SVAR (Mis-)Identification and the Real Effects of Monetary Policy Shocks*

Christian K. Wolf
Princeton University

This version: November 3, 2019
First version: August 25, 2018

Abstract: I argue that the seemingly disparate findings of the recent empirical literature on monetary policy transmission are in fact all consistent with the same standard macro models. Weak sign restrictions, which suggest that contractionary monetary policy if anything boosts output, present as policy shocks what actually are expansionary demand and supply shocks. Classical zero restrictions are robust to such misidentification, but miss short-horizon effects. Two recent approaches – restrictions on Taylor rules and external instruments – instead work well. My findings suggest that empirical evidence is consistent with models in which the real effects of monetary policy are larger than commonly estimated.

Keywords: Monetary policy, SVARs, sign restrictions, zero restrictions, Haar prior, external instruments, invertibility. *JEL codes:* C3, E3, E5.

*Email: ckwolf@princeton.edu. I received helpful comments from Jonas Arias, Thorsten Drautzburg, Jim Hamilton, Marek Jarociński, Peter Karadi, Matthias Meier, Ben Moll, Ulrich Müller, Emi Nakamura, Mikkel Plagborg-Møller, Chris Sims, Harald Uhlig, Mark Watson, and seminar participants at several venues. I am particularly grateful to the editor, Giorgio Primiceri, as well as the anonymous referees. Part of this research was conducted while I was visiting the European Central Bank, whose hospitality is gratefully acknowledged. I also acknowledge support from the Alfred P. Sloan Foundation and the Macro Financial Modeling Project. A previous version of this paper circulated under the title "Masquerading Shocks in Sign-Restricted VARs".

1 Introduction

A central question in empirical macroeconomics is the response of the economy to changes in monetary policy. Going back to Sims (1980), a long literature has tackled this question using structural vector autoregressions (SVARs), with policy shocks identified through zero restrictions on the contemporaneous response of macro aggregates to policy changes. This early literature suggests that a policy tightening indeed reduces real activity, if only moderately so and with a delay. Recent work challenges this consensus. Uhlig (2005) casts doubt on the conventional timing restrictions, proposes a weaker identification procedure based on uncontroversial sign restrictions, and finds that, if anything, contractionary monetary shocks *boost* output. Yet more recently, refinements of Uhlig's identification scheme (Arias et al., 2019) or identification based on external instruments (Gertler & Karadi, 2015) tend to qualitatively re-store conventional wisdom, and in fact suggest somewhat larger and faster real effects than previously believed. At the same time, more and more studies have started to outright question the ability of SVARs to reliably identify shock transmission (e.g. Plagborg-Møller, 2019; Nakamura & Steinsson, 2018b), raising concerns about the informativeness of macroeconomic aggregates for hidden structural shocks – the so-called non-invertibility problem. Evidently, a consensus on the real effects of monetary policy remains elusive.

In this paper I show that, when viewed through the lens of standard structural models, these apparent inconsistencies across different empirical methods are not at all surprising, but exactly what we should expect. The argument is simple: I fix a single common structural model as my data-generating process, characterize the probability limits of various popular empirical strategies, and show that the estimators disagree in exactly the same fashion as they do in real data. Sign restrictions, as in Uhlig (2005), are vulnerable to expansionary demand and supply shocks "masquerading" as contractionary monetary policy shocks, which then seemingly boost – rather than depress – output. Standard impact zero restrictions on output impulse responses give classical results because they implicitly safeguard against this particular form of mis-identification, but at the cost of understating the (short-horizon) real effects of monetary policy. Direct restrictions on the implied Taylor rule of the monetary authority (as in Arias et al. (2019)) or external instruments (IVs) instead robustly estimate the true model-implied aggregate effects of monetary policy shocks. In principle, both approaches are vulnerable to non-invertibility concerns, but in practice either solution can work well, as monetary policy shocks are often *near*-invertible.

My analysis builds on a fully specified structural model of monetary policy transmission in the mold of Woodford (2003), Galí (2008) or Smets & Wouters (2007). In line with empirical practice, I assume that the econometrician observes data on output, inflation and the policy rate generated from the model, estimates their VAR representation, and identifies structural shocks using different identification schemes. To avoid conflating estimation and identification uncertainty, I allow the econometrician to observe an infinitely long sample, so she is able to perfectly recover the true population reduced-form VAR representation. Against this reduced-form VAR, I then characterize the probability limits of various popular estimators of monetary policy transmission. In particular, for each estimator, I am able to write the identified shocks – what the researcher will call a "monetary policy shock" – as a linear combination of the true shocks of the underlying model. Perfect identification corresponds to a coefficient of 1 on the shock of interest, and 0 on all other shocks.

I first address the non-invertibility problem. I show that, for *any* possible SVAR identification scheme, the coefficient on the actual monetary policy shock is bounded above by the R^2 in a regression of that shock on past and current values of the observed macro variables. Under invertibility ($R^2 = 1$), identification can thus in principle succeed; if instead the R^2 is small, then conventional SVAR identification schemes will invariably fail. My first result is that, because monetary policy shocks are – at least in my models – the only shock to drive interest rates and inflation in opposite directions, any VAR that includes these macro aggregates is likely to give a high R^2. For example, even for a small VAR in (y, π, i), estimated on data generated by the model of Smets & Wouters, the R^2 is 0.8702. With the invertibility assumption nearly satisfied, we know that *some* SVAR identification scheme will at least approximately recover true impulse responses. In the remainder of the paper, I ask whether any of the popular standard SVAR estimators attain this near-perfect identification.

I begin with the identification scheme of Uhlig (2005). He defines as a candidate monetary policy shock any shock that moves interest rates and inflation in opposite directions.[1] Since different linear combinations of reduced-form forecasting errors are consistent with these restrictions, his procedure will only provide set identification – it will not identify a single SVAR, but a set, and thus a set of candidate "monetary policy shocks." As discussed above, the true monetary policy shock *uniquely* moves policy rates and inflation in opposite directions; appealingly, this not only ensures near-invertibility, but also implies that the actual true shock will lie in the set of acceptable candidate shocks, while any of the other pure shocks will not. Troublingly, however, accurate identification is still not guaranteed, as the

[1] He imposes further restrictions, designed chiefly to disentangle monetary policy shocks from money demand shocks. My models feature no such shocks, so I abstract from his additional restrictions.

identified set may contain *linear combinations* of other structural shocks. I find that this "masquerading" problem is prominent in my structural models, where many acceptable candidate "monetary policy shocks" counterfactually increase aggregate output. Intuitively, the right linear combination of expansionary demand and supply shocks can also push inflation and interest rates in opposite directions, but of course boosts output.

In large-sample Bayesian analysis of sign-restricted VARs, posterior uncertainty over the identified set is exclusively governed by the prior (Baumeister & Hamilton, 2015; Watson, 2019). I show that the Haar prior – the most popular prior in applied work (Uhlig, 2005; Rubio-Ramírez et al., 2010) – automatically puts more mass on more volatile structural shocks. But since, in my structural models, demand and supply shocks are more volatile than monetary policy shocks, most posterior mass is automatically put on the "masquerading" shock combinations that counterfactually increase real output. This conclusion agrees exactly with posterior uncertainty over identified sets reported in Uhlig's analysis.[2]

I next consider the performance of zero or near-zero identifying restrictions. Uhlig (2005), in his review of the classical zero restriction literature, finds the zero *output* restriction to be central to the old conventional wisdom. My model-based analysis reveals that this key role for the impact output restriction is not an accident, but in fact an economically sensible feature of identified sets. The logic is simple: In purely sign-identified SVARs, counterfactual positive output responses are generated by masquerading expansionary supply and demand shocks. These shocks move interest rates in opposite directions, but output in the same direction, and thus imply very large impact output *multipliers* of monetary policy interventions. Even moderate bounds on these multipliers are enough to eliminate most combinations of positive demand and supply shocks from the identified set, thus substantially tightening inference around the truth. Literal zero restrictions of course afford most tightening, but – at least in models with small, but non-zero impact effects – lead the researcher to robustly *understate* the short-horizon real effects of monetary policy shocks.

Two alternative recently proposed identification schemes do not suffer from this defect. First, Arias et al. (2019) combine the benchmark identification scheme of Uhlig with additional sign restrictions on implied Taylor rule coefficients, and show that their identification scheme restores conventional wisdom. Yet again, this finding can be rationalized through standard structural models: I show that, in regions of the identified set where demand and

[2] My results should not be taken to imply that the Haar prior is incorrectly imposed in popular work, nor that the derived Bayesian posterior sets are invalid. I merely clarify the role of this particular choice of prior in shaping posterior uncertainty over the identified sets implied by sign restrictions alone.

supply shocks masquerade as contractionary monetary policy shocks, the coefficient on output in the implied mis-identified "Taylor rule" is invariably (and counterfactually) negative. Restricting the coefficient to be positive thus markedly improves identification. Second, several researchers have proposed to identify monetary policy SVARs using external instruments (e.g. Gertler & Karadi, 2015). Plagborg-Møller & Wolf (2019a) show that, even with a valid external instrument, the standard SVAR-IV estimator is biased under non-invertibility. The bias, however, is proportional to the reciprocal of the R^2 in a regression of the monetary policy shock on lags of the macro aggregates, and so, by near-invertibility, likely to be small.

My results have important implications for macro-econometric practice in general, and the study of monetary policy transmission in particular. The review of the agnostic identification scheme in Uhlig (2005) reveals that, for tight and reliable inference, it is not enough to ensure that the imposed sign restrictions are uniquely satisfied by the shock of interest. Linear combinations of other structural shocks can masquerade as the shock of interest and thus lead inference astray. In particular, if these rival shocks are more volatile than the shock of interest, then the popular Haar prior is likely to focus attention on the mis-identified region of the identified set. For monetary policy transmission, my results encouragingly suggest that, first, recent advances in identification effectively address the masquerading problem, and second, even small sets of macro observables may carry a lot of information about policy shocks. Viewed in this light, I conclude that existing empirical work quite consistently paints the picture of significant, medium-sized effects of monetary policy on the real economy.

Literature. My work relates to several strands of literature. First, I provide a unifying model-based perspective on recent advances in the empirical study of monetary policy transmission. In particular, my results lend support to recent empirical work identifying medium-sized real effects of monetary policy via a variety of quite different identification schemes: the narrative evidence reviewed in Coibion (2012), the external SVAR-IV approach of Gertler & Karadi (2015) and Jarociński & Karadi (2018), the Taylor rule restrictions of Arias et al. (2019), and heteroskedasticity-based identification of Brunnermeier et al. (2017). In its attempt to reconcile different empirical findings, my work shares similarities with Mertens & Ravn (2014) and Caldara & Kamps (2017). I show that restrictions on either impact output responses or on the VAR-implied Taylor rule parameter are robustly sufficient to generate negative output responses in monetary policy SVARs.

Second, I offer several novel results on the relation between structural macro models and SVAR representations, in particular for the non-invertible case. The mapping from model parameters to VAR coefficients, and from primitive structural shocks to SVAR-identified shocks, is characterized in detail in Fernández-Villaverde et al. (2007) and Giacomini (2013). Relative to those papers, I offer additional insights by tying the connection between SVAR and model shocks to quantitative measures of the degree of invertibility. In particular, and perhaps somewhat surprisingly, I show that standard macro aggregates can be informative for monetary policy shocks even if those shocks are largely irrelevant for aggregate business-cycle fluctuations (Ramey, 2016; Plagborg-Møller & Wolf, 2019a).

