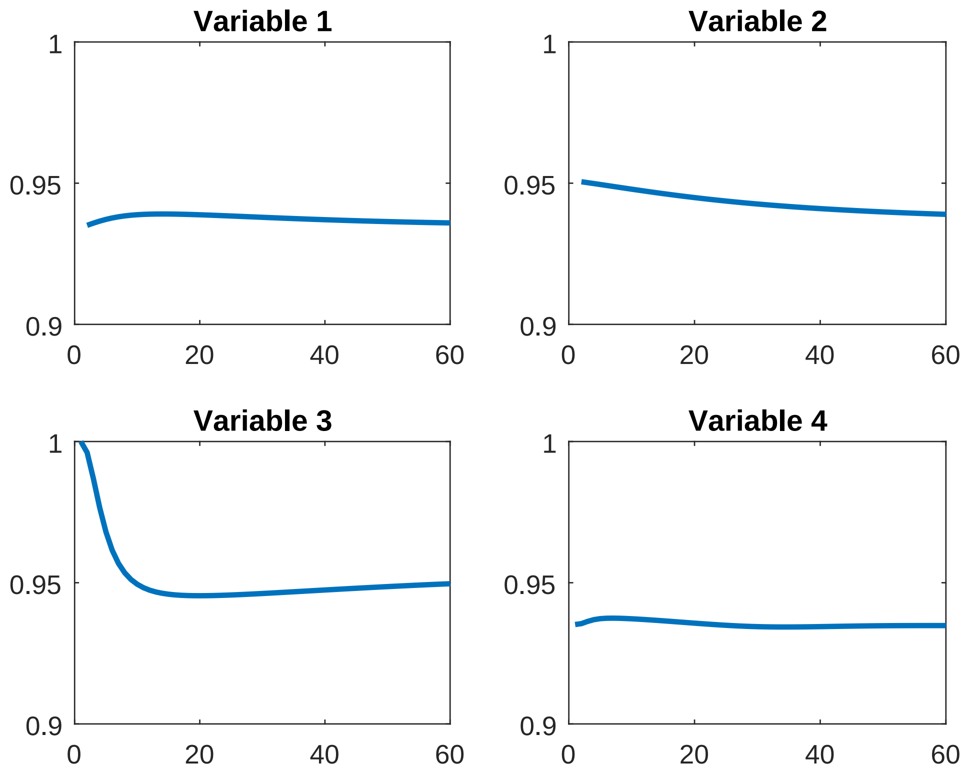
Stock, J. H. and M. W. Watson (2012): "Disentangling the channels of the 200709 recession," *Brookings Papers on Economic Activity*, 1, 81135.

Figure 1: Asymptotic Relative Standard Errors (Normal Scores/LS): External Instruments



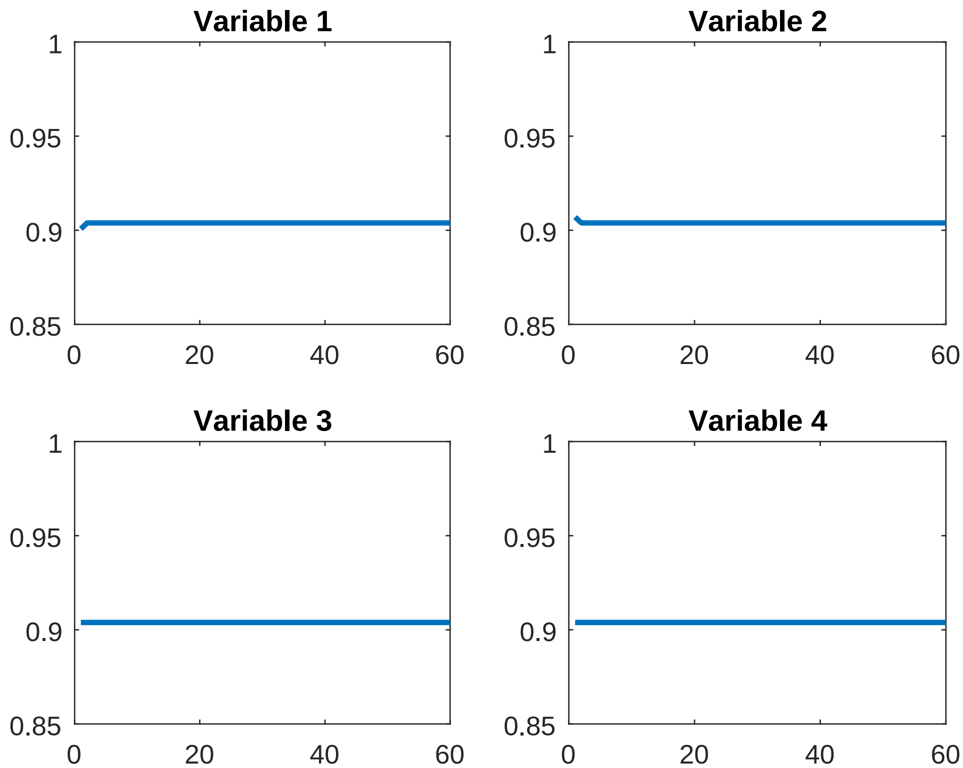
Notes: This figure shows the asymptotic standard errors for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the asymptotic standard errors using conventional least squares estimation. Identification is by external instruments and the shocks are t(5) distributed.

