Proxy Ridge SVAR

Luca Preuße

January 2025

1 SVAR: Cholesky and Proxy identification

The standard B-Model of a VAR(p) of dimension k reads as
$y_t = a + A_1 y_{t-1} + A_2 y_{t-2} ... + A_p y_{t-p} + u_t$
$u_t = B\epsilon_t,\ \epsilon_t \sim N(0, I)$
where ϵ_t are the structural shocks of interest. Since u_t is gaussian the data on y_t identifies a, A_i, Σ_u which can be estimated consistently and efficiently with Maximum Likelihood. We can write $\Sigma_u = BB'$ which is an equation with k^2 unknowns and a maximum of $k^2 - (k^2 - k)/2$ non-redundant equations so that one needs a minimum of $(k^2 - k)/2$ additional restrictions for identification of B. A common identification strategy that delivers exact identification (up to sign normalization) is assuming B to be lower triangular, or estimating B as the Cholesky factor of $\hat{\Sigma}_u$. For this approach the order of variables matters and needs to be justified economically. If interest is only in the effect of one specific shock i, the order between variables 1 to i-1 and i+1 to k does not make a difference in estimating the ith column of B.
Such an approach could be undesirable as a lower triangular B may not always be justified. For example in standard Newkeynesian DSGEs with VAR representation B has no elements that are 0 hence every possible Cholesky ordering is misspecified and leads to inconsistent estimates of B. An alternative identification approach is to use external variables/instruments for identification. Say there exists an instrument z_t with the properties

$E(z_t\epsilon_{-i,t}) = 0, \ E(z_t\epsilon_{i,t}) \neq 0$ so an instrument that is correlated with the target shock but uncorrelated with the non-target shocks. Then the column of B corresponding to the target shock is identified (Angelini and Fanelli, 2019). We can write the assumptions on the instrument as moment conditions and perform GMM estimation minimizing the criterion with respect to B

$Q = g(B, y)'Wg(B, y) \ s.t. \ BB' = \hat{\Sigma}_u$

$g(B, y) = \frac{1}{T}\sum_{t=1}^{T} z_t\hat{\epsilon}_{-i,t}$

$\hat{\epsilon}_t = B^{-1}\hat{u}_t$

$E(g(B, y)) = 0$ by assumption and under exact identification $g(\hat{B}, y) = 0$ so that the weighting matrix W does not matter but the extension to several instruments and overidentification is straightforward.

An obvious drawback of using external variables for identification compared to Cholesky is that it increases estimation variance leading to potentially larger MSE. Consider for simplicity $\hat{u} = u$ and a correctly specified Cholesky order then the MSE of estimating B is 0 while for the Proxy approach still estimation variance remains due to the imperfect correlation between shock and instrument.

2 Proxy Ridge SVAR

To estimate shock i identified by the instrument z_t I propose to minimize the following criterion with respect to B given the reduced form VAR estimates:

$Q = g(B, y)'Wg(B, y) + \lambda\sum_{i,j} v_{i,j}(B(i, j) - B_0(i, j))^2 \ s.t. \ BB' = \hat{\Sigma}_u$

$g(B, y) = \frac{1}{T}\sum_{t=1}^{T} z_t\hat{\epsilon}_{-i,t}$

$\hat{\epsilon}_t = B^{-1}\hat{u}_t$

The first part is equivalent to the GMM criterion. Note that the choice of W is not irrelevant even in the case of exact identification because of the additional shrinkage term. Asymptotically efficient in GMM is to set $W = E((z_t\epsilon_{-i,t})(z_t\epsilon_{-i,t})')^{-1}$. Strengthening the assumption of uncorrelatedness across shocks and instruments to statistical independence this simplifies to a diagonal with the inverse of the instrument variance as elements.