Third, I add several cautionary results to the fast-growing literature on sign-based set identification in empirical macro-econometrics. The sign restrictions methodology for the identification of SVARs was pioneered by Faust (1998), Canova & De Nicolo (2002) and Uhlig (2005). A comprehensive algorithm for inference, relying on the Haar prior, is developed in Rubio-Ramírez et al. (2010). Similar to Paustian (2007) and Castelnuovo (2012), my results reveal the common minimal requirement for sign-based analysis – that *only* the shock of interest satisfy all imposed sign restrictions – to be necessary, but not sufficient for reliable inference (also see Uhlig, 2017). Relative to these earlier contributions, my analysis adds further insights by explicitly characterizing the model-implied (mis-)identified set in terms of the underlying true structural disturbances, and then using the "masquerading shocks" interpretation to rationalize the importance of the Haar prior in shaping posterior uncertainty over this identified set. Relatedly, Paustian (2007) and Canova & Paustian (2011) emphasize that sign restrictions are likely to perform well for sufficiently volatile shocks. My analysis of Bayesian posteriors over identified sets shows that this conclusion is *exclusively* driven by the prior: If, in a given model and with a given SVAR identification scheme, the researcher is unable to sign the response of a variable of interest to a certain shock, then she would be unable to sign the response even if the shock of interest were arbitrarily volatile. Equivalently, for a judiciously chosen prior, the Bayesian posterior probability assigned to a positive (say) impulse response for the variable of interest can always be made arbitrarily large or small, whatever the underlying relative shock volatilities (also see Giacomini & Kitagawa, 2016).

Outline. Section 2 presents my model laboratories, characterizes the mapping from structural model to SVAR estimand, and argues for robust (near-)invertibility of conventional monetary policy shocks. Sections 3 to 5 then interpret recently popular empirical estimators through the lens of the model laboratories. Finally Section 6 concludes. Appendix A provides further details and selected proofs, and a supplementary appendix is available online.[3]

[3]See https://www.christiankwolf.com/research. My webpage also contains codes for replication.

2 VAR analysis in structural models

A wide class of popular structural models admits a VAR representation for observable macro aggregates, and so gives well-defined population estimands for different SVAR estimators of structural shock transmission. In Section 2.1, I outline two particular model laboratories. Section 2.2 characterizes the probability limits of generic SVAR estimators applied to artificial model-generated data. Finally, in Section 2.3, I leverage knowledge of the underlying data-generating process to link SVAR-estimated "structural" shocks to the true disturbances of the structural model, and in particular connect my results to SVAR non-invertibility.

This section mostly reviews relatively standard material; in particular, the only result novel to this paper is my characterization of SVAR estimands under non-invertibility. As such, the analysis here merely collects the tools necessary for my model-based interpretation of SVAR-implied identified sets in Sections 3 to 5.

2.1 Model laboratories

For most of this paper, I will study the properties of popular SVAR identification strategies through the lenses of two structural models. First, I consider a simple variant of the canonical three-equation New Keynesian model (Galí, 2008; Woodford, 2003). This model is simple enough to conveniently and transparently provide closed-form illustrations of my results. Second, I use the quantitatively more realistic model of Smets & Wouters (2007) to show that all intuitions also survive in a richer environment. In particular, the Smets-Wouters model allows me to judge the likely importance of VAR mis-specification due to non-invertibility.

Section B.1 of the Online Appendix extends my results to other environments; notably, I consider a dynamic three-equation model as well as alternative model variants with *passive* monetary policy rules (Castelnuovo & Surico, 2010; Leeper & Leith, 2016).

THE THREE-EQUATION MODEL. Detailed derivations of the conventional three-equation New Keynesian model are offered in Galí (2008) and Woodford (2003). To allow the cleanest possible study of the various popular SVAR estimators, I consider a particularly simple static variant of this model, without any exogenous or endogenous persistence:

$$y_t = \mathbb{E}_t(y_{t+1}) - (i_t - \mathbb{E}_t(\pi_{t+1})) + \sigma^d \varepsilon_t^d \tag{IS}$$

$$\pi_t = \beta \mathbb{E}_t(\pi_{t+1}) + \kappa y_t - \sigma^s \varepsilon_t^s \tag{NKPC}$$

$$i_t = \phi_\pi \pi_t + \phi_y y_t + \sigma^m \varepsilon_t^m \tag{TR}$$

where $(\varepsilon_t^d, \varepsilon_t^s, \varepsilon_t^m)' \sim N(0, I)$. y is real output, i is the nominal interest rate (the federal funds rate), and π is inflation. The model has three structural disturbances: a demand shock ε^d, a supply shock ε^s and a monetary policy shock ε^m. The first equation is a standard IS-relation (demand block), the second equation is the New Keynesian Phillips curve (supply block), and the third equation is the monetary policy rule (policy block). For most of my analysis, I do not rely on any specific assumptions on model parameterization; I only make the conventional assumptions $\beta \in (0, 1)$, $\kappa > 0$, $\phi_\pi > 1$, $\phi_y \geq 0$, and $\sigma^d, \sigma^s, \sigma^m > 0$.

It is straightforward to show that this benchmark model is static and admits the closed-form solution

$$\begin{pmatrix} y_t \\ \pi_t \\ i_t \end{pmatrix} = \frac{1}{1 + \phi_y + \phi_\pi \kappa} \begin{pmatrix} \sigma^d & \phi_\pi \sigma^s & -\sigma^m \\ \kappa \sigma^d & -(1 + \phi_y)\sigma^s & -\kappa \sigma^m \\ (\phi_y + \phi_\pi \kappa)\sigma^d & -\phi_\pi \sigma^s & \sigma^m \end{pmatrix} \begin{pmatrix} \varepsilon_t^d \\ \varepsilon_t^s \\ \varepsilon_t^m \end{pmatrix} \tag{1}$$

For the study of different SVAR estimators, I assume that the econometrician observes data on output, inflation, and the policy rate. However, she is not aware that the data are actually generated according to (1), and so does not exploit the structure of the model for inference.

The Smets-Wouters Model. The structural model of Smets & Wouters (2007) is perhaps the most well-known example of an empirically successful business-cycle model. For further details, I refer the reader to the original paper. In most of my analysis here, I consider their posterior mode parameterization; as a further robustness check, Section B.4 in the Online Appendix presents results taking into account posterior estimation uncertainty.[4] Exactly as before I assume that the econometrician observes data on aggregate output, inflation, and the interest rate, but does not know the true underlying model.

2.2 Structural models and VAR analysis

When solved through standard first-order perturbation techniques, my laboratories – as well as many other business-cycle models – give linear evolution equations for all model variables. Splitting variables into observables and unobservable states generates a linear state-space

[4] My implementation of the Smets-Wouters model is based on *Dynare* replication code kindly provided by Johannes Pfeifer. The code is available at `https://sites.google.com/site/pfeiferecon/dynare`.

model. As is conventional, I restrict attention to *Gaussian* linear state-space models:[5]

$$s_t = As_{t-1} + B\varepsilon_t \tag{2}$$

$$x_t = Cs_{t-1} + D\varepsilon_t \tag{3}$$

where s_t is a n_s-dimensional vector of state variables, x_t is an $n_x \times 1$ vector of observables and ε_t is an $n_\varepsilon \times 1$ vector of structural shocks. The disturbances ε_t are Gaussian white noise, with $\mathbb{E}[\varepsilon_t] = 0$, $\mathbb{E}[\varepsilon_t \varepsilon_t'] = I$ and $\mathbb{E}[\varepsilon_t \varepsilon_{t-j}'] = 0$ for $j \neq 0$.

Under weak conditions, the state-space system (2) - (3) implies a VAR representation for the observables x_t. As most material in this section is relatively standard, I only state the main results here, and refer the interested reader to Section B.2 of the Online Appendix and the literature referenced therein. The implied reduced-form VAR representation is[6]

$$x_t = \sum_{j=1}^{\infty} B_j x_{t-j} + u_t \tag{4}$$

where the coefficient matrices $B_j, j = 1, 2, \dots$ are complicated functions of the fundamental model matrices (A, B, C, D), and the u_t are the (Gaussian) forecast errors on observables x_t given information up to time $t-1$, with disturbance variance $\mathbb{E}(u_t u_t') \equiv \Sigma_u$.

The Computational Experiment. I assume that the econometrician observes macro aggregates x_t, but does not exploit the structure of the model – the matrices (A, B, C, D) – for inference. Since I allow her to observe an infinitely large sample generated from (2) - (3), I simply treat the reduced-form VAR representation (4) as known. Details on the computation of this VAR(∞) representation are presented in Section B.7 of the Online Appendix.

SVAR Identification. Structural VAR analysis posits that the true structural shocks ε_t can be obtained as a linear combination of contemporaneous reduced-form disturbances u_t. With e_t denoting the SVAR-identified "structural" shocks, a structural VAR representation of the same system is thus

$$A_0 x_t = \sum_{j=1}^{\infty} A_j x_{t-j} + e_t \tag{5}$$

[5]The Gaussianity assumption is made for notational simplicity only. Equivalently, I could restrict structural identification to only come from the second-moment properties of the data.

[6]Note that I use $\{B_j\}_{j=1}^{\infty}$ for VAR coefficient matrices and B for the shock impact matrix in the state equation (2). I do so to be as close as possible to textbook notation.

where $A_0^{-1}A_0^{-1'} = \Sigma_u$, $e_t \equiv A_0 u_t$ is Gaussian white noise with $\mathbb{E}[e_t] = 0$, $\mathbb{E}[e_t e_t'] = I$ and $\mathbb{E}[e_t e_{t-j}] = 0$ for $j \neq 0$, and $A_j \equiv A_0 B_j$.

As is well-known, a continuum of SVARs are consistent with a given reduced-form VAR representation. It is straightforward to see that, under the Gaussianity assumption, the SVAR (5) is identified up to orthogonal rotations – pre-multiplying both sides of (5) with a matrix Q in the space of n_x-dimensional orthogonal rotation matrices $O(n_x)$ does not change the likelihood of the model. In other words, SVARs are identified up to $\frac{n_x(n_x-1)}{2}$ restrictions (Rubio-Ramírez et al., 2010). Outside identifying information is then used to restrict attention to a strict subset of the set $O(n_x)$, often a singleton. In what follows, I will refer to this smaller set of SVARs as the *identified set*, and to outside identifying information as the *identification scheme*. A formal definition of identified sets is relegated to Appendix A.1. Given a model-implied reduced-form VAR representation and a structural identification scheme, it is straightforward to numerically characterize the (population) identified set of SVARs, as well as any corresponding impulse response functions, forecast error variance decompositions, or other objects of interest. Again, further computational details are provided in Section B.7 of the Online Appendix.

2.3 Interpreting SVAR estimands

The analysis in Section 2.2 did little to exploit the structure of the underlying model – *any* reduced-form VAR can be mapped into identified sets, both those estimated on actual data and those derived from fully-specified structural models. A controlled data-generating process does, however, have one important advantage: It allows us to interpret SVARs by linking their identified "structural" shocks, e_t, to the true shocks of the model, ε_t.