Figure 2: Asymptotic Relative Standard Errors (Normal Scores/LS): Cholesky



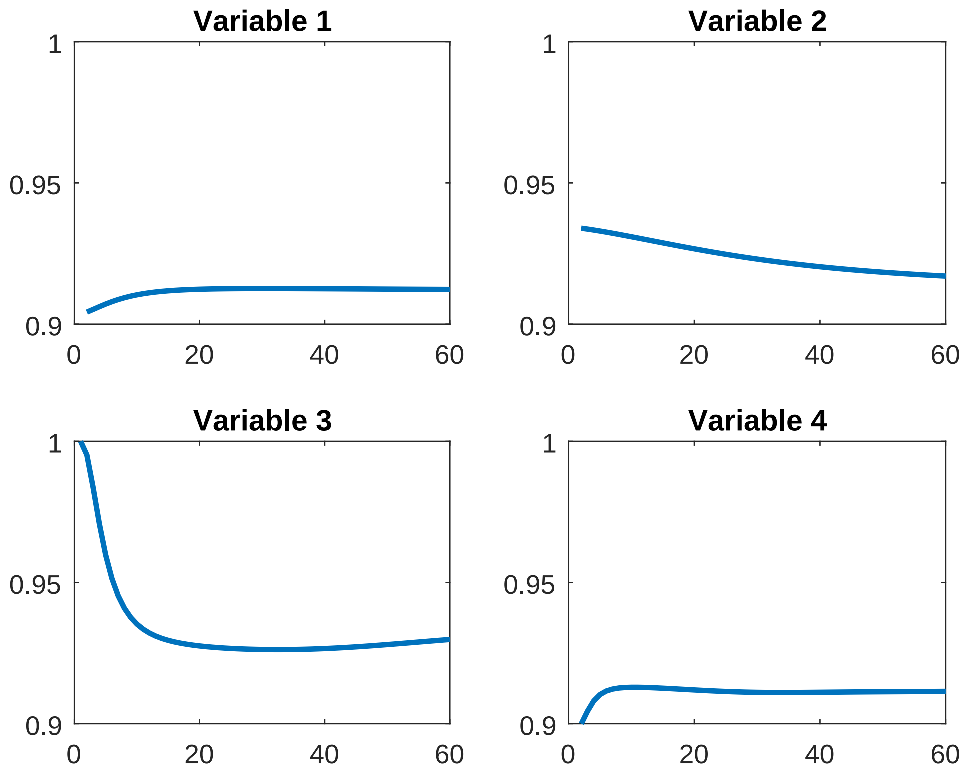
Notes: This figure shows the asymptotic standard errors for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the asymptotic standard errors using conventional least squares estimation. Identification is by Cholesky ordering and the shocks are t(5) distributed.

Figure 3: Asymptotic Relative Standard Errors (Wilcoxon/LS): External Instruments



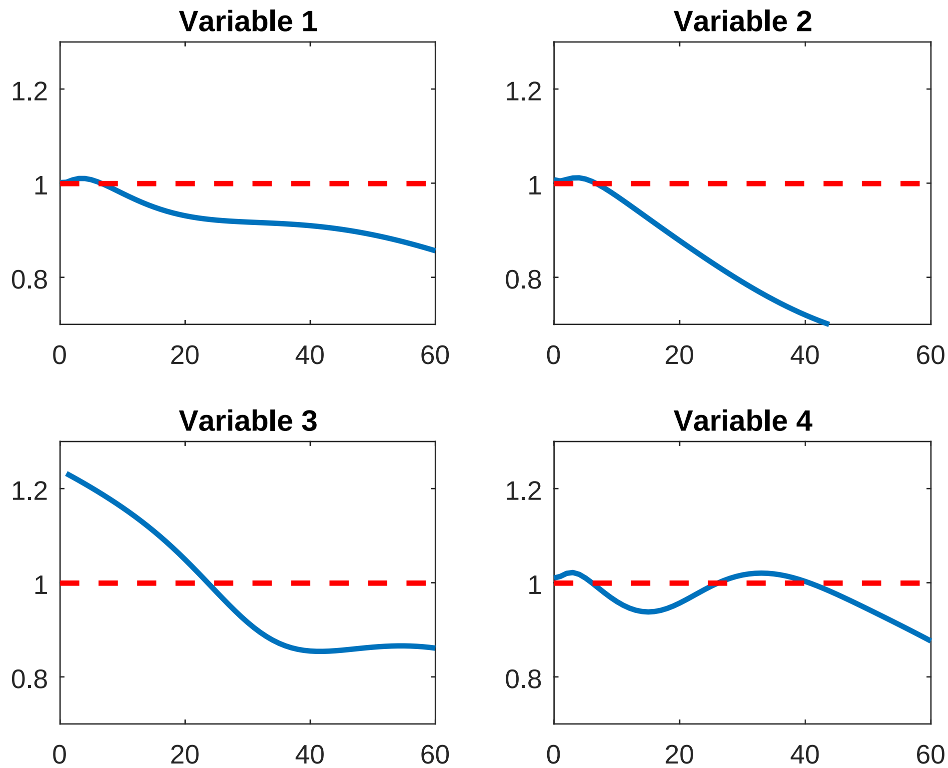
Notes: This figure shows the asymptotic standard errors for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the asymptotic standard errors using conventional least squares estimation. Identification is by external instruments and the shocks are t(5) distributed.

Figure 4: Asymptotic Relative Standard Errors (Wilcoxon/LS): Cholesky



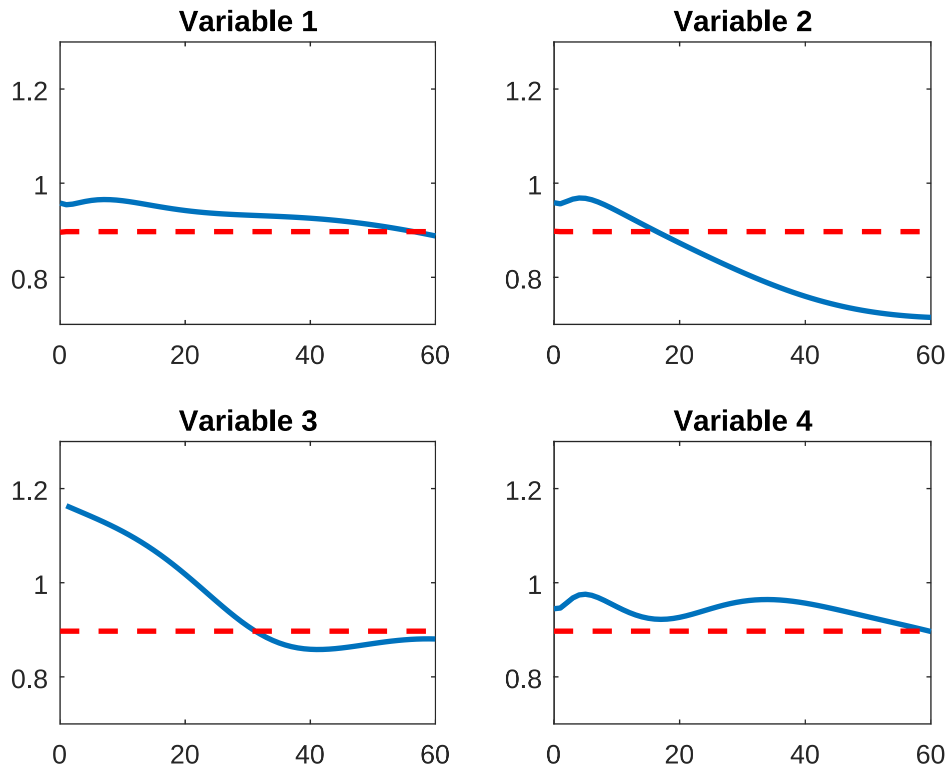
Notes: This figure shows the asymptotic standard errors for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the asymptotic standard errors using conventional least squares estimation. Identification is by Cholesky ordering and the shocks are t(5) distributed.

