What Can We Learn From Sign-Restricted VARs?

By Christian K. Wolf*

Traditionally, the identification of business-cycle shocks through Vector Autoregressions (VARs) has relied on *exact* exclusion restrictions: impulse responses are assumed to be zero either on impact or in the long run, and shock proxies are assumed to only correlate with the structural shock of interest (see Ramey, 2016, for a review).[1] Economic theory, of course, rarely justifies such exact exclusion restrictions. In response, a by-now voluminous literature considers a less dogmatic, more credible source of identifying information: restrictions imposed only on the *signs* of impulse responses (e.g. Uhlig, 2005) .

In this short note I use a simple structural business-cycle model to analytically illustrate the workings and limitations of this popular sign restrictions approach. I try to summarize the insights from my analysis in the form of a couple of concrete takeaways for applied macroeconometric practice.

I. Model

I follow Wolf (2020) and study a variant of the canonical three-equation New Keynesian model. The equations are

$$\text{(IS)}\quad y_t \;=\; E_t\left(y_{t+1}\right)-\left(i_t-E_t\left(\pi_{t+1}\right)\right)+\sigma^d\varepsilon_t^d$$

$$\text{(NKPC)}\quad \pi_t \;=\; \kappa y_t+\beta E_t\left(\pi_{t+1}\right)-\sigma^s\varepsilon_t^s$$

$$\text{(TR)}\quad i_t \;=\; \phi_\pi\pi_t+\sigma^m\varepsilon_t^m$$

where $\varepsilon_t \equiv (\varepsilon_t^d,\varepsilon_t^s,\varepsilon_t^m)' \sim N(0,I)$; y is real output; i is the nominal interest rate; and π is inflation. The model consists of three equations and features three shocks: an IS-relation (demand block) subject to a demand shock ε_t^d; a New Keynesian Phillips curve (supply block) subject to a supply/cost-push shock ε_t^s; and a monetary policy rule subject to a monetary shock ε_t^m. Under standard restrictions on parameter values, the solution satisfies

$$(1)\qquad \begin{pmatrix} y_t \\ \pi_t \\ i_t \end{pmatrix} = \underbrace{\begin{pmatrix} + & + & - \\ + & - & - \\ + & - & + \end{pmatrix}}_{\Theta} \times \begin{pmatrix} \varepsilon_t^d \\ \varepsilon_t^s \\ \varepsilon_t^m \end{pmatrix}$$

I give the full expression for the impulse response matrix Θ in Appendix A. For ease of reference, I will denote the entry of Θ corresponding to the response of variable $j \in \{y,\pi,i\}$ to shock $k \in \{d,s,m\}$ by Θ_{jk}.

I consider a researcher who wishes to learn about the effects of monetary policy shocks—arguably *the* canonical problem in the vast structural VAR literature. The researcher observes model-generated data $x_t = (y_t,\pi_t,i_t)'$, and seeks to estimate Θ_{ym}, the response of output to a monetary shock. For simplicity, I assume that the researcher has access to an arbitrarily large sample, so I can treat the covariance matrix

$$(2)\qquad \Sigma_x \;\equiv\; \text{var}(x_t) \;=\; \Theta\Theta'$$

as known. Since $\tilde{\Theta}\tilde{\Theta}' = \Theta\Theta'$ for $\tilde{\Theta} = \Theta P'$ with P an orthogonal matrix, it follows that knowledge of (2) identifies the model only up to orthogonal rotations. Given any such P, the implied impulse response matrix $\tilde{\Theta}$ is in fact the impulse response matrix corresponding to some *linear combination* $\tilde{\varepsilon}_t$ of the true structural shocks, where

$$(3)\qquad \tilde{\varepsilon}_t \;\equiv\; P\varepsilon_t$$

To learn about Θ_{ym}, the researcher needs to find additional identifying information that (i) is consistent with $P = I$ (giving $\tilde{\varepsilon}_t = \varepsilon_t$ and $\tilde{\Theta}_{ym} = \Theta_{ym}$) and (ii) rules out many (ideally all) other P's.