The nature of the link depends on the invertibility (or lack thereof) of the model (2) - (3). A linear state-space system is said to be invertible for the structural shocks ε_t if and only if, given knowledge of the system matrices (A, B, C, D), the infinite past of observables $\{x_{t-\ell}\}_{\ell=0}^{\infty}$ is sufficient to perfectly identify the hidden shocks ε_t.[7] A natural quantitative measure of invertibility is the R^2 in an (infeasible) regression of a structural shock $\varepsilon_{j,t}$ on current and past macro aggregates $\{x_\tau\}_{-\infty<\tau\leq t}$ (Plagborg-Møller & Wolf, 2019a).

[7] Strictly speaking, for SVAR analysis of a given shock $\varepsilon_{j,t}$ to work, the system needs to only be invertible for that shock. In that case, the static relation (6) will apply for shock j, while the richer dynamic relation (8) will apply to other shocks, with potentially non-zero weights for higher horizons $\ell > 0$.

THE INVERTIBLE CASE. Under invertibility – that is, if the R^2 is 1 for all structural shocks ε_t –, the link between SVAR shocks and true disturbances is very simple:

$$e_t \quad = \quad P \times \varepsilon_t \tag{6}$$

where $P \in O(n_x)$ is an orthogonal matrix. In short, with invertibility, the identified shocks are linear combinations of the true (contemporaneous) underlying structural shocks, with the weights given by the entries of the rotation matrix P. The right kind of SVAR identification scheme then identifies $P = I$ as the true rotation, with SVAR-identified structural shocks e_t equal to the true structural shocks ε_t.

It is straightforward to show that, with (y_t, π_t, i_t) observable, the three-equation model of Section 2.1 is invertible. In particular, writing out (6), we see that the identified set associated with any SVAR identification scheme is simply a collection of unit-length weight vectors $p_m = (p_{md}, p_{ms}, p_{mm})'$, with implied "monetary policy shocks"

$$e_t^m = p_{md} \times \varepsilon_t^d + p_{ms} \times \varepsilon_t^s + p_{mm} \times \varepsilon_t^m \tag{7}$$

Whether or not we are close to the ideal of $p_m = (0, 0, 1)'$ depends, of course, on the details of the chosen identification scheme.

THE NON-INVERTIBLE CASE. In the general non-invertible case, the link between identified and true structural shocks is more complicated. Following Fernández-Villaverde et al. (2007) and Lippi & Reichlin (1994):

$$e_t \quad = \quad P(L) \times \varepsilon_t \quad = \quad \sum_{\ell=0}^{\infty} P_\ell \times \varepsilon_{t-l} \tag{8}$$

where the entries of the matrix polynomial $P(L)$ are complicated functions of the chosen identification scheme as well as the fundamental model matrices (A, B, C, D). In a natural generalization of (6), the (k, j)th entry of P_ℓ is now the weight of kth identified SVAR shock on the ℓth lag of the jth true underlying structural disturbance.[8]

Standard small-scale VARs induced by the large Smets-Wouters model are not invertible, so the more complicated expression (8) applies. Two new results, developed in more detail in

[8] Since ε_t and e_t are both orthonormal white noise, we see immediately, following Lippi & Reichlin (1994), that $P(L)$ is a Blaschke matrix – that is, the matrix-polynomial generalization of an orthogonal matrix.

Appendix A.2, clarify when and how non-invertibility threatens SVAR-based identification of (monetary policy) shock transmission.

First, I tie the weights in the matrix polynomial $P(L)$ of (7) to the R^2 in a regression of shock j on current and lagged macro aggregates x_t.

Proposition 1. *Let the SVAR (5) be derived from a structural model (2) - (3). The weight of the kth identified SVAR shock on the jth contemporaneous structural shock, $P_0(k,j)$ is subject to the following upper bound:*

$$P_0(k,j) \leq \sqrt{R_j^2} \equiv \sqrt{1 - \text{Var}\left(\varepsilon_{j,t} \mid \{x_\tau\}_{-\infty<\tau\leq t}\right)} \tag{9}$$

Furthermore, there exists a SVAR, consistent with the model-implied reduced-form VAR (4), such that this upper bound is attained.

If R_j^2 is small, then – for *any possible* SVAR identification scheme – the identified shock $e_{k,t}$ will bear little relation to the true structural shock $\varepsilon_{j,t}$. Conversely, if R_j^2 is close to 1, then *some* SVAR identification scheme will (nearly) identify the true shock. Thus, in a precise sense, SVAR identification can work *if and only if* R_j^2 is sufficiently close to 1.

Second, I establish that, in the Smets-Wouters model, the R_m^2 for monetary policy shocks in a VAR in (y, π, i) is robustly close to 1. In this model, macro fluctuations are driven by seven distinct shocks; out of these, monetary policy shocks are among the least important, as measured by conventional forecast error variance decompositions. It thus seems *a priori* unlikely that a small trivariate VAR should contain much information about monetary policy shocks, casting doubt on the viability of SVAR inference. This simple intuition, however, turns out to be incorrect. At my benchmark parameterization, the R_m^2 is 0.8702, so the maximal attainable weight on the monetary policy shock is $\sqrt{0.8702} = 0.9328$.[9]

The intuition underlying this result is subtle: Monetary policy shocks are not important drivers of any *individual* macro aggregate, but they induce highly atypical *co-movement* patterns. Notably, monetary policy shocks are unique in that they push interest rates and inflation in opposite directions. Thus, while a divergence of interest rates and inflation is not definitive proof, it is at least suggestive of monetary policy shocks. I provide further details for this argument in Appendix A.2. In particular, the appendix discusses an instructive

[9] As I show in Section B.3 of the Online Appendix, this result is not sensitive to the assumption of infinite VAR lag lengths. For example, the R_m^2 is already equal to 0.8662 for a trivariate VAR with four lags. Furthermore, I also show that a high R_m^2 is not special to the model's posterior mode, but is in fact a feature of most draws from the estimated model posterior.

illustration using forward guidance shocks: In response to a credible promise of an interest rate hike *tomorrow*, interest rates and inflation move in the same direction *today*. Upon observing this co-movement, the econometrician initially concludes that the economy was almost surely hit by a conventional demand shock, and so the monetary policy R^2 is small. As soon as the promised rate hike materializes, however, inflation and policy rate diverge, the econometrician realizes that actually a forward guidance shock may have occurred, and the R^2 jumps back up.

Outlook. The results in this section establish that, for both model laboratories sketched in Section 2.1, SVAR-based inference can *in principle* succeed. Whether any *given* identification scheme succeeds is, of course, a different question. In the remainder of this paper, I will use the structural shock decompositions in (6) and (9) to evaluate and economically interpret the performance of several popular approaches to SVAR identification.

3 Sign restrictions and masquerading shocks

This section uses the controlled model laboratories of Section 2.1 to judge and economically interpret the popular agnostic sign identification scheme of Uhlig (2005). In Section 3.1, I characterize the entire SVAR-implied identified set; in particular, I study the largest and smallest output responses possibly consistent with the imposed sign-identifying information. Section 3.2 then analyzes the distribution *over* this identified set induced by the popular Bayesian implementation of sign restrictions – that is, the Haar prior.

3.1 The identified set

Uhlig (2005) proposes an agnostic identification scheme. He defines as a monetary policy shock any shock that, for a pre-specified (often quite large) number of periods, moves interest rates and inflation in opposite directions.[10] Notably, the response of output is left unrestricted, contrary to popular recursive schemes. In my candidate data-generating processes, monetary policy shocks indeed are the only shocks to satisfy these restrictions, so the proposed identification scheme is in principle promising.

[10]In his benchmark analysis, Uhlig considers a few additional constraints, designed chiefly to disentangle monetary policy and money demand shocks. As my candidate models feature no such shocks, I ignore these restrictions. Also, it is well-known that Uhlig's results continue to hold with my smaller set of restrictions on estimated three-variable SVARs (e.g Castelnuovo, 2012; Wolf, 2017).

Static Model. I begin with the simple three-equation model. Since the model is static, I only restrict the inflation and interest rate responses on impact. It is straightforward to see that the proposed sign restrictions are not strong enough to uniquely pin down the sign of the output response. I provide an informal discussion here, and relegate the formal proof to Section B.4.1 of the Online Appendix.

By construction, the monetary policy shock is the only *pure* shock to lie in the identified set. However, linear combinations of (expansionary) demand and supply shocks can do so as well and thus "masquerade" as contractionary policy shocks. By definition, any candidate "structural" shock $e_t^m \equiv p_{md}\varepsilon_t^d + p_{ms}\varepsilon_t^s + p_{mm}\varepsilon_t^m$, where the unit-length vector of weights $p_m = (p_{md}, p_{ms}, p_{mm})'$ is such that

$$p_{md} \times \kappa\sigma^d - p_{ms} \times (1+\phi_y)\sigma^s - p_{mm} \times \kappa\sigma^m \leq 0 \tag{10}$$

$$p_{md} \times (\phi_y + \phi_\pi\kappa)\sigma^d - p_{ms} \times \phi_\pi\sigma^s + p_{mm} \times \sigma^m \geq 0, \tag{11}$$

will lie in the identified set. The corresponding (scaled) output response is

$$p_{md} \times \sigma^d + p_{ms} \times \phi_\pi\sigma^s - p_{mm} \times \sigma^m \tag{12}$$

Clearly, as long as the impact impulse response matrix displayed in (1) is full rank, the two inequality restrictions cannot possibly be informative about the sign of the impact output response. In particular, straightforward algebra shows that, as long as $\phi_y > 0$, *positive* weights on demand and supply shocks are consistent with the imposed identifying restrictions, but with the obvious incorrect implications for the output response to the identified shock. It is important to note that this logic works completely independently of relative shock volatilities. In particular, even if the monetary shock were the overwhelming driver of macro fluctuations ($\sigma^m \gg \sigma^d, \sigma^s$), the sign-restricted identified set would continue to contain incorrect positive output responses. Thus, at least in this simple model, sign-identifying information *alone* can never be enough to pin down the sign of the unrestricted output response.

Smets-Wouters. The previous conclusions may appear particular to the simple model considered so far. Realistic data-generating processes are not static, and actual applications of sign-identifying schemes restrict impulse responses for many periods, not just one. I thus extend the inflation and interest rate restrictions to hold for six quarters, and apply them to identify structural VARs generated from the more realistic medium-scale DSGE model of Smets & Wouters (2007). Figure 1 displays identified sets of impulse response functions.

Identified Set of Impulse Responses: Uhlig (2005) Sign Restrictions

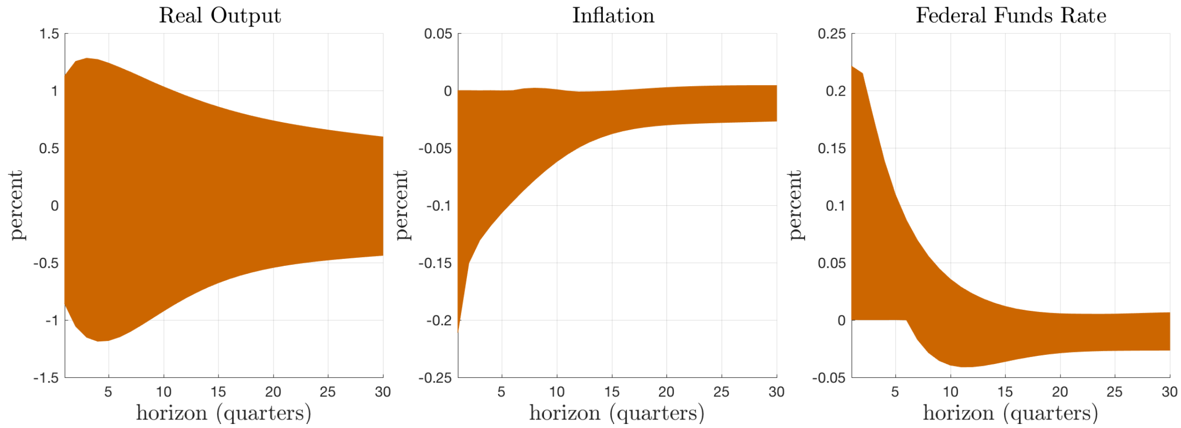

Figure 1: Identified sets of the responses of output, inflation and policy rate to a one standard deviation shock to the monetary policy rule, identified through sign restrictions on inflation and policy rate (imposed for six quarters).