Figure 5: Simulated RRMSEs (Normal/LS): T=200, Gaussian shocks



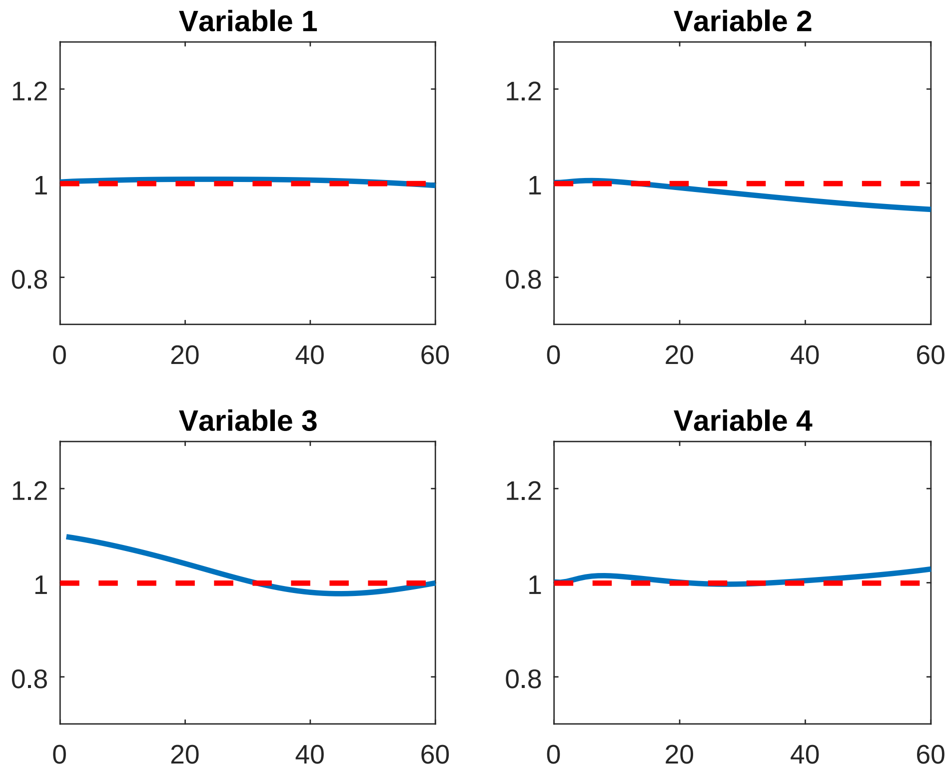
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 200 with Gaussian shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 6: Simulated RRMSEs (Normal/LS): T=200, t(5) shocks



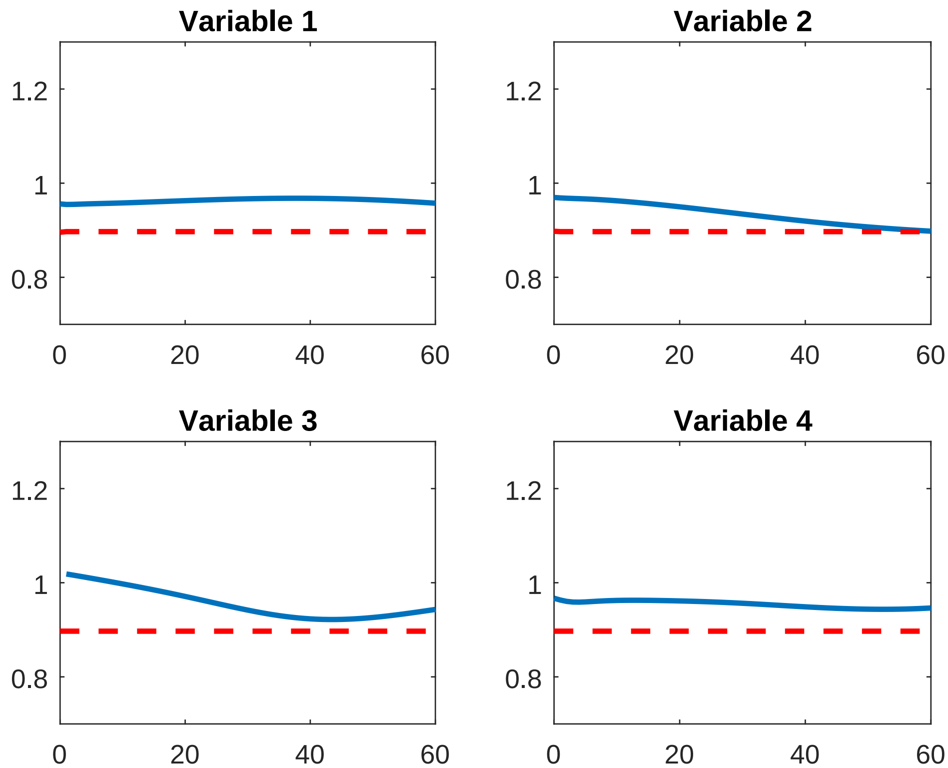
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 200 with t(5) shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 7: Simulated RRMSEs (Normal/LS): T=1,000, Gaussian shocks



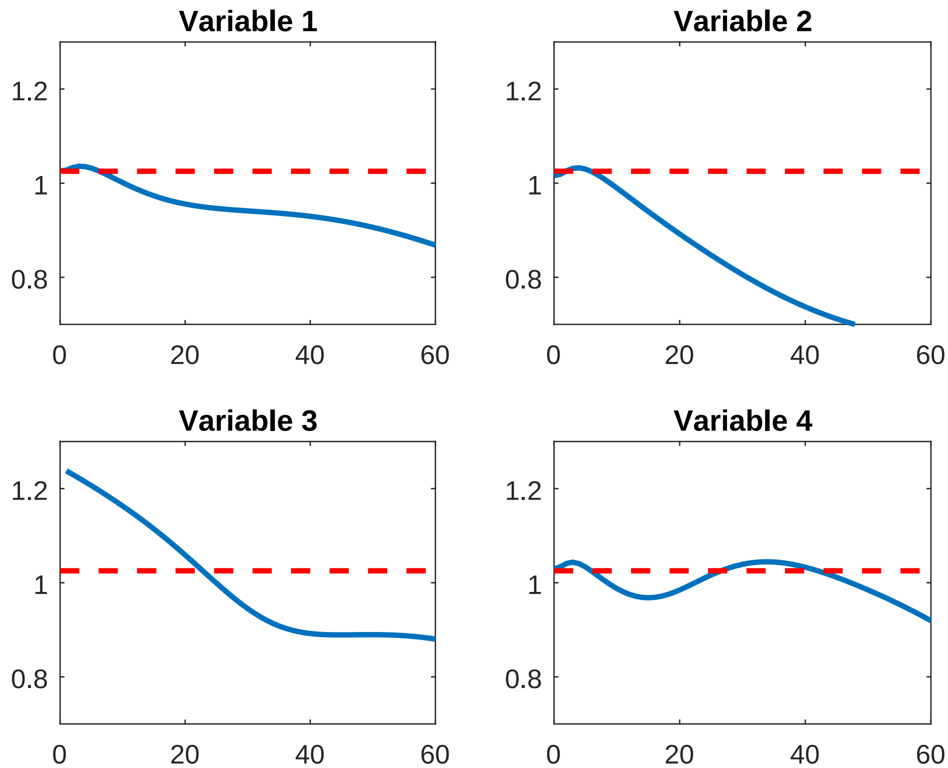
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 1,000 with Gaussian shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 8: Simulated RRMSEs (Normal/LS): T=1,000, t(5) shocks