* MIT Department of Economics, 50 Memorial Drive, Cambridge, MA 02142 (email: ckwolf@mit.edu). I received helpful comments from Mikkel Plagborg-Møller and Jim Hamilton.

[1] More precisely, identification in the first two cases relies on the stated exclusion restrictions together with the assumption of *invertibility*: the shocks are assumed to lie in the span of observed macro aggregates.

II. Sign Restrictions on Impulse Responses

Following Uhlig (2005), I assume that the researcher is willing to impose that contractionary monetary policy shocks increase interest rates and lower inflation. Thus, in addition to (2), she also knows that the impulse response matrix Θ must satisfy the following sign pattern:

$$\Theta = \begin{pmatrix} ? & ? & ? \\ ? & ? & - \\ ? & ? & + \end{pmatrix} \tag{4}$$

This restriction is promising: monetary shocks are unique in that they move interest rates and inflation in opposite directions, so requiring a sign pattern as in (4) will certainly prevent the researcher from mislabeling the demand and supply shocks ε_t^d and ε_t^s as monetary shocks ε_t^m. But is that enough to learn something about Θ_{ym}?

A. The Identified Set

Rather than pinning down a single value for Θ_{ym}, the sign restrictions in (4) give us a *range*: they cannot rule out any $\tilde{\Theta}_{ym} \in [\underline{\Theta}_{ym}, \overline{\Theta}_{ym}]$, where the lower and upper bounds are defined as the solutions to the optimization programs

$$\inf_p/\sup_p \quad \Theta_{1,\bullet} \cdot p \tag{5}$$

subject to the unit-length requirement $||p|| = 1$ (from the covariance matrix (2)) and the two inequality constraints (from the sign restrictions (4))

$$\Theta_{2,\bullet} \cdot p < 0, \quad \Theta_{3,\bullet} \cdot p > 0 \tag{6}$$

It is easy to see from this program that we will generally have $\underline{\Theta}_{ym} < 0 < \overline{\Theta}_{ym}$: as long as Θ is invertible, both strictly positive as well as strictly negative values of $\Theta_{1,\bullet} \cdot p$ are consistent with (6); the requirement that $||p|| = 1$ then simply rescales impulse responses, without changing their signs.[2]

Why did the sign restrictions in (4) fail to pin down the sign of Θ_{ym}? Clearly the vector $p = (0, 0, 1)'$—which corresponds to perfect identification, i.e. $\tilde{\varepsilon}_t^m \equiv p'\varepsilon_t = \varepsilon_t^m$—is consistent with the imposed restrictions, so $\Theta_{ym} \in [\underline{\Theta}_{ym}, \overline{\Theta}_{ym}]$. Furthermore, neither *pure* demand nor supply shocks satisfy the imposed restrictions, so $p = (1, 0, 0)'$ and $p = (0, 1, 0)'$ are ruled out. Unfortunately, however, linear combinations of expansionary demand and supply shocks can satisfy (6), yet will increase output—an identification challenge that I have referred to as shock "masquerading" in prior work (Wolf, 2020). Panel (a) of Figure 1 provides a graphical illustration, displaying the identified set in terms of the shock weights $p = (p_d, p_s, p_m)'$ for $\tilde{\varepsilon}_t^m \equiv p'\varepsilon_t$. The orange region indicates shock weights that are consistent with the imposed sign restrictions and give a negative identified response of output, $\tilde{\Theta}_{ym}$. As expected, the vector $p = (0, 0, 1)'$ corresponding to perfect identification lies in this orange region. However, there is also a blue region: positive weights on demand and supply shocks can also be consistent with the sign restrictions, but counterfactually increase output.

Additional restrictions.

Two popular refinements of the simple baseline restrictions in (4) are (i) restrictions on other shocks (i.e., the other columns of Θ), and (ii) restrictions on entire impulse response *paths* (applicable for richer, dynamic environments). I provide a brief discussion of both cases here, with details relegated to Appendix B.