Consistent with the intuition from the static model, and similar to the earlier simulation-based evidence of Castelnuovo (2012), I conclude that the impact output response is not well-identified; in particular, the identified set again contains both positive and negative values. To allow an economic interpretation of this identified set, I use the results of Section 2.3 to link the mis-identified SVAR shocks to the true underlying structural shocks. Figure 2 visualizes my results by matching the impact output response of identified shocks to their decomposition in terms of the underlying disturbances.[11]

Recall that the Smets-Wouters model features seven shocks; to ease visual interpretation, I have summed the weights on the three demand and supply shocks, respectively. The plot reveals that the right tail of positive output responses largely reflects positive demand and supply shocks *masquerading* as contractionary monetary policy shocks. The right linear combination of these shocks also pushes inflation down and interest rates up, but of course boosts output. Section B.4.5 of the Online Appendix shows that the exact same masquerading shocks logic features just as prominently in a dynamic three-equation model.

[11] To ease visual interpretability, I adjust raw shock weights in two ways. First, I only show impact weights, and ignore any weights on *lagged* true structural shocks. Appendix A.2 explains why this simplification is harmless. Second, there is in fact no strict one-to-one mapping between impact output responses and corresponding shock weights. I thus draw many entries from the model's identified set, and smooth the resulting series of shock weights as a function of the impact output response. I show a plot of unsmoothed sampled weights in Section B.4 of the Online Appendix. Finally, note that the weight vector $(0, 0, 1)$ does not lie in the identified set, precisely because the model is non-invertible.

Identified Set of Shock Weights: Uhlig (2005) Sign Restrictions

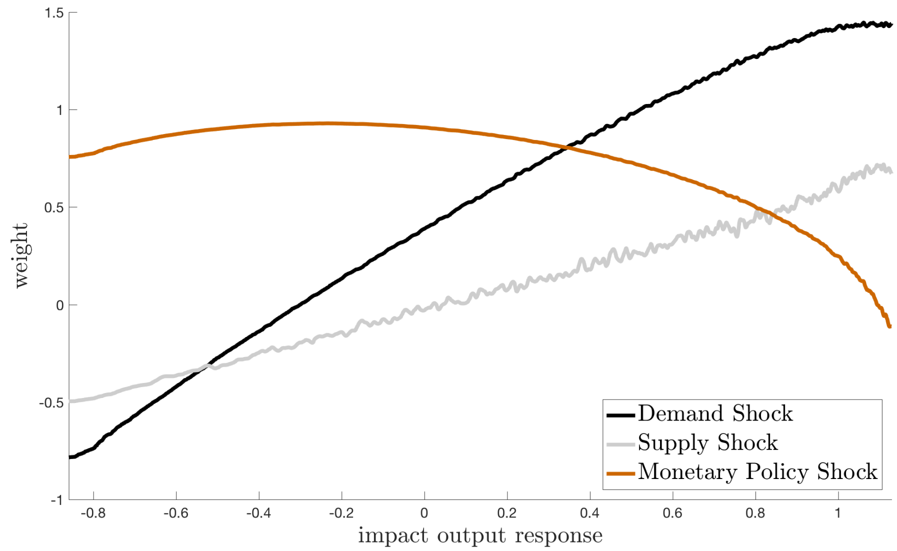

Figure 2: Identified set of static shock weights as a function of the output response at horizon 0. For the demand and supply shocks I sum all relevant weights, adjusting for the fact that the cost-push and wage-push shocks are negative supply shocks. I smooth the resulting series to facilitate interpretation. The true impact response of output is -0.219.

Implications. The analysis in this section has implications for sign-based SVAR inference in general and for the identification of monetary policy shocks in particular. First, both model laboratories suggest that the common minimal requirement of sign restrictions – that they be *exclusively* satisfied by the shock of interest – is necessary, but not sufficient for successful identification.[12] Monetary policy shocks are arguably unique in having opposite effects on inflation and interest rates (Uhlig, 2005), but, unfortunately, this is only enough to ensure that SVAR analysis can *in principle* succeed (cf. Section 2.3), not that weak sign restrictions alone give tight identified sets.

Second, the wide identified sets in Uhlig (2005) can be given an economic "masquerading shocks" interpretation. Through the lens of popular structural models, the large positive output responses in Uhlig's identified sets are readily explained as particular linear combinations of *positive* demand and supply shocks masquerading as contractionary policy shocks. I also showed that, at least in the static three-equation model, such contamination of identified

[12] Wolf (2017) studies the identification of technology shocks as a second illustration, and Section B.4.3 shows that even the simultaneous identification of multiple structural shocks does not safeguard against the masquerading threat. Kilian & Murphy (2012) arrive at a similar conclusion in the context of oil shocks.

sets will persist even when monetary policy shocks are counterfactually volatile. I further expand on this point in the next subsection.

3.2 The Haar prior

In addition to its width, a second defining feature of the identified set for output responses in Uhlig (2005) is that – at least under the Haar prior – most posterior mass is actually put on *positive* output responses (see Figure 7 in his paper). The model-based perspective taken here can also rationalize this finding and offer broader lessons for the role of the Haar prior in applied macro-econometrics with sign restrictions.

The Haar prior is a uniform prior over orthogonal rotation matrices $P \in O(n_x)$. Under invertibility, by (6), we can directly interpret the entries of these rotation matrices as weights on the underlying true structural shocks.[13] For example, in the static model of Section 2.1, the uniform Haar prior randomly draws shock weights p, spaced uniformly over the unit sphere. But if all shocks receive equal prior weight, yet some shocks have much larger effects on macro aggregates than others, then the prior distribution for impulse responses of these aggregates is automatically dominated by the most volatile shocks. In the remainder of this section, I show that this observation has two important implications. First, it can rationalize the substantial posterior mass on positive output responses observed in Uhlig (2005). Second, it clarifies that earlier results on the promise of sign restrictions for volatile shocks (Paustian, 2007; Canova & Paustian, 2011) are *exclusively* driven by the imposition of a probabilistic prior over the identified set, and not by the sign-identifying information itself.

Static Model. I again begin with an illustration in the simple static model, summarized in Figure 3. Panel (a) shows the top right part of the unit circle, corresponding to candidate "structural shocks" that assign positive weights to the true demand shock (x-axis) and the true supply shock (y-axis); the weight on the true monetary policy shock is implicitly assumed to be positive and then simply recovered residually (recall that the weight vector p must have unit length). The light grey region – the interior of the unit circle – is the set of all *possible* shock vectors with positive weights on true demand and supply shocks. The orange region shows, for a benchmark parameterization chosen to replicate the relative shock volatilities in Smets & Wouters (2007), combinations of those shock weights that (i) lie in the

[13]Formally, this result uses translation-invariance of the Haar prior, ensuring uniformity for any basis matrix $b(\Sigma_u)$ (see the discussion in Appendix A.1).

identified set and (ii) *increase* output – that is, the undesirable masquerading shocks. The dark grey region gives the analogous combinations of masquerading shocks for a different model parameterization, now with more volatile monetary policy shocks. Finally, panel (b) shows the posterior probability of a negative output response to identified monetary policy shocks (under the Haar prior) as a function of relative shock volatilities.

(a) Mis-Identification Regions (b) Probability of Negative Output Response

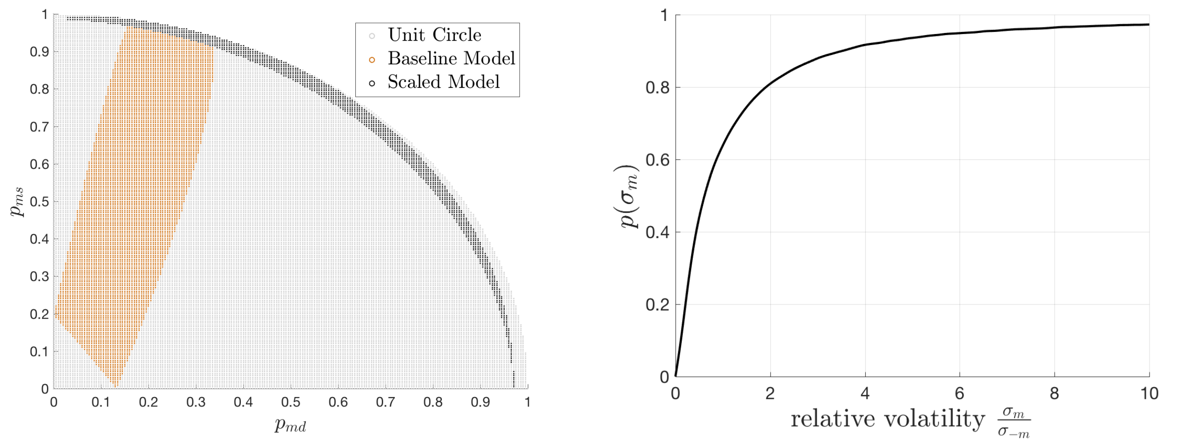

Figure 3: Identified sets in the static three-equation model, with $\phi_\pi = 1.5$, $\phi_y = 0.2$, $\kappa = 0.2$, $\sigma^d = 1$, $\sigma^s = 1$. In the benchmark calibration $\sigma^m = 0.2$; in the high-volatility calibration, $\sigma^m = 6$. Panel (a) shows regions of masquerading shocks giving positive output responses; panel (b) gives the probability of a negative identified output response as a function of relative shock volatilities.

Figure 3 illustrates the two main results of this section. First, in the baseline calibration, the orange region of "masquerading" demand and supply shocks features prominently in the top right part of the unit circle, and the posterior probability of correctly signing the output response is small. Intuitively, because demand and supply shocks are much more volatile than monetary policy shocks, very large weights on monetary policy shocks are needed to dominate the output response. Such large weights are unlikely according to the prior, so most posterior mass will instead be put on the large orange area of masquerading shock combinations. Exactly in line with this intuition, the posterior distribution over the identified set in Uhlig (2005) is dominated by positive output responses.

Second, as relative shock volatilities are re-scaled, the shape of the posterior over the identified set changes dramatically. Consider first the two identified sets of masquerading demand and supply shocks in panel (a), constructed for two different values of monetary policy shock volatility. From the inequality constraints (10) - (11), we know that there exists a simple one-to-one mapping between all points in these two identified sets.[14] Their posterior probabilities, however, are very different. In the benchmark parameterization, positive demand and supply shocks in the identified set occupy a large region in the unit circle, and so are regarded as likely by the Haar measure. As the monetary policy shock becomes more volatile, the associated weights on demand and supply shocks necessarily become larger – graphically, the orange area maps into a smaller and smaller sliver of the unit circle, and thus the masquerading shock combinations are regarded as increasingly unlikely. In the limit, as the monetary policy shock becomes infinitely more volatile than demand and supply shocks, the orange and dark grey regions actually get mapped into a measure-0 subspace at the boundary of the unit circle, and so receive a posterior probability of 0. Panel (b) provides an illustration across a large range of possible relative shock volatilities.