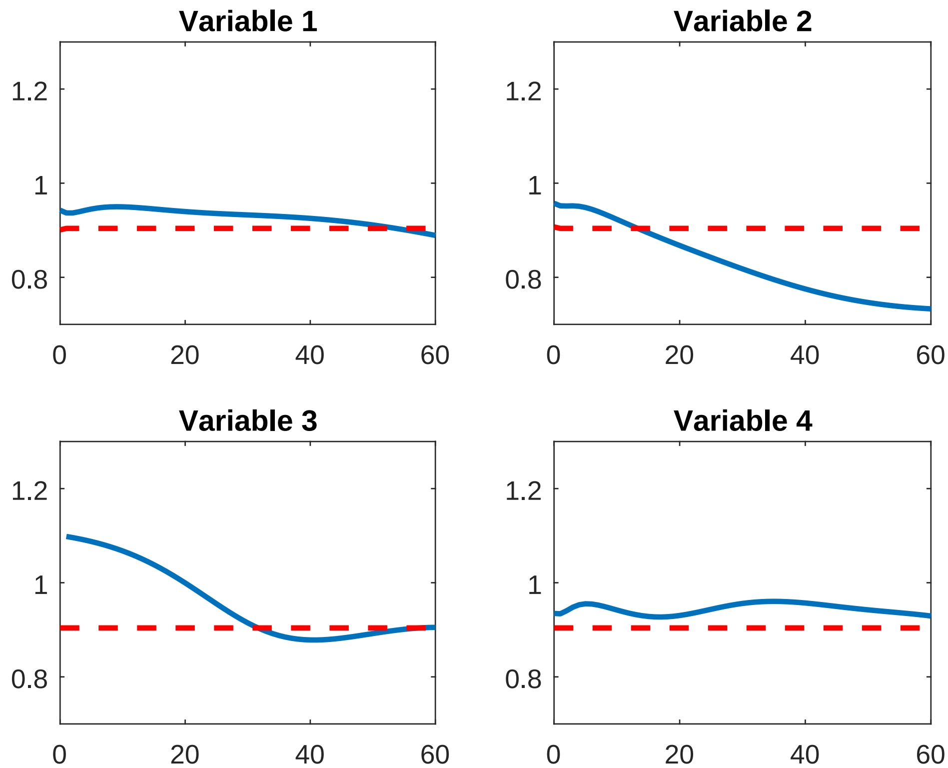
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 1,000 with t(5) shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 9: Simulated RRMSEs (Wilcoxon/LS): T=200, Gaussian shocks



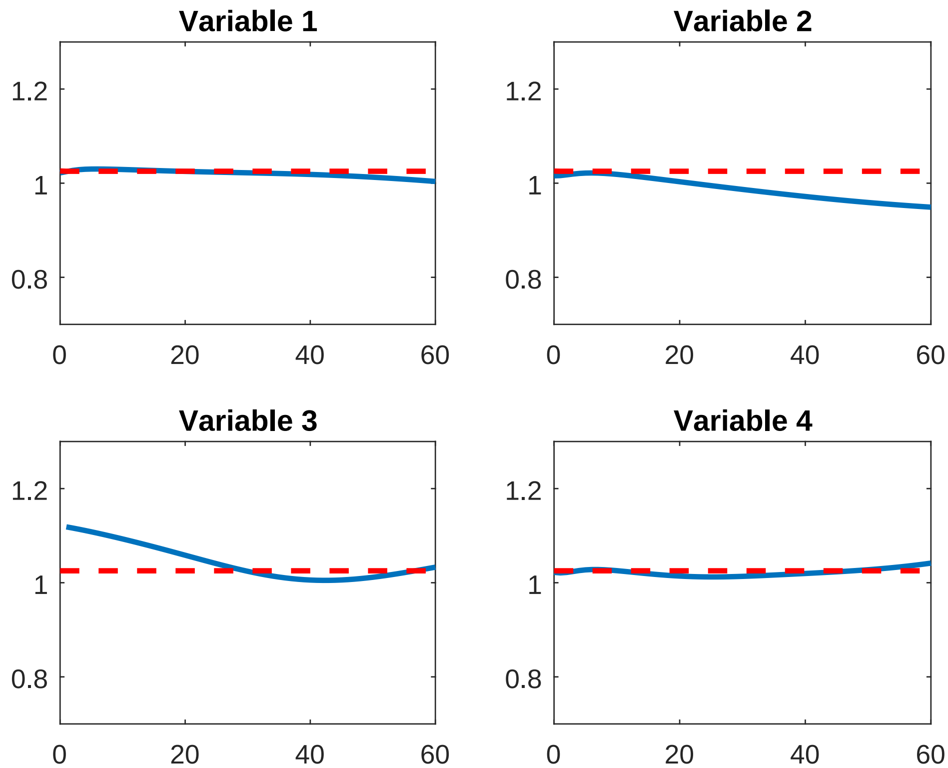
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 200 with Gaussian shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 10: Simulated RRMSEs (Wilcoxon/LS): T=200, t(5) shocks