First, suppose that a researcher is also willing to restrict the signs of all impulse responses to demand and supply shocks, requiring that

$$\Theta = \begin{pmatrix} + & + & ? \\ + & - & - \\ + & - & + \end{pmatrix} \tag{7}$$

To see how such simultaneous identification of multiple shocks may help, take any of the "masquerading" p's in the mis-identified blue region in panel (a) of Figure 1. For the proposed additional restrictions to rule those out, it must *not* be possible to find any other vectors $p^\perp$ orthogonal to p and

[2] I show in Appendix A that, for standard model parameterizations, Θ is in fact invertible.

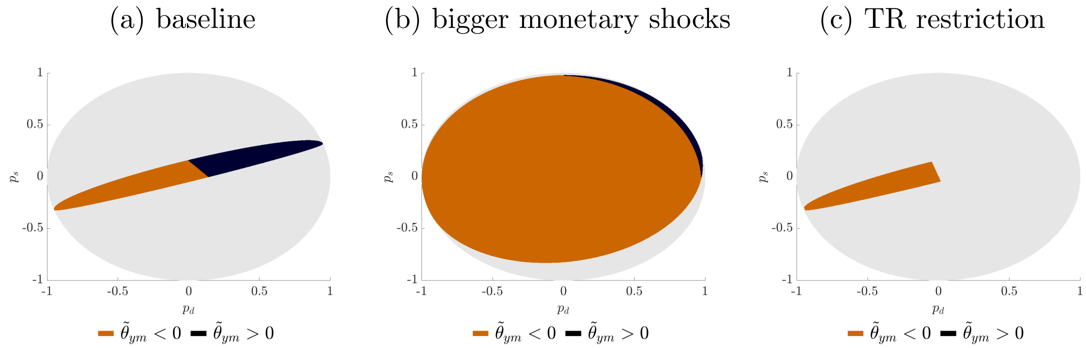

Figure 1. : Identified sets of monetary policy shocks.

Note: Regions of the identified set derived from (5) that give $\tilde{\Theta}_{ym} < 0$ (orange) and $\tilde{\Theta}_{ym} > 0$ (blue) vs. all possible shock weights (grey). p_d and p_s are on the axes, and I residually set $p_m = \sqrt{1 - p_d^2 - p_s^2}$. Baseline model in (a), more volatile monetary shocks in (b), and adding the restriction (8) in (c). Parameterization in Appendix A.

with impulse responses as in the first and second columns shown in (7). One simple sufficient condition for this to indeed not be possible is that monetary shocks are dominant: if $\sigma^m \gg \{\sigma^d, \sigma^s\}$, then only pure demand and supply shocks (i.e., $p^\perp = (1,0,0)'$ and $p^\perp = (0,1,0)'$) satisfy the restrictions in the first two columns of (7). But then the identified monetary policy vector p would have to equal $(0,0,1)'$, so the model is in fact point-identified.

Second, a researcher may impose the restrictions (4) for many time periods. It is straightforward to see how such dynamic restrictions may prove useful: for example, suppose that demand and supply shock impulse responses erratically change their sign from period to period, while monetary impulse responses show a gradual decay. In that case, the masquerading demand and supply shocks would satisfy the imposed restrictions on impact, but not along the entire impulse response path, thus again tightening identified sets.

Neither of these conditions is however particularly plausible empirically: monetary shocks are unlikely to be a main driver of cyclical fluctuations, and impulse responses in both theory and data tend to not quickly alternate in sign. It is thus not surprising that, in practice, allowing for additional restrictions on many shocks or many time periods does not usually help much with the masquerading problem. I provide an illustration in Figure B.1.