My results can also help to rationalize the conclusions in Paustian (2007) and Canova & Paustian (2011). If the shock of interest is sufficiently volatile, then conventional Bayesian posteriors over identified sets are likely to put most mass on correctly signed impulse responses. However, it is also immediate that this conclusion is *exclusively* driven by the particular choice of prior. As I show in Appendix A.3, it is always possible to construct an alternative prior such that, whatever the relative shock volatility, the posterior probability of a correctly signed impulse response remains arbitrarily small.[15]

SMETS-WOUTERS. The insights from the simple static model generalize without change to the environment of Smets & Wouters (2007). Figure 4 provides a graphical illustration.

Since monetary policy shocks are on average relatively small, most posterior mass over the identified set concentrates on positive output responses, fully consistent with the empirical findings in Uhlig (2005). As the relative volatility of the monetary policy shock increases, posterior mass mostly shifts to negative output responses. Nevertheless, even for an extreme counterfactual increase of monetary policy shock volatility, the identified set itself continues to include strictly positive output responses, so any conclusions about *statistical* significance

[14] Let p be a weight vector in the original identified set, giving a positive output response. Now let $\tilde{p}_i^* = p_i \times \frac{\sigma^i}{\tilde{\sigma}^i}$ ($i \in (d, s, m)$, and where σ and $\tilde{\sigma}$ are the old and new shock volatilities, respectively). Then the vector $\tilde{p} \equiv \frac{\tilde{p}^*}{||\tilde{p}^*||}$ lies in the identified set for the rescaled model, and also gives a positive output response.

[15] Finally, my results are also informative about the role played by the uniform Haar prior in allowing sign-restricted inference to be informative about *quantities*. For example, as the monetary policy shock becomes dominant relative to other shocks, the identified set for the output response converges to $[-\frac{1}{1+\phi_y+\phi_\pi\kappa}\sigma^m, 0]$, and the distribution over this identified set can be derived following the steps in Baumeister & Hamilton (2015). The quantity information contained in pure sign restrictions is thus simply that the impulse response is somewhere between zero and the truth; all further information comes from the prior.

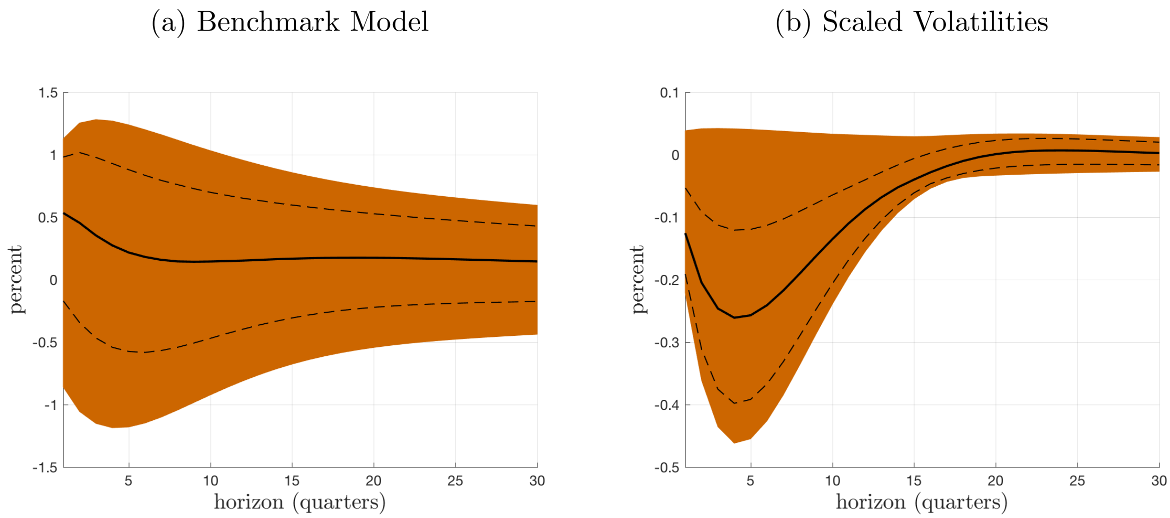

Figure 4: Identified set of the output response, with identifying restrictions as in Figure 1. Posterior uncertainty via imposition of the uniform Haar prior; the solid and dotted lines give 16th, 50th and 84th percentile bands. In the model with scaled volatilities, the relative volatility of the monetary policy shock is scaled up by a factor of 30.

of a negative output response are necessarily exclusively driven by the prior.

Implications. Large-sample Bayes inference over identified sets is dominated by the prior (Moon & Schorfheide, 2012; Baumeister & Hamilton, 2015; Watson, 2019). Taking a population perspective, my analysis precisely characterizes the additional probabilistic identifying information embedded in the popular Haar prior. In particular, I show that this flat prior over orthogonal rotation matrices can equivalently be interpreted as a flat prior over hidden shock weights, thus automatically over-weighting particularly volatile macro shocks.

Whether or not the Haar prior is a sensible prior is invariably an application-dependent question. In my analysis of monetary policy shock identification, I find that, due to the relatively low volatility of policy shocks, researchers relying on the Haar prior are likely to mis-characterize the sign of the aggregate output response. Since pure sign-identifying information is also consistent with (correct) negative output responses, any conclusions about statistical significance of positive responses are exclusively driven by the prior.[16]

[16] In actual empirical practice, such identification uncertainty is further conflated with reduced-form parameter estimation uncertainty. My analysis is exclusively concerned with population limits, and so only speaks to one part of the inference problem. Nevertheless, the larger Monte Carlo exercise in Section B.4.4 suggests that identification uncertainty is, in relevant applications, large relative to estimation uncertainty.

4 Zero restrictions

The classical approach to monetary policy shock identification is the imposition of zero impact response restrictions on output and inflation. Uhlig (2005) shows that the zero output restriction is central to recovering the conventional negative – if small – output effects of a contractionary monetary shock. In this section I provide a rationale for this centrality of the zero output restriction, but also show that, at least at short horizons, the estimated output response is likely to understate the policy's true aggregate effects.

THE IMPACT OUTPUT RESTRICTION. Uhlig (2005) shows that conventional wisdom (e.g. Christiano et al., 1996) relies sensitively on the impact zero output restriction – in his words a "rather spurious identification restriction." Expanding on the analysis of Section 3, I will now show that the strong bite of the impact output restriction is not an accident, but an intuitively sensible feature of identified sets.

In my structural models, the large positive output responses identified by the pure sign-restricting scheme of Uhlig (2005) correspond to large weights on positive demand and supply shocks. These shocks both push output up, but move interest rates in opposite directions, and so necessarily imply a large ratio (or multiplier) $|\frac{dy_0}{di_0}|$. Restricting this impact multiplier thus promises to chop off the right tail of mis-identified masquerading shocks displayed in Figure 2; as Figure 5 shows, this is exactly what happens in the model of Smets & Wouters. Panel (a) shows the identified set for the output response if the baseline sign restrictions of Uhlig (2005) were to be complemented with a hard zero restriction on the impact response of output. Consistent with the intuition given above, the identified set is now tight around the familiar hump-shaped negative response of output to an identified monetary policy shock. The plot of shock weights in panel (b) confirms that the large mis-identified region of masquerading expansionary demand and supply shocks is eliminated. Section B.4.6 of the Online Appendix further shows that, even with moderate *bounds* on the impact multiplier $|\frac{dy_0}{di_0}|$, identified sets tighten significantly around negative output responses.[17]

My analysis suggests that the centrality of a zero impact output restriction – or of weaker bounds – to the *sign* of the identified output response path should not come as a surprise. However, to the extent that the true impact output restriction is not literally zero, monetary policy shocks will still be mis-identified. In particular, for the first few quarters, real effects will mechanically be understated. Panel (a) in Figure 5 shows exactly this. At the same time,

[17] The same happens in actual data, as shown in Wolf (2017).

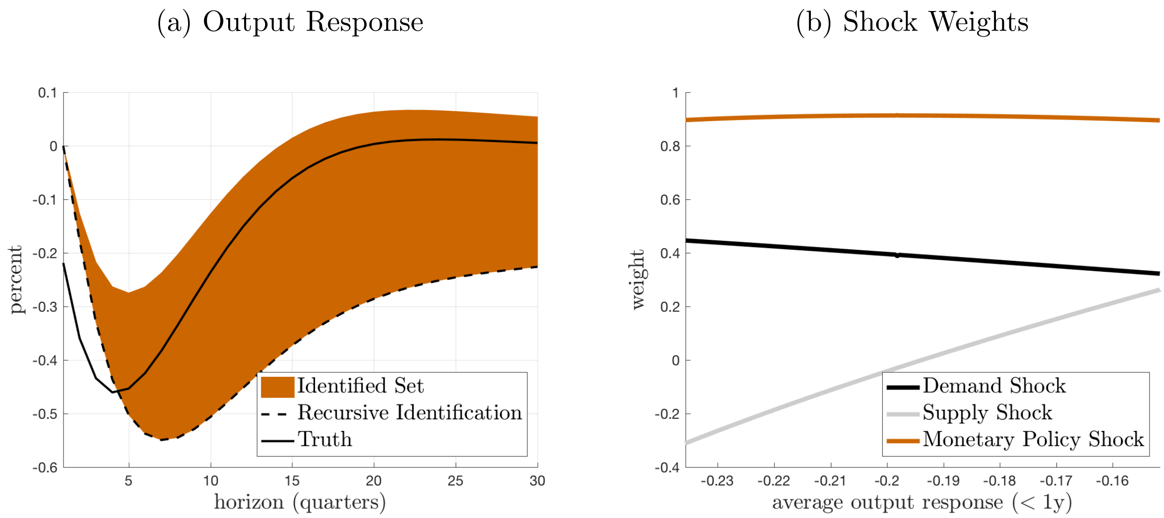

Figure 5: Identified sets of output and shock weights. Inflation and interest rates are restricted to move in opposite directions for six quarters; additionally, the impact output response is restricted to be 0. Panel (a) also shows the point-identified recursive impulse responses (with the policy shock ordered last) as well as the true impulse response. Panel (b) shows shock weights as a function of the average output response over the first year.

farther-out dynamics may be mis-identified in other less obvious ways; in the model of Smets & Wouters, the identified sets indicate greater persistence of policy shocks than is actually the case. As it turns out, the economics underlying this long-horizon mis-identification are particularly transparent for standard recursive identification schemes.

Recursive identification and shock persistence. Recursive identification of monetary policy shocks complements the impact zero output restriction with an impact zero inflation restriction. Together, these two pieces of identifying information are enough to provide point identification. The corresponding output impulse response is also displayed in panel (a) of Figure 5.[18] Recursively identified monetary policy shocks appear to depress output, but, relative to the true model-implied impulse response, real effects are understated at short horizons and overstated at long horizons. In fact, recursive identification distorts long-horizon impulse responses more than any other SVAR in the identified set of Figure 5.