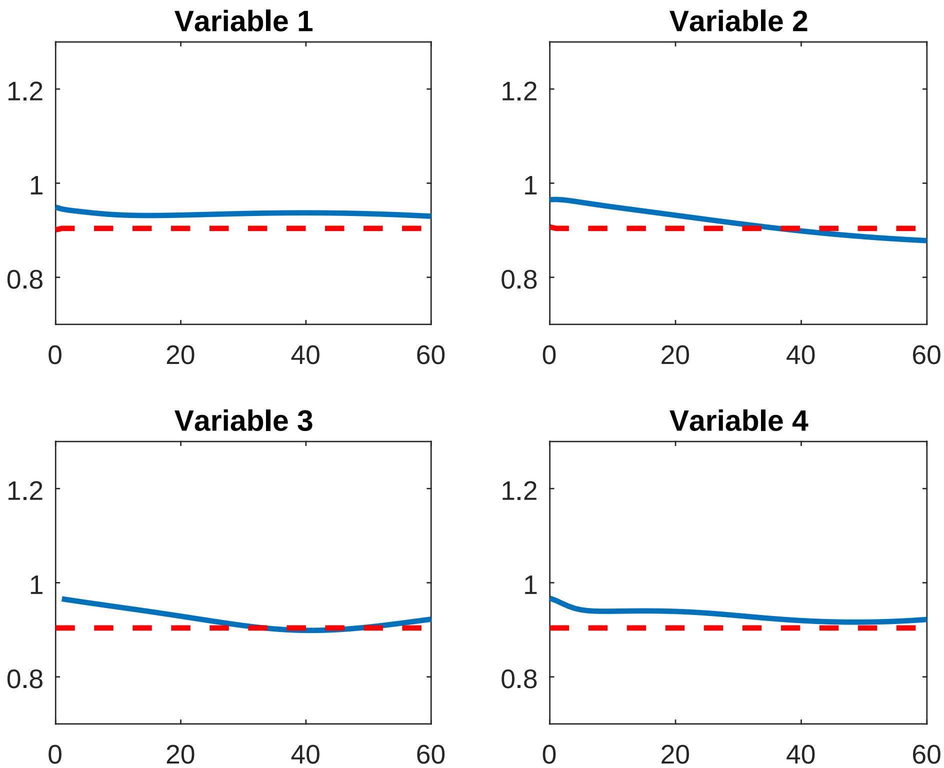
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 200 with t(5) shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 11: Simulated RRMSEs (Wilccoxon/LS): T=1,000, Gaussian shocks



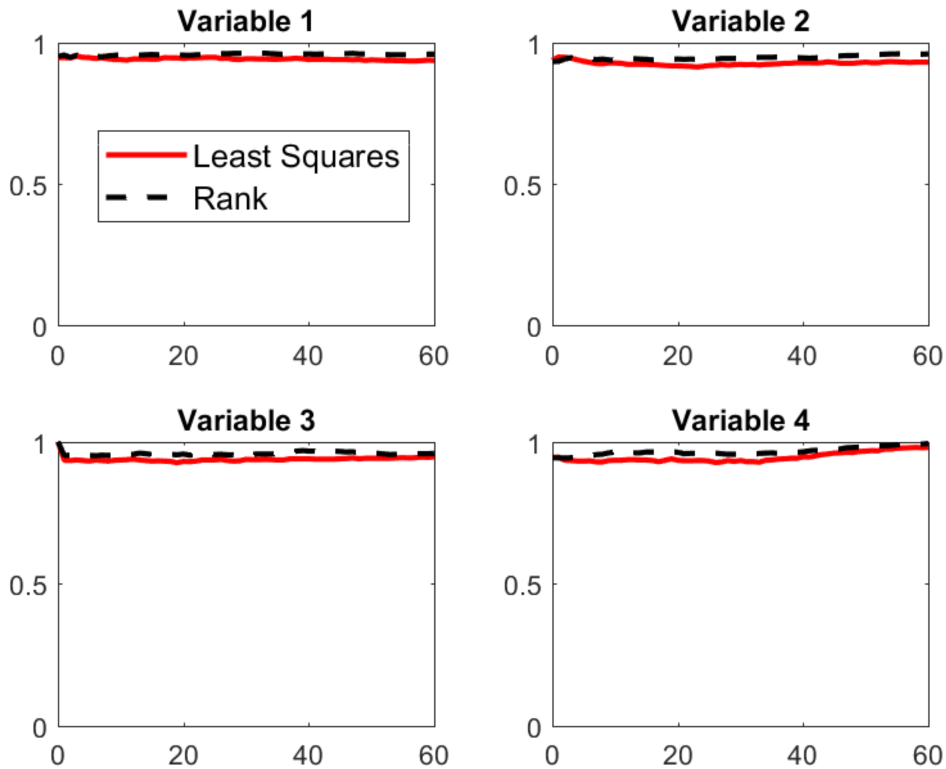
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 1,000 with Gaussian shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 12: Simulated RRMSEs (Wilcoxon/LS): T=1,000, t(5) shocks



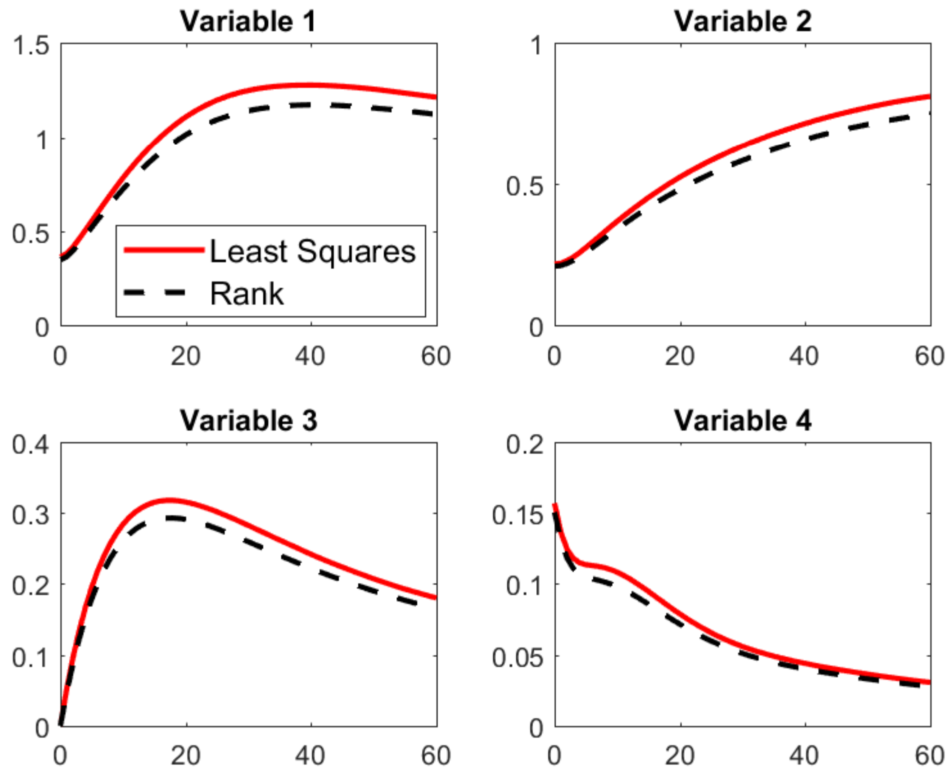
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 1,000 with t(5) shocks and identification by external instruments. The asymptotic relative root mean square errors are also reported for comparison (red dashed lines).