The second term in the criterion allows for shrinkage towards prior restrictions summarized in B_0 by punishing deviations from these restrictions. $v_{i,j}$ allows to weigh different restrictions differently. λ determines the overall strength of the shrinkage. To ensure consistency $\lambda = o(1)$ (Rajkumar, 2019) so that asymptotically the moment conditions dominate the criterion function. A more rigorous analysis on choosing λ by methods of cross-validation I leave for future research. Here I propose to use $\lambda = log(T)/T$. Following Keweloh (2024) I choose

$$v_{i,j} = \begin{cases} 0 & \text{if } B_0(i,j) \text{ unrestricted} \\ \frac{1}{(\hat{B}(i,j)-B_0(i,j))^2} & \text{else} \end{cases}$$

with $\hat{B}$ being some first-stage estimator, for example using a Proxy estimate, so that shrinkage is strong in elements that are close to the restrictions in a first-stage estimator that doesn't use these restrictions for estimation.

3 Simulation Design

The DGP for the SVAR follows a VAR representation of a standard monetary policy DSGE as in Herwartz et al. (2022) and is a 3-dimensional VAR(1) with matrices

$$A = \begin{bmatrix} 0.74 & -0.09 & -0.16 \\ 0.13 & 0.44 & -0.06 \\ 0.24 & 0.30 & 0.53 \end{bmatrix}, \quad B = \begin{bmatrix} 2.32 & -0.48 & -0.41 \\ 0.72 & 2.32 & -0.22 \\ 0.98 & 1.57 & 0.76 \end{bmatrix}.$$

with the third shock being a monetary policy shock of interest. The instrument is generated by $z_t = \beta_0 + \beta_1 \epsilon_{3,t} + \eta_t, \quad \eta_t \sim N(0, \sigma^2)$ where all the coefficients are the estimates of the regression of the Proxy-identified shock in the data application outlined in section 6. I let the instrument strength vary by considering $\sigma = [0.5\hat{\sigma}, 2\hat{\sigma}]$. As sample sizes I consider T=270 which is the sample size in the data application but I let it vary from T=150 to T=1000. As restrictions B_0 I use Cholesky assuming the upper triangle elements of B to be 0 which is a wrong identification assumption in this DGP. I evaluate the MSE of the impulse response function averaged over 500 simulations.

4 Simulation Results

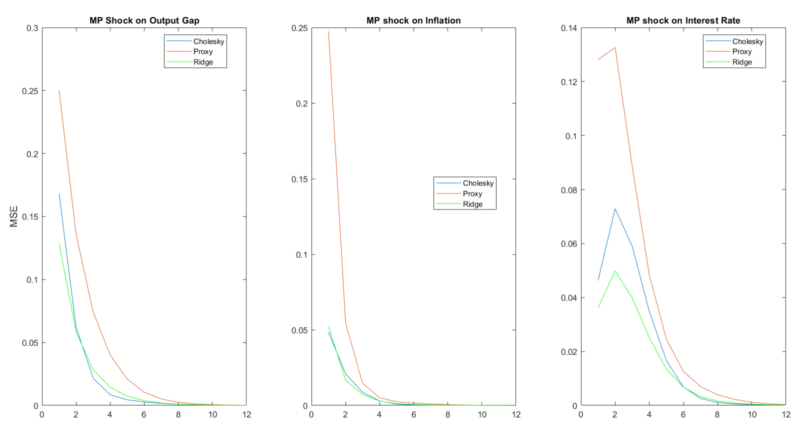

Figure 1: Impulse Response MSE of 500 Simulations for Cholesky, Proxy and Proxy-Ridge identification. T=270 and $\sigma = \hat{\sigma}$.

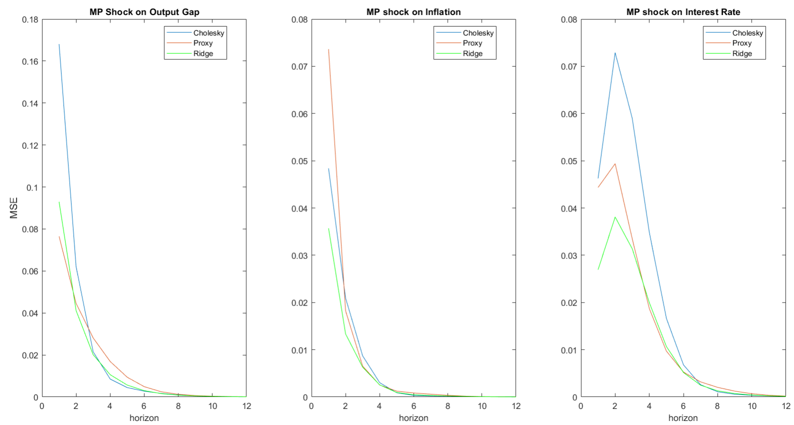

Figure 2: Impulse Response MSE of 500 Simulations for Cholesky, Proxy and Proxy-Ridge identification. T=270 and $\sigma = 0.5\hat{\sigma}$.