As before, understatement at short horizons is simply a mechanical implication of the

[18] Since this paper is chiefly concerned with the real effects of monetary policy shocks, I do not further discuss the "price puzzle" – another well-known anomaly of recursively identified monetary policy SVARs. In Section B.6 I study the identified inflation response and discuss the extent to which the model-based perspective taken here can also rationalize the price puzzle.

impact zero restriction. The subsequent pattern of dynamic mis-identification is more subtle and intimately related to the relative persistence of the underlying structural shocks. In Section B.5 of the Online Appendix I show that, if all true model shocks had equally persistent effects on macro aggregates, then the recursively identified "monetary policy" impulse responses for output and inflation would be exactly 0 at all times. Intuitively, if a given linear combination of shocks – all with equally persistent *dynamic* effects – implies a zero response on impact, then it will necessarily imply a zero response forever. This simple logic can help clarify the dynamics displayed in Figure 5: Relative to all other SVARs in that identified set, a recursive SVAR gives the largest possible (i.e., zero) inflation impact response, and it achieves this zero impact response through a large positive weight on contractionary supply shocks. Crucially, in the structural model of Smets & Wouters, technology shocks – which account for most low-frequency variation in macroeconomic aggregates – are extremely persistent. These persistent supply shocks then dominate long-run dynamics, and in particular result in the displayed substantial overstatement of the output drop at long horizons.

Implications. The theory presented here rationalizes the centrality of zero impact output restrictions to conventional wisdom, but cautions against interpretation of the resulting estimates as accurate representations of the economy's true shock propagation. Recursively identified shocks are likely to *understate* the true aggregate effects of policy interventions at short horizons, and their implied long-horizon dynamics are sensitive to the persistence of the various other underlying macro shocks.[19] The next section considers alternative identification schemes that are less vulnerable to such criticisms.

5 Recent advances in identification

Following the concerns expressed in Uhlig (2005), the past few years have seen a flurry of research trying to identify the real effects of monetary policy without any direct restrictions on the response of output. Two particularly prominent examples are sign restrictions on the VAR-implied Taylor rule, as in Arias et al. (2019), and the use of external instruments, as in Gertler & Karadi (2015) or Jarociński & Karadi (2018). Most of these methods indicate somewhat larger effects of policy shocks on real outcomes, in particular at short horizons. In this section, I argue that, first, these results are again not at all surprising through a model

[19] Of course, these concerns would be less acute in models like Christiano et al. (2005), which have been explicitly constructed to ensure consistency of the usual recursive estimators.

lens, and second, the resulting identified sets are likely to be quite informative about the true real effects of monetary policy disturbances.

5.1 Taylor rule restrictions

Arias et al. (2019) show that restrictions on the output coefficient in an implied Taylor rule substantially tighten Uhlig's identified set around negative effects of monetary policy shocks; equivalently, their analysis suggests that many of the candidate "monetary policy shocks" in Uhlig (2005) must imply Taylor rules with a negative output response. This section makes two observations. First, I show that the long tail of masquerading supply and demand shocks in my structural models also induces implied Taylor rules with negative output coefficients. It is thus not surprising that, just like in the data, an additional restriction on implied Taylor rule coefficients materially tightens identified sets in my models. Second, I show that this conclusion is a robust implication of basic properties of New Keynesian models.

Figure 6 displays the identified set under the identification scheme of Arias et al. (2019). Building on the baseline sign restrictions of Uhlig (2005) – but then additionally imposing that the output and inflation coefficients in the SVAR-implied Taylor rule are strictly positive – leads to a substantial tightening of the identified set around significant negative output responses. Exactly as in Arias et al. (2019), I find that this tightening is almost exclusively driven by the restriction on the implied Taylor rule output coefficient.[20]

The results of Figure 6 indicate that restrictions on implied Taylor rule coefficients contain substantial additional identifying information. The intuition is as follows. As shown in Figure 2, most of the mis-identified "masquerading" shock combinations that counterfactually increase aggregate output are in fact mixtures of positive demand and supply shocks. Equivalently, the mis-identified monetary policy shocks are linear combinations of residuals in the model's IS and NKPC curves. But if the shocks are a linear combination of these residuals, then the implied Taylor rule itself is a linear combination of those same IS and NKPC equations. For example, in the static three-equation model of Section 2.1, straightforward manipulations show that the implied (mis-identified) Taylor rule is

$$i_t \quad = \quad \underbrace{\frac{p_{mm}\phi_\pi + p_{ms}}{p_{md} + p_{mm}}}_{\tilde{\phi}_\pi} \times \pi_t \quad + \quad \underbrace{\frac{p_{mm}\phi_y - p_{md} - p_{ms}\kappa}{p_{md} + p_{mm}}}_{\tilde{\phi}_y} \times y_t \quad + \quad e_t^m$$

[20]Differently from their analysis, and consistent with the discussion in Section 3.2, I do not impose the uniform Haar prior, but instead display entire identified sets.

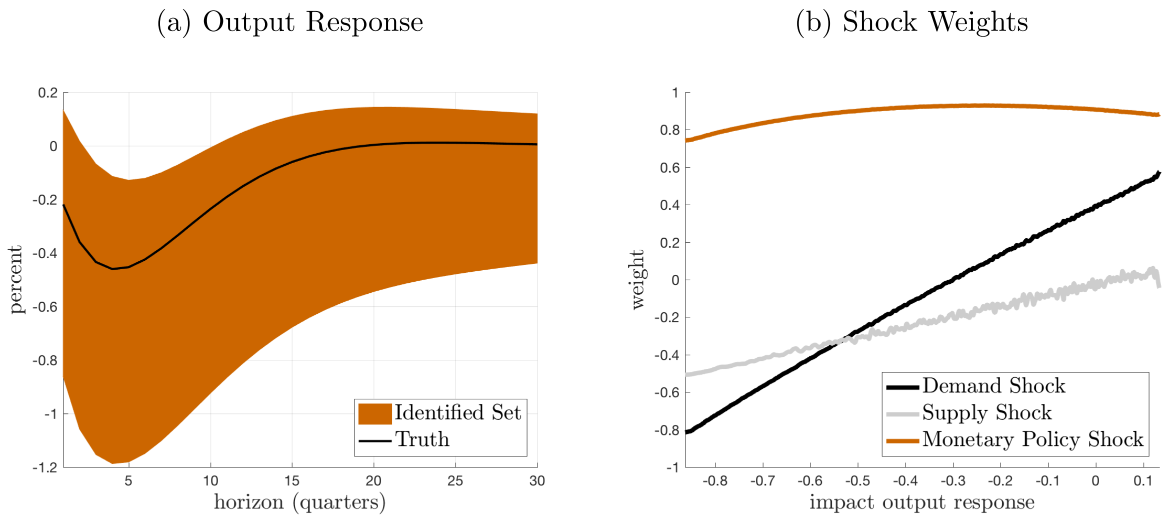

Figure 6: Identified sets of output and shock weights. Inflation and interest rates are restricted to move in opposite directions for six quarters; additionally, the implied Taylor rule coefficients on inflation and output are restricted to be positive.

Importantly, for mis-identified masquerading shocks with $p_{mm} \approx 0$ and $p_{md}, p_{ms} > 0$, the SVAR-implied Taylor rule coefficient $\tilde{\phi}_y$ is necessarily negative. It is thus unsurprising that the additional restriction $\tilde{\phi}_y > 0$ substantially tightens identified sets and largely removes masquerading supply and demand shocks.

5.2 External instruments

A popular alternative to the use of direct identifying restrictions – "internal instruments," in the language of Stock & Watson (2017) – is the use of instrumental variables, or "external instruments." An external instrument is a variable correlated with the shock of interest, and uncorrelated with any other structural shocks. For the study of monetary policy shocks, the most popular instrument is that of Gertler & Karadi (2015).

In this section I study the performance of the popular SVAR-IV estimator (Stock, 2008; Stock & Watson, 2012; Mertens & Ravn, 2013) in the structural model of Smets & Wouters (2007). Specifically, I assume that, in addition to the usual macro aggregates x_t, the econometrician now also observes an artificially generated *external instrument* z_t satisfying

$$z_t = \sum_{\ell=1}^{\infty} (\Psi_\ell z_{t-\ell} + \Lambda_\ell x_{t-\ell}) + \alpha \varepsilon_{m,t} + \sigma_v v_t, \tag{13}$$

where (i) all roots of the lag polynomial $1 - \sum_{\ell=1}^{\infty} \Psi_\ell L^\ell$ are outside the unit circle, (ii) $\{\Lambda_\ell\}_\ell$

is absolutely summable, and (iii) v_t is uncorrelated at all leads and lags with the structural shocks ε_t. The econometrician then implements the SVAR-IV estimator using the linear projection $\tilde{z}_t \equiv z_t - \mathbb{E}\left[z_t \mid \{z_\tau, x_\tau\}_{-\infty<\tau<t}\right] = \alpha\varepsilon_{m,t} + \sigma_v v_t$ as an external instrument.[21]

Even with a valid instrument, however, non-invertibility can threaten the consistency of the SVAR-IV estimator. In Plagborg-Møller & Wolf (2019a), we prove two related results. First, we show that, under non-invertibility, the weight of identified on true contemporaneous monetary policy shock is

$$P_0(k,m) = \sqrt{R_m^2} = \sqrt{1 - \text{Var}\left(\varepsilon_{m,t} \mid \{x_\tau\}_{-\infty<\tau\leq t}\right)} \tag{14}$$

Thus, the SVAR-IV estimator attains the theoretical bound in (9), and so – in a very particular sense – provides the best possible approximation to the true unknown monetary policy shock. Of course, with a low R_m^2, this approximation could still be quite poor. Second, we partially characterize the resulting bias. In particular, we show that impact impulse response estimates are biased up (in absolute value) by a factor of $1/\sqrt{R_m^2}$.

Taken together, these theoretical results as well as my earlier conclusions about likely near-invertibility of monetary policy shocks imply that SVAR-IV estimators of monetary policy transmission are likely to perform reasonably well. Figure 7 shows that this is exactly what happens in the model of Smets & Wouters (2007). Even with only three macro observables, the impact bias is small, and the dynamics of output and inflation are captured adequately, at least at short horizons. If anything, the real effects of policy shocks are somewhat overstated. With an augmented set of observables, identification obviously improves further; for example, if the researcher were to additionally include investment and a measure of total labor, then the $R_{0,m}^2$ would rise to 0.9302, and the weight on the true shock would increase to 0.9645. If the researcher were to go even further and include measures of consumption and real wages, then the system becomes invertible and identification is perfect.

6 Conclusion

In this paper I interpret various different *empirical* approaches to the study of monetary policy transmission through the lens of fully specified *structural models*. This model-based perspective suggests two important conclusions. First, theory and empirics are internally

[21] Note that the probability limit of the SVAR-IV estimator is independent of the particular numerical values of α and σ_v (as long as $\alpha \neq 0$). I thus do not need to take a stand on what those values actually are.