Figure 13: Effective Coverage of Confidence Intervals: T=1,000, t(5) shocks



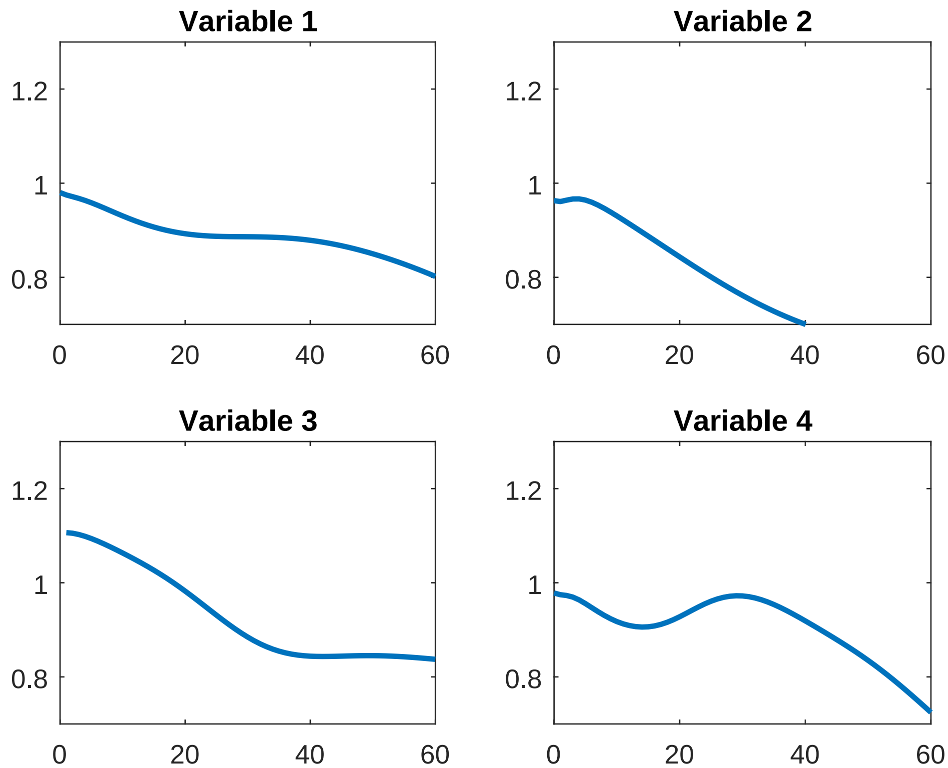
Notes: This figure shows the simulated effective coverage of confidence intervals (95% nominal level) for impulse responses in the VAR given by equation (2.18) using a bias-adjusted bootstrap. Results are shown both for least squares and rank-based estimates, with the latter using Wilcoxon scores. The simulation is based on a sample of size 1,000 with t(5) shocks and identification by external instruments.

Figure 14: **Average Width of Confidence Intervals: T=1,000, t(5) shocks**



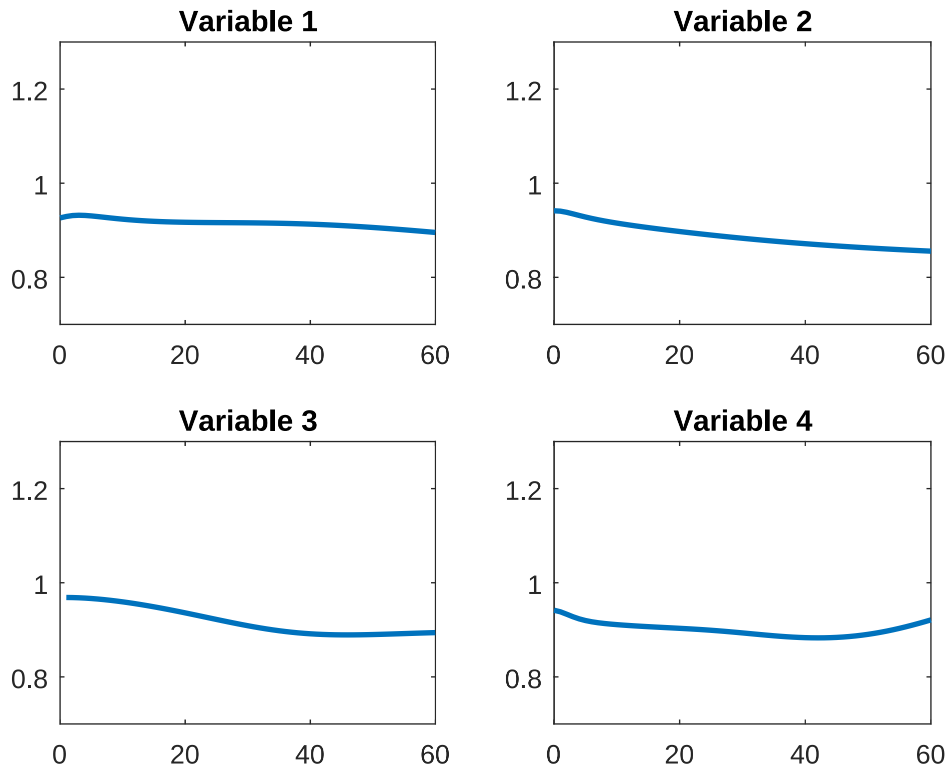
Notes: This figure shows the simulated average width of confidence intervals (95% nominal level) for impulse responses in the VAR given by equation (2.18) using a bias-adjusted bootstrap. Results are shown both for least squares and rank-based estimates, with the latter using Wilcoxon scores. The simulation is based on a sample of size 1,000 with t(5) shocks and identification by external instruments.