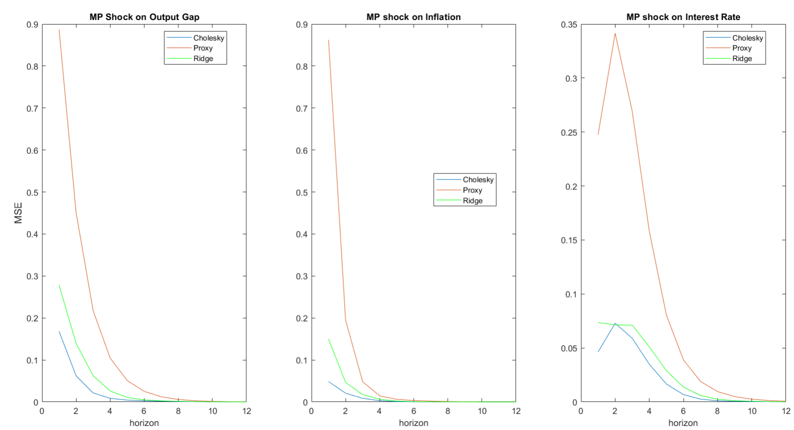

Figure 3: Impulse Response MSE of 500 Simulations for Cholesky, Proxy and Proxy-Ridge identification. T=270 and $\sigma = 2\hat{\sigma}$.

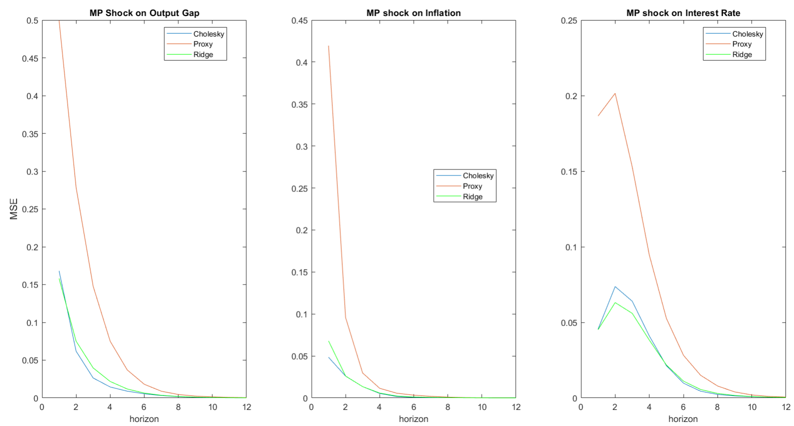

Figure 4: Impulse Response MSE of 500 Simulations for Cholesky, Proxy and Proxy-Ridge identification. T=150 and $\sigma = \hat{\sigma}$.

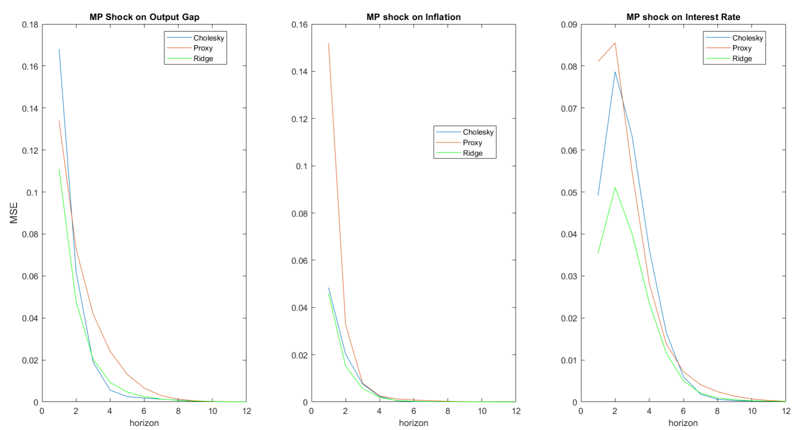

Figure 5: Impulse Response MSE of 500 Simulations for Cholesky, Proxy and Proxy-Ridge identification. T=500 and $\sigma = \hat{\sigma}$.