Identified Set of Impulse Responses

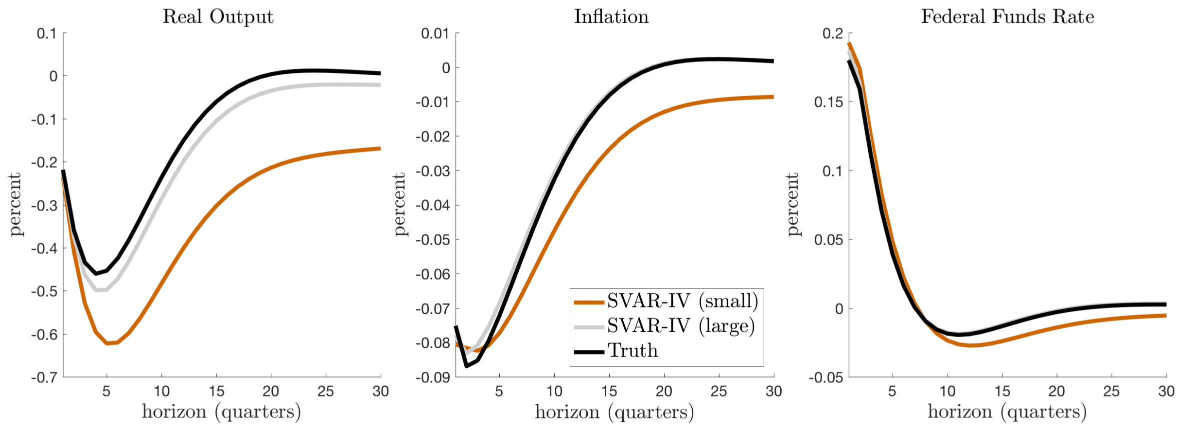

Figure 7: Impulse response functions identified via valid external instruments. The small-scale VAR contains output, inflation and the interest rate; the large-scale VAR adds investment and hours worked.

consistent. I find that different estimators, all applied to the same standard structural model, can give estimates of impulse responses that look as disparate as those estimated on actual data. I conclude that the data are consistent with monetary policy having significant real effects, and in fact somewhat larger and somewhat less persistent than often estimated. Second, pure sign restrictions are quite weak identifying information. Because of what I call masquerading shocks, the common minimal requirement for sign-based inference – that the shock of interest be the only one to simultaneously satisfy all imposed sign restrictions – is not sufficient. The masquerading shock problem is particularly acute when the shock of interest is not very volatile, as then the uniform Haar prior will concentrate most posterior mass on rival large masquerading shocks.

The identification of monetary policy transmission can, of course, be improved further. A valid external instrument is clearly the ideal solution, implemented either using LP-IV or SVAR-IV methods (Stock & Watson, 2017; Plagborg-Møller & Wolf, 2019b). Existing high-frequency instruments, however, may fail to adequately disentangle true policy shocks and information effects (Jarociński & Karadi, 2018; Nakamura & Steinsson, 2018a). Alternatively, model-consistent set-identifying information in the spirit of Uhlig (2005) and Arias et al. (2019) promises to be robust and, in the latter case, informative across a wide range of structural models, but may not yield tight enough inference. Further identifying restrictions may thus be needed.

References

Arias, J. E., Caldara, D., & Rubio-Ramirez, J. F. (2019). The systematic component of monetary policy in SVARs: An agnostic identification procedure. *Journal of Monetary Economics*, *101*, 1–13.

Baumeister, C. & Hamilton, J. D. (2015). Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica*, *83*(5), 1963–1999.

Brunnermeier, M., Palia, D., Sastry, K. A., & Sims, C. A. (2017). Feedbacks: Financial Markets and Economic Activity. Manuscript, Princeton University.

Caldara, D. & Kamps, C. (2017). The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. *Review of Economic Studies*, *84*(3), 1015–1040.

Canova, F. & De Nicolo, G. (2002). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, *49*(6), 1131–1159.

Canova, F. & Paustian, M. (2011). Business cycle measurement with some theory. *Journal of Monetary Economics*, *58*(4), 345–361.

Castelnuovo, E. (2012). Monetary Policy Neutrality? Sign Restrictions Go to Monte Carlo. Working Paper.

Castelnuovo, E. & Surico, P. (2010). Monetary policy, inflation expectations and the price puzzle. *The Economic Journal*, *120*(549), 1262–1283.

Christiano, L., Eichenbaum, M., & Evans, C. (1996). The effects of monetary policy shocks: evidence from the flow of funds. *Review of Economics and Statistics*, *78*, 16–34.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, *113*(1), 1–45.

Coibion, O. (2012). Are the Effects of Monetary Policy Shocks Big or Small? *American Economic Journal: Macroeconomics*, *4*(2), 1–32.

Faust, J. (1998). The Robustness of Identified VAR Conclusions about Money. *Carnegie-Rochester Conference Series on Public Policy*, *49*, 207–244.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J., & Watson, M. W. (2007). ABCs (and Ds) of Understanding VARs. *American Economic Review*, *97*(3), 1021–1026.

Forni, M., Gambetti, L., & Sala, L. (2019). Structural VARs and noninvertible macroeconomic models. *Journal of Applied Econometrics*, *34*(2), 221–246.

Gafarov, B., Meier, M., & Montiel Olea, J. L. (2017). Delta-Method Inference for a Class of Set-Identified SVARs. Manuscript, Columbia University.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications*. Princeton University Press.

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

Giacomini, R. (2013). The relationship between DSGE and VAR models. Cemmap Working Paper CWP21/13.

Giacomini, R. & Kitagawa, T. (2016). Robust inference about partially identified SVARs. Working Paper.

Jarociński, M. & Karadi, P. (2018). Deconstructing monetary policy surprises: the role of information shocks. ECB Working Paper.

Kilian, L. & Murphy, D. P. (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.

Leeper, E. M. & Leith, C. (2016). Understanding inflation as a joint monetary–fiscal phenomenon. In *Handbook of Macroeconomics*, volume 2 (pp. 2305–2415). Elsevier.

Lippi, M. & Reichlin, L. (1994). VAR analysis, nonfundamental representations, Blaschke matrices. *Journal of Econometrics*, *63*(1), 307–325.

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

Mertens, K. & Ravn, M. O. (2014). A reconciliation of SVAR and narrative estimates of tax multipliers. *Journal of Monetary Economics*, *68*, S1–S19.

Moon, H. R. & Schorfheide, F. (2012). Bayesian and frequentist inference in partially identified models. *Econometrica*, *80*(2), 755–782.

Nakamura, E. & Steinsson, J. (2018a). High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics*, *133*(3), 1283–1330.

Nakamura, E. & Steinsson, J. (2018b). Identification in macroeconomics. *Journal of Economic Perspectives*, *32*(3), 59–86.

Paustian, M. (2007). Assessing sign restrictions. *The BE Journal of Macroeconomics*, *7*(1).

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

Plagborg-Møller, M. & Wolf, C. K. (2019a). Instrumental Variable Identification of Dynamic Variance Decompositions. Manuscript, Princeton University.

Plagborg-Møller, M. & Wolf, C. K. (2019b). Local projections and VARs estimate the same impulse responses. Working Paper.

Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of Macroeconomics*, volume 2 chapter 2, (pp. 71–162). Elsevier.

Rubio-Ramírez, J. F., Waggoner, D. F., & Zha, T. (2010). Structural vector autoregressions: Theory of identification and algorithms for inference. *Review of Economic Studies*, *77*(2), 665–696.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*(1), 1–48.

Smets, F. & Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, *97*(3), 586–606.

Stock, J. H. (2008). What's New in Econometrics: Time Series, Lecture 7. Lecture slides, NBER Summer Institute.

Stock, J. H. & Watson, M. W. (2012). Disentangling the Channels of the 2007–09 Recession. *Brookings Papers on Economic Activity*, *2012*(1), 81–135.

Stock, J. H. & Watson, M. W. (2017). Identification and Estimation of Dynamic Causal Effects in Macroeconomics. Manuscript based on 2017 Sargan Lecture, Royal Economic Society.

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

Uhlig, H. (2017). Shocks, Sign Restrictions, and Identification. In *Advances in Economics and Econometrics, Chapter 4*, volume Eleventh World Congress (pp. 95–127). Cambridge University Press.

Watson, M. W. (2019). Comment on "The Empirical (Ir)Relevance ofthe Zero Lower Bound Constraint". In M. S. Eichenbaum, E. Hurst, & J. A. Parker (Eds.), *NBER Macroeconomics Annual 2019, volume 34*. University of Chicago Press.

Wolf, C. K. (2017). Masquerading Shocks in Sign-Restricted VARs. Manuscript, Princeton University.

Woodford, M. (2003). *Interest and Prices.* Princeton University Press.

A Appendix

A.1 Identified sets

Throughout the paper I informally refer to identified sets of SVARs and impulse response functions. A proper definition of identified sets requires a formal treatment of identifying information. Following Rubio-Ramírez et al. (2010), I allow identifying information to take the form of linear restrictions on transformations of the structural parameter space into $q \times n_x$ matrices, where $q > 0$. Denote the transformation by $f(\cdot)$. Linear restrictions on the transformation can then be represented via $q \times q$ matrices Z_j and S_j, with $j = 1, \ldots, n_x$. Here the Z_j allow us to impose exact linear restrictions on $f(\cdot)$ through the requirement $Z_j f(\cdot) e_j = 0$, and the S_j allow us to impose linear sign restrictions through the requirement $S_j f(\cdot) e_j \geq 0$. In the SVAR literature, most identifying restrictions take the form of restrictions on impulse response functions. Formally, the impulse response of variable i to shock j at horizon h is defined recursively as the (i, j)th element of the matrix

$$\text{IRF}_h = \begin{cases} A_0^{-1} & \text{if } h = 0 \\ \sum_{\ell=1}^{h} A_0^{-1} A_\ell \text{IRF}_{h-\ell} & \text{if } h > 0 \end{cases}$$

Transformation functions $f(\cdot)$ are then typically of the following form:

$$f(\{A_j\}) = \begin{pmatrix} \ldots \\ IRF_h \\ \ldots \end{pmatrix}$$

where $\{A_j\}$ collects the structural VAR matrices $(A_0, A_1, \ldots)$. The matrices Z_j and S_j simply consist of 0's, 1's and -1's, placed so as to ensure that the desired zero and sign restrictions are imposed. Simple examples of such matrices are provided, for example, in Rubio-Ramírez et al. (2010) or Arias et al. (2019). Finally, covariance restrictions and outside information are appended by sign normalizations on A_0: We require the diagonal elements of A_0 to be non-negative, which just means that a unit positive change in the ith structural shock is interpreted as a one standard deviation positive innovation to the ith variable in the VAR.