Figure 15: Simulated RRMSEs (Normal/LS): T=200, common stochastic volatility shocks



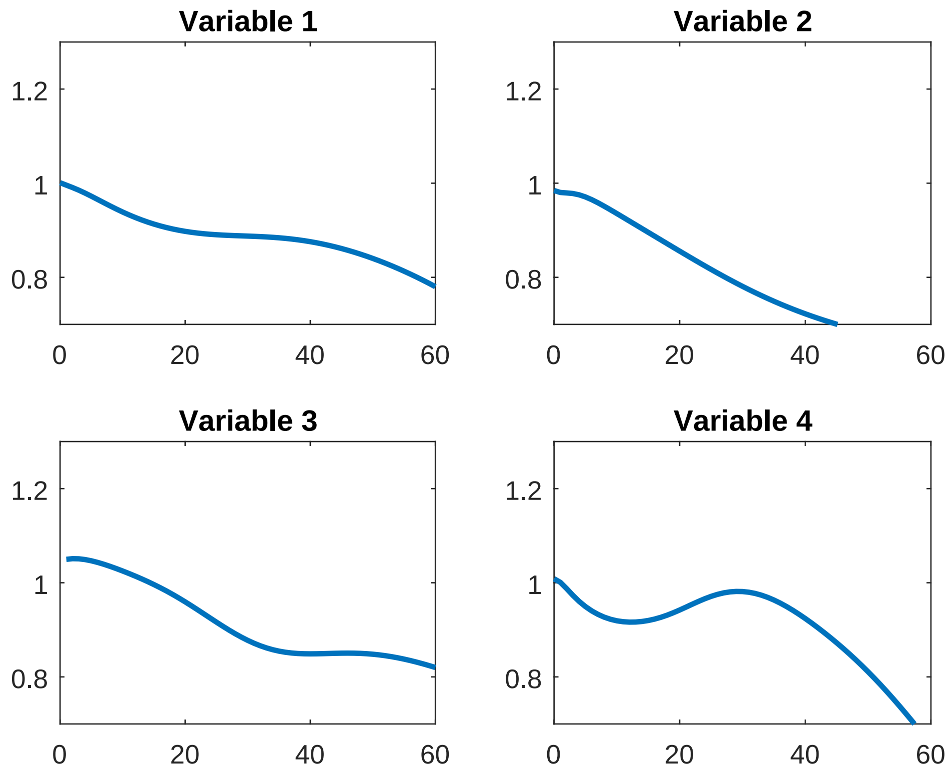
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 200 with shocks that are conditionally Gaussian and independent, but with common stochastic volatility.

Figure 16: Simulated RRMSEs (Normal/LS): T=1,000, common stochastic volatility shocks



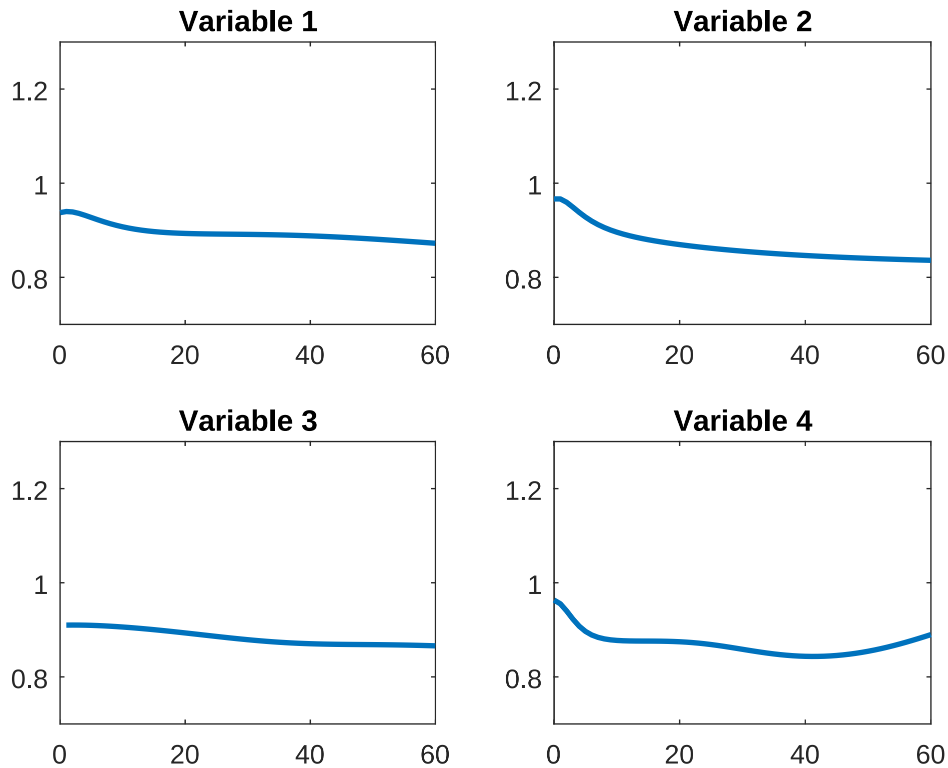
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using normal scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 1,000 with shocks that are conditionally Gaussian and independent, but with common stochastic volatility.

Figure 17: Simulated RRMSEs (Wilcoxon/LS): T=200, common stochastic volatility shocks



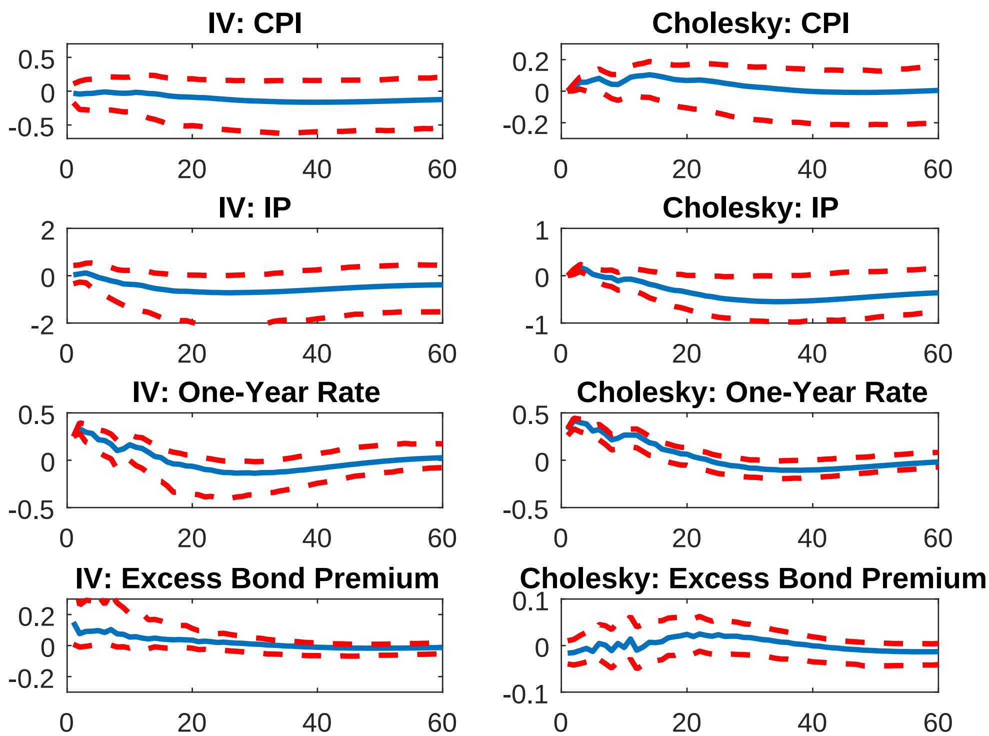
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 200 with shocks that are conditionally Gaussian and independent, but with common stochastic volatility.