B. *The Haar Prior*

In Section II.A, I studied the entire identified set induced by the sign restrictions (4). I did so because, without further information, sign restrictions alone simply do not allow us to distinguish between any of the $\tilde{\Theta}_{ym}$ in the set $[\underline{\Theta}_{ym}, \overline{\Theta}_{ym}]$. Much of the applied sign restrictions literature, however, proceeds rather differently, reporting instead the *distribution* over $[\underline{\Theta}_{ym}, \overline{\Theta}_{ym}]$ induced by a particular prior over the space of orthogonal rotation matrices P—the so-called uniform Haar measure (Uhlig, 2005; Rubio-Ramírez, Waggoner and Zha, 2010).[3]

Why should we care about the distribution induced by this Haar prior? The advantage of the model-based perspective taken here is that it allows us to give an explicit structural interpretation to that prior. It follows from (3) that the Haar prior is a prior over the matrix of shock weights P; for a single shock, it is a prior over the unit-length shock weight vector p. The Haar measure—as the natural notion of a uniform measure over the unit sphere—regards each of these shock weight vectors

[3]Since I treat Σ_x as known, my analysis in this paper should be interpreted as applying to the large-sample limit of the conventional Bayesian procedure.

p as ex-ante equally likely.[4] This implies that the Haar-induced posterior over impulse responses is necessarily dominated by the most volatile shocks: those give the largest entries in the impulse response matrix Θ, and so also dominate the identified responses $\tilde{\Theta}_{\bullet m} = \Theta \cdot p$. Panel (b) of Figure 1 illustrates. There I report the same objects as in panel (a), but now for a model with a much higher relative volatility of monetary shocks. The dark blue area of masquerading positive demand and supply shocks is still present, but it now corresponds to a much smaller *share* of the identified set, and so will receive relatively little weight in the researcher's posterior. Thus, even though the identified set continues to include both negative and positive output responses to identified contractionary monetary shocks, the imposed prior now implies that a Bayesian researcher regards negative responses as overwhelmingly likely.

What do those observations imply for applied practice? SVAR methods are often used to study the transmission of policy shocks (either monetary or fiscal), and those shocks are not regarded as important sources of business-cycle fluctuations. In those cases, the Haar prior can be expected to amplify the masquerading problem: the posterior will be dominated by other, more volatile shocks, thus potentially centering the researcher's posterior away from the true effects of the policy shocks. For example, in my monetary policy application, typical values for the relative volatilities of policy and non-policy shocks imply that the Haar-induced posterior has most mass concentrated on counterfactual positive output responses, precisely because of the Haar prior's over-weighting of the more volatile demand and supply shocks. More generally, this sensitivity of the researcher's posterior to the relative importance of shocks—a feature of shocks that is orthogonal to their dynamic causal effects, and so to what we actually want to learn about—means that the computational strategy of sampling from the Haar prior does not in fact give us an agnostic default inference procedure.

[4] See Baumeister and Hamilton (2015) for a visual representation of this prior. In particular, their illustrations reveal that uniformity over points on the unit sphere does not necessarily correspond to uniformity about the individual entries of the weight vector p.

C. *Restricting Structural Elasticities*

The analysis in Sections II.A and II.B relied on restrictions on impulse responses Θ. Alternatively, researchers may also entertain restrictions on the structural elasticities in the equilibrium system (IS) - (TR)—i.e., restrictions on entries of Θ^{-1}.

In my model environment, one such restriction—a sign restriction on the implied monetary policy rule, as in Arias, Caldara and Rubio-Ramirez (2019)—turns out to sharply tighten the identified set. Specifically, suppose that the researcher complements the sign restrictions in (4) with the requirement that the coefficient on output in the monetary authority's rule is weakly positive (the true value is zero). The identified set can then be derived by an extended version of the program (5), now also featuring the constraint that

$$(8) \qquad (-p' \cdot \Theta^{-1}_{\bullet,1})/(p' \cdot \Theta^{-1}_{\bullet,3}) \geq 0$$

Panel (c) in Figure 1 reveals that this restriction is enough to eliminate the blue region of masquerading expansionary demand and supply shocks.[5] The intuition is simple: if positive demand and supply shocks are mis-identified as a contractionary monetary shock, then a linear combination of the (IS) and (NKPC) equations is mis-identified as the Taylor rule. These two equations, however, together postulate a negative relationship between output and interest rates, with an implied Taylor rule output coefficient proportional to $-(\frac{p_d}{\sigma^d} + \kappa\frac{p^s}{\sigma^s})$ and so inconsistent with (8) for $p_d > 0$ and $p_s > 0$.