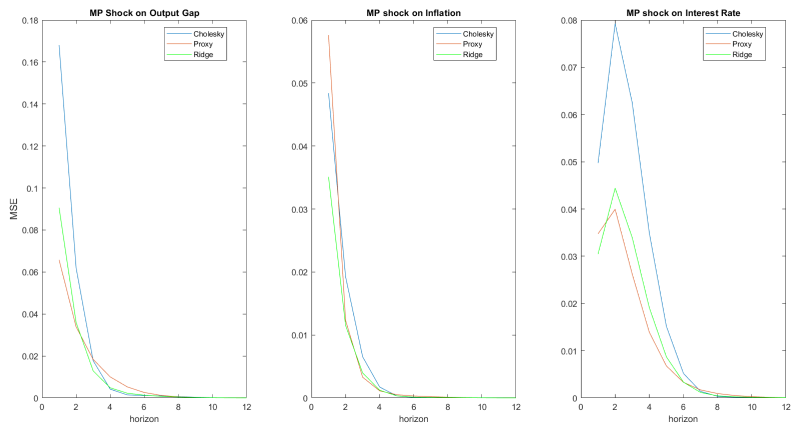

Figure 6: Impulse Response MSE of 500 Simulations for Cholesky, Proxy and Proxy-Ridge identification. T=1000 and $\sigma = \hat{\sigma}$.

As one may expect in large sample sizes and cases where the instrument is strong the Proxy identification performs better than the Cholesky identification and vice versa in small samples and with weak instruments. The suggested Proxy Ridge approach however seems to always be close to the dominating approach and is for most of the simulations better than both competing identification approaches. Also unlike Cholesky, which uses misspecified assumptions, the Proxy-Ridge approach is consistent with the MSE approaching 0 as the sample size increases which one may desire as a property for an estimator even when the inconsistent estimator has better finite sample performance. Combining several identification approaches seems to be recommended based on these simulation results, if one is interested in reducing the MSE.

5 Application

The application is based on Gertler and Karadi (2015) who use high frequency surprises around policy announcements as an instrument for monetary policy shocks. I estimate a VAR(12) with log-Industrial Production, log-CPI, the

One-Year Government Bond Rate and Excess Bond Premium in that order with the third shock being the MP shock. The data is monthly from 1991:1 to 2012:6. A larger sample is available for the variables but not for the instrument. In contrast to me Gertler and Karadi (2015) use the full sample to estimate the reduced form VAR which explains differences in results. However, keeping the sample for variables and instrument consistent is closer to our simulation design.

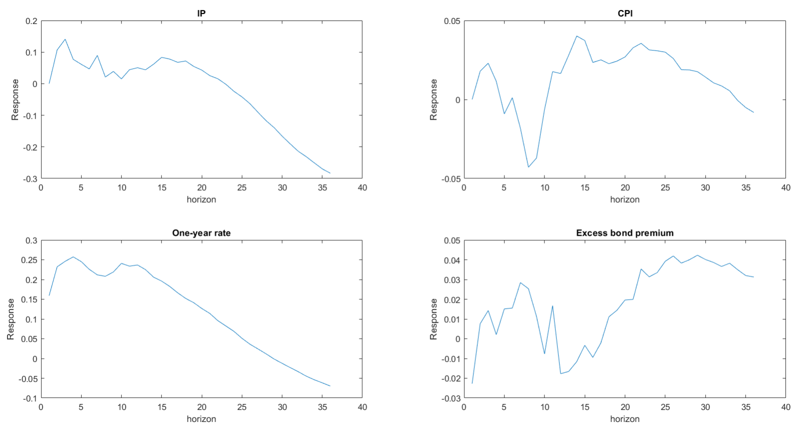

Figure 7: Impulse Responses to Monetary Policy Shock identified with Cholesky.