Figure 18: Simulated RRMSEs (Wilcoxon/LS): T=200, common stochastic volatility shocks



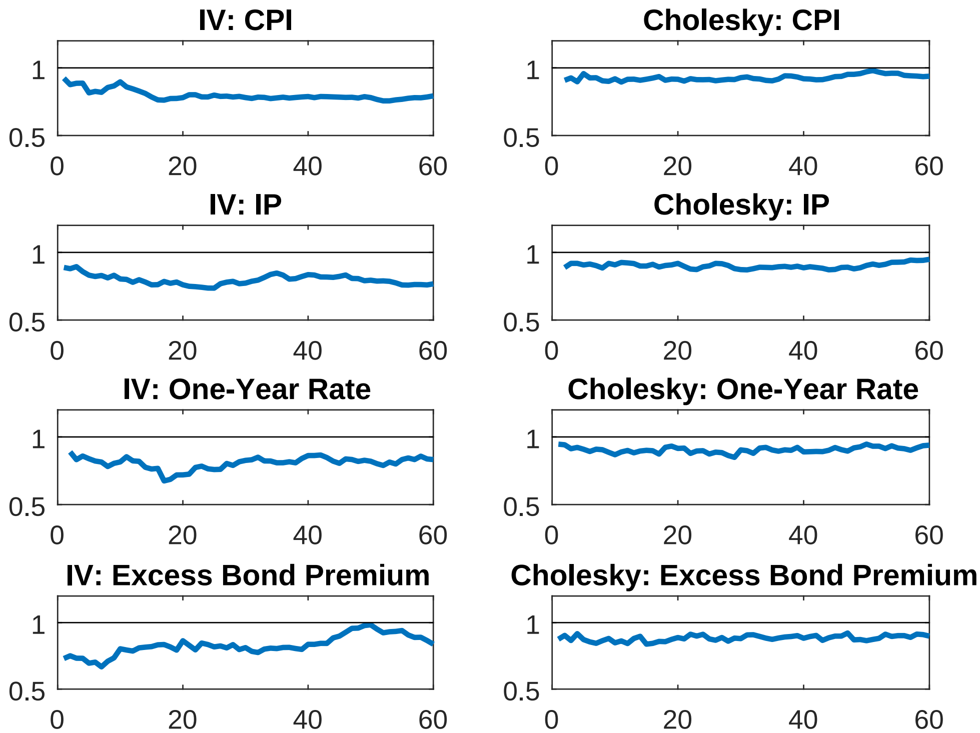
Notes: This figure shows the ratio of simulated root mean square error for impulse responses in the VAR given by equation (2.18) using Wilcoxon scores, divided by the counterparts using conventional least squares estimator. The simulation is based on a sample of size 1,000 with shocks that are conditionally Gaussian and independent, but with common stochastic volatility.

Figure 19: Estimated Impulse Responses to a Monetary Policy Shock



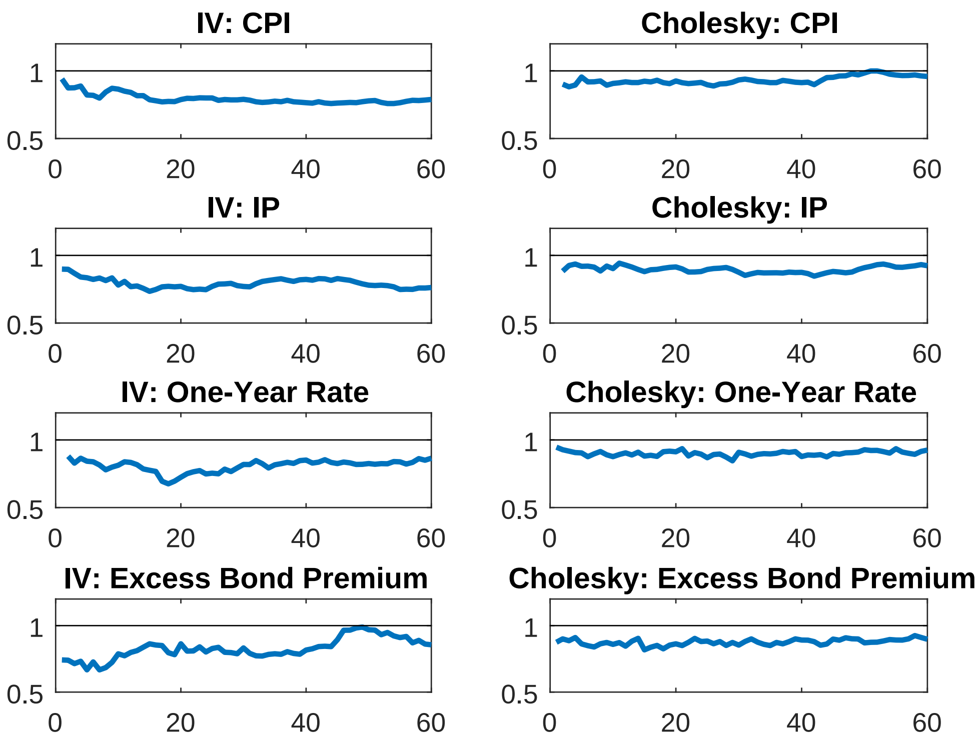
Notes: This figure shows the estimated impulse responses to a monetary policy shock using conventional least squares inference in the VAR considered by Gertler and Karadi (2015). Impulse responses using external instruments (IV) and Cholesky ordering are both shown. The monetary policy shock is normalized to raise the one-year yield by 25 basis points. 95 percent bias-adjusted bootstrap confidence intervals are included.

Figure 20: Relative Width of Confidence Intervals (Normal/LS)



Notes: This figure shows the width of confidence intervals using normal scores divided by the width of the confidence intervals using least squares in the VAR considered by Gertler and Karadi (2015). Impulse responses using external instruments (IV) and Cholesky ordering are both considered. The monetary policy shock is normalized to raise the one-year yield by 25 basis points. All confidence intervals are formed by the bias-adjusted bootstrap with 95 percent nominal coverage.

Figure 21: Relative Width of Confidence Intervals (Wilcoxon/LS)



Notes: This figure shows the width of confidence intervals using Wilcoxon scores divided by the width of the confidence intervals using least squares in the VAR considered by Gertler and Karadi (2015). Impulse responses using external instruments (IV) and Cholesky ordering are both considered. The monetary policy shock is normalized to raise the one-year yield by 25 basis points. All confidence intervals are formed by the bias-adjusted bootstrap with 95 percent nominal coverage.