III. Conclusions

The relative agnosticism of sign restrictions makes them an attractive addition to the SVAR toolkit. In this short note, I have used a simple structural model to gain some

[5] This sharp result relies on my assumption that the true output coefficient in the monetary rule is zero. I consider the more general case in Wolf (2020).

insights into the workings and limitations of sign restrictions as identifying information. Three lessons emerged.

1) To identify a given structural shock, it is necessary but generally not sufficient to impose sign restrictions that are satisfied only by that shock—intuitively, the identification challenge is that linear combinations of other shocks can also satisfy the imposed restrictions. In my model, this shock "masquerading" problem was severe enough to imply that sign restrictions on interest rates and inflation do not allow us to say anything about the sign of the output response to monetary shocks. The problem remains hard to escape even if sign restrictions are imposed at multiple horizons or on multiple shocks.

2) Sampling from the Haar prior, as done in standard Bayesian implementations of sign restrictions, does not yield an agnostic default inference procedure, as it implicitly over-weights the most volatile underlying structural shocks—a feature of the data that has little to do with the dynamic causal effects that the researcher is trying to estimate. Inference on entire identified sets, using for example the procedures of Giacomini and Kitagawa (2021), sidesteps these concerns.

3) Nevertheless, SVAR identification does not need to go all the way back to exact exclusion restrictions to make progress. Recently, the SVAR toolkit has been expanded to allow for: (i) sign restrictions on structural elasticities; (ii) explicitly stated and defended probabilistic priors on either impulse response functions (Plagborg-Møller, 2019) or structural elasticities (Baumeister and Hamilton, 2015), replacing the Haar prior; and (iii) range restrictions on impulse response magnitudes (Kilian and Murphy, 2012). Put together with sign restrictions on impulse responses, this enlarged set of tools may offer an attractive middle ground for the estimation of dynamic causal effects.

REFERENCES

Arias, Jonas E, Dario Caldara, and Juan F Rubio-Ramirez. 2019. "The systematic component of monetary policy in SVARs: An agnostic identification procedure." *Journal of Monetary Economics*, 101: 1–13.

Baumeister, Christiane, and James D Hamilton. 2015. "Sign restrictions, structural vector autoregressions, and useful prior information." *Econometrica*, 83(5): 1963–1999.

Giacomini, Raffaella, and Toru Kitagawa. 2021. "Robust Bayesian inference for set-identified models." *Econometrica*, 89(4): 1519–1556.

Kilian, Lutz, and Daniel P Murphy. 2012. "Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models." *Journal of the European Economic Association*, 10(5): 1166–1188.

Plagborg-Møller, Mikkel. 2019. "Bayesian inference on structural impulse response functions." *Quantitative Economics*, 10(1): 145–184.

Ramey, Valerie A. 2016. "Macroeconomic Shocks and Their Propagation." In *Handbook of Macroeconomics*. Vol. 2, , ed. John B. Taylor and Harald Uhlig, Chapter 2, 71–162. Elsevier.

Rubio-Ramírez, Juan F., Daniel F. Waggoner, and Tao Zha. 2010. "Structural vector autoregressions: Theory of identification and algorithms for inference." *Review of Economic Studies*, 77(2): 665–696.

Uhlig, Harald. 2005. "What are the effects of monetary policy on output? Results from an agnostic identification procedure." *Journal of Monetary Economics*, 52(2): 381–419.

Wolf, Christian K. 2020. "SVAR (Mis) identification and the Real Effects of Monetary Policy Shocks." *American Economic Journal: Macroeconomics*, 12(4): 1–32.