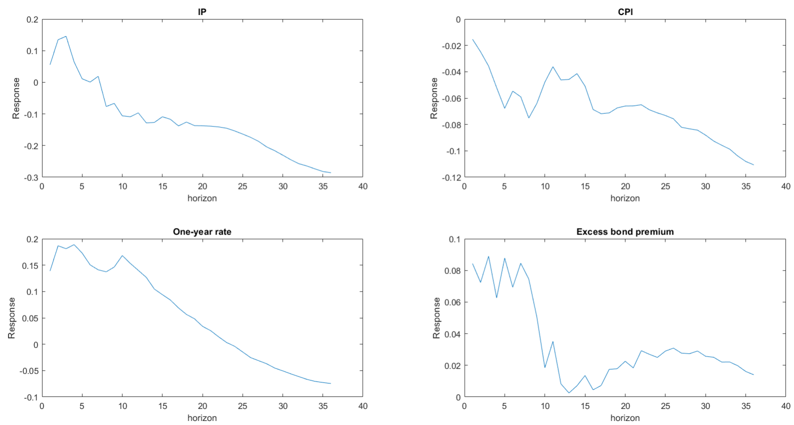

Figure 8: Impulse Responses to Monetary Policy Shock identified with External Instrument.

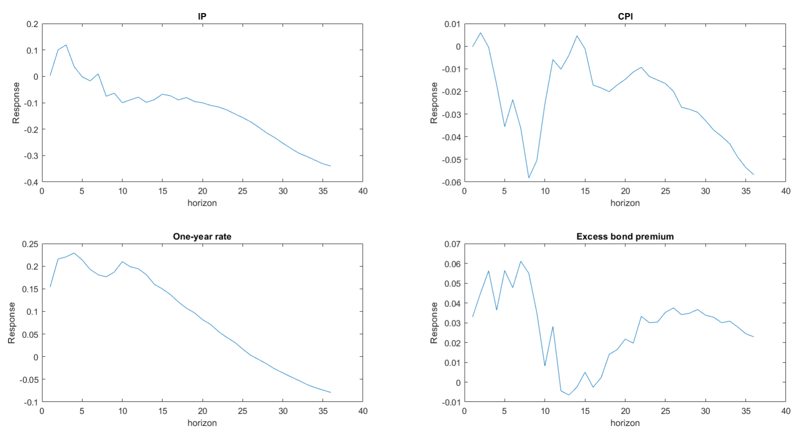

Figure 9: Impulse Responses to Monetary Policy Shock identified with External Instrument and Ridge penalty shrinking towards Cholesky.

The dynamics shown in the impulse responses of the Ridge-Proxy approach are largely similar to the dynamics of the Proxy approach. The cumulated effect on CPI is estimated to be weaker with the Ridge approach estimate being even positive after around 15 months so we do see a bit of a prize puzzle although not as bad as with Cholesky. The cumulated effect of a contractionary monetary policy shock on Industrial Production is estimated to be stronger with Ridge than with Proxy.

6 Summary

In this paper I suggest to combine SVAR identification through external instruments with more classical restrictions like Cholesky through a ridge penalty. A simulation study calibrated around real data and a theoretical DSGE shows that this may lead to reduction in the MSE of estimating impulse responses of monetary policy shocks. For a variety of sample sizes and instrument strengths in simulations this estimator dominates or is close to as good as the better of a Cholesky-only or Proxy-only identification. Future research may look at cross-validation techniques to determine MSE-optimal shrinkage parameter λ as well as inference methods for this type of estimator. The goal of this paper was to motivate the potential of such an estimator for relevant and realistic finite samples and instrument strengths in economic data.

References

G. Angelini and L. Fanelli. Exogenous uncertainty and the identification of structural vector autoregressions with external instruments. *Journal of Applied Econometrics*, 34(6):951–971, 2019.

M. Gertler and P. Karadi. Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76, 2015.

H. Herwartz, H. Rohloff, and S. Wang. Proxy svar identification of monetary policy shocks - monte carlo evidence and insights for the us. *Journal of Economic Dynamics and Control*, 139:104457, 2022.

S. A. Keweloh. Uncertain short-run restrictions and statistically identified structural vector autoregressions, 2024. URL https://arxiv.org/abs/2303.13281.

K. Rajkumar. Ridge regularization for mean squared error reduction in regression with weak instruments, 2019. URL https://arxiv.org/abs/1904.08580.