Definition 1. *Consider the reduced-form VAR $(\{B_j\}, \Sigma_u)$, and let $b(\Sigma_u)$ denote an invertible $n_x \times n_x$ matrix such that $b(\Sigma_u) b(\Sigma_u)' = \Sigma_u$. Let R be an identifying restriction, defined by a transformation function $f(\cdot)$ and a set of restriction matrices Z_j, S_j, $j = 1, \ldots, n$. Then the* *identified set of rotation matrices with respect to the basis* $b(\Sigma_u)$*,* $\mathcal{Q}_R^b$*, is defined as the set of orthogonal rotation matrices* $Q \in O(n_x)$ *such that* $A_0 \equiv Q \times b(\Sigma_u)^{-1}$ *is consistent with the reduced-form covariance matrix, the normalization rule and the identifying restriction* R*:*

$$\mathcal{Q}_R^b \equiv \{Q \mid Q \in O(n_x),\ A_0 \equiv Qb(\Sigma_u)^{-1},\ A_0^{-1}A_0^{-1'} = \Sigma_u,\ \mathrm{diag}(A_0) > 0,\ \textit{and} \\ Z_j f(\{A_j\})e_j = 0,\ \ S_j f(A)e_j \geq 0 \textit{ for } 1 \leq j \leq n_x\}$$

Identified sets can be empty, have a single member, or they can have multiple members. Generically, sign restrictions are set-identifying, and the identified set of rotation matrices has strictly positive Haar measure (independent of basis). It is now straightforward to define the identified set of impulse responses:[22]

Definition 2. *Consider the reduced-form VAR* $(\{B_j\}, \Sigma_u)$*. Let* R *be an identifying restriction, defined by a transformation* $f(\cdot)$ *and a set of restriction matrices* Z_j, S_j*,* $j = 1, \ldots, n_x$*. Then the identified set of impulse responses for variable* i *in response to shock* j *at horizon* h*,* $IS_{i,j,h}$*, is defined as the set of impulse responses generated by some rotation matrix* Q *in the identified set of SVARs:*

$$IS_{i,j,h} \equiv \{a \in \mathbb{R} \mid a = IRF_{i,j,h}(\{A_j\}(\{B_j\}, \Sigma_u, Q)),\ Q \in \mathcal{Q}_R^b\}$$

The upper bound of the identified set of impulse responses for variable i *in response to shock* j *at horizon* h *is defined as follows:*

$$\overline{IS}_{i,j,h} \equiv \sup_{Q \in \mathcal{Q}_R^b} IRF_{i,j,h}(\{A_j\}(\{B_j\}, \Sigma_u, Q))$$

The lower bound $\underline{IS}_{i,j,h}$ *is defined analogously.*

The bounds can be obtained using the closed-form expressions provided in Gafarov et al. (2017). Strictly speaking, an exclusive focus on bounds is, of course, only justified if the identified set is convex. In almost all applications considered in this paper, this is easy to establish using Lemma 5.1 in Giacomini & Kitagawa (2016). The sole exception is the case of multiple simultaneously identified structural shocks, considered in Section B.4.3; even in that case, numerical explorations suggest convexity of the identified set.

[22] It is trivial to show that identified set and in particular boundaries are independent of the chosen basis $b(\Sigma_u)$. No further reference to the basis is thus needed.

A.2 Supplementary results on invertibility

This appendix complements the discussion of Section 2.3. First, I prove Proposition 1. Second, I discuss in more detail why the R_m^2 in the structural model of Smets & Wouters (2007) is high. And third I conclude that, by near-invertibility, we can, without much loss of generality, restrict attention to the first entry (P_0) of the matrix polynomial $P(L)$.

Shock Weights & Non-Invertibility. I will prove a slightly generalized version of Proposition 1, in fact asserting that

$$|P_\ell(k,j)| \leq \sqrt{R_{\ell,j}^2 - R_{\ell-1,j}^2}$$

where $R_{\ell,j}^2 = 1 - \text{Var}\left(\varepsilon_{j,t} \mid \{x_\tau\}_{-\infty<\tau\leq t+\ell}\right)$. At $\ell = 0$, this statement is identical to that of Proposition 1 (since $R_{-1,j}^2$ is trivially 0).

Proof. By Lemma B.1 and Lemma B.2

$$\begin{aligned} R_{0,j}^2 &= \text{Var}\left(\mathbb{E}\left(\varepsilon_{j,t} \mid \{x_\tau\}_{-\infty<\tau\leq t}\right)\right) \\ &= \text{Cov}\left(u_t, \varepsilon_{j,t}\right)' \Sigma_u^{-1} \text{Cov}\left(u_t, \varepsilon_{j,t}\right) \\ &= M_{\bullet,j,0}' \Sigma_u^{-1} M_{\bullet,j,0} \end{aligned}$$

Similarly, it can be established that

$$R_{\ell,j}^2 = R_{\ell-1,j}^2 + M_{\bullet,j,\ell}' \Sigma_u^{-1} M_{\bullet,j,\ell} \quad \forall \ell \geq 2$$

Next, we know from Lemma B.1 and the definition of the VAR structural shocks e_t that

$$e_t = Q \times \Sigma_u^{-1/2} \times \sum_{\ell=0}^{\infty} M_\ell \varepsilon_{t-\ell} \equiv \sum_{\ell=0}^{\infty} P_\ell \varepsilon_{t-\ell}$$

where $Q \in O(n_x)$ is an orthogonal matrix. Summing over all identified shocks k at some given horizon ℓ, the total squared weights on shock j are given as the (j,j)th element of

$$M_\ell' \Sigma_u^{-1} M_\ell$$

But this is just $R_{\ell,j}^2 - R_{\ell-1,j}^2$. Any *individual* squared weight is thus bounded above by the difference in R^2's.

To assign this maximal weight to a single identified shock, it suffices to have the jth column of P_0 be proportional to the standard basis vector e_j. We can ensure this by setting q_j, the jth row of Q, proportional to $M'_{\bullet,j,0}\Sigma_u^{-1/2}$ (with normalization to ensure unit length), and the other $n_x - 1$ rows orthogonal to q_j and to each other, again with unit length. □

By linking shock weights to a quantitative measure of the degree of invertibility, Proposition 1 formalizes the notion that SVAR inference can succeed if and only if the $R^2_{0,j}$ is sufficiently close to 1.

Near-Invertibility in Smets & Wouters (2007). I find that, in a trivariate VAR in (y_t, π_t, i_t) induced by the structural model of Smets & Wouters (2007), the $R^2_{0,m}$ for monetary policy shocks is robustly close to 1. It is equal to 0.8702 at my benchmark parameterization, and it remains high for most draws from the model's posterior, as well as for much shorter VAR(p) representations (see Section B.3 in the Online Appendix).

To understand why the $R^2_{0,m}$ is so robustly high, it will prove useful to re-write it in a slightly more interpretable fashion. Following Forni et al. (2019) and Plagborg-Møller & Wolf (2019a), we can equivalently re-write the $R^2_{0,j}$ as follows:

$$R^2_{0,j} = \text{Cov}(u_t, \varepsilon_{j,t})' \times \text{Var}(u_t)^{-1} \times \text{Cov}(u_t, \varepsilon_{j,t}) \tag{A.1}$$

In words, the $R^2_{0,j}$ is large if and only if the true structural shock $\varepsilon_{j,t}$ is responsible for a lot of the variation in the reduced-form VAR forecasting error u_t made by the econometrician. Equivalently, the system is invertible for shock $\varepsilon_{j,t}$ if and only if

$$\text{Var}(u_t) - \underbrace{\text{Cov}(u_t, \varepsilon_{j,t})\,\text{Cov}(u_t, \varepsilon_{j,t})'}_{\equiv \text{Var}_j(u_t)} \tag{A.2}$$

is reduced-rank. To interpret (A.1) and (A.2), suppose first that the econometrician only observes a single macro aggregate. In that case, the $R^2_{0,j}$ for a shock j would be high if and only if that shock accounts for almost all forecasting uncertainty in that single aggregate x_t. Formally, if shock j drives all of the forecasting uncertainty in x_t, then the unconditional variance $\text{Var}(u_t)$ is equal to the variance induced *only* by shock j, $\text{Var}_j(u_t)$. Thus, by (A.2), we have invertibility. Of course, since monetary policy shocks are arguably unimportant for the evolution of most macro aggregates, the $R^2_{0,m}$ in univariate ARs is invariably small.

Matters are much more subtle in the multivariate case, however. Here, macro data can be very (in fact even fully) informative about a quite unimportant shock, if that shock induces an atypical covariance pattern, in the sense that the induced *conditional* variance $\text{Var}_j(u_t)$ is far from being proportional to $\text{Var}(u_t)$. For example, if *unconditionally* two entries in u_t co-move, but *conditional* on shock j they diverge, then intuitively the data should be quite informative about j; formally, the residual covariance matrix should become nearly singular.

This logic explains the robustly high $R^2_{0,m}$ in Smets & Wouters (2007). If the econometrician observes a negative forecast error in nominal interest rates and a positive forecast error in inflation, then – because only monetary policy shocks induce such conditional co-movement – she concludes that a contractionary monetary policy shock is reasonably likely to have occurred. Consistent with this intuition, the $R^2_{0,m}$ is already large in a bivariate VAR in (π_t, i_t) (it is equal to 0.7800, so the maximal weight is 0.8832), but remains stubbornly low in VARs that omit either macro aggregate, in particular the policy rate.[23]

Another way to see this logic is to study monetary policy R^2's in a variant of the Smets-Wouters model with forward guidance shocks, as done in Plagborg-Møller & Wolf (2019a).

R^2 for Forward Guidance Shock

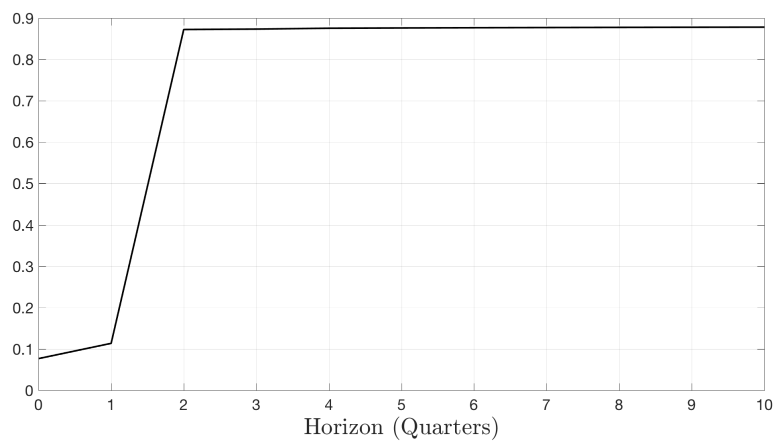

Figure A.1: I take the benchmark model of Smets & Wouters (2007), but delay the monetary policy shock by two periods. The figure shows the implied $R^2_{\ell,FG}$ with the conventional small set of macro observables: output, inflation, and interest rates.

Figure A.1 (which is taken straight from that paper) shows various R^2's for a forward guidance shock two quarters ahead. In addition to the $R^2_{0,FG}$ relevant for SVAR inference, I additionally consider the implied R^2's using current, past and some *future* values of macro aggregates – that is, the more general $R^2_{\ell,j}$ terms of Proposition 1. The x-axis in Figure A.1

[23]For example, in a four-variable VAR in consumption, investment, output and inflation, the $R^2_{0,m}$ at my benchmark parameterization of Smets & Wouters (2007) is equal to 0.0934.

varies the number of periods ℓ that the econometrician is allowed to look into the future. At horizon 0, the R^2 is very low, for two reasons. First, forward guidance shocks do not drive much of the variation in aggregate data (just like conventional policy shocks), and second, they move inflation and interest rates in the same direction (unlike conventional policy shocks). As a result, any econometrician observing the conventional small set of macro aggregates hit by a forward guidance shock would presumably conclude that a demand shock must have occurred. Two periods out, however, when the shock materializes, interest rates and inflation are untied, so the econometrician realizes that the economy was presumably hit by a forward guidance shock – accordingly, the R^2 jumps to the high level familiar for conventional monetary policy shocks.

WEIGHT DECAY. Since the structural model of Smets & Wouters (2007) does not feature news shocks, the incremental information in *future* realizations of macro aggregates – that is, $R^2_{\ell,j} - R^2_{\ell-1,j}$ – is generically small. For example, for monetary policy shocks, we have $R^2_{0,m} = 0.8702$ and $R^2_{\infty,m} = 0.8763$. Consistent with this intuition, the coefficients in the higher-order entries of the lag polynomial $P(L)$ in (9) are generically small.

Figure B.3 in Section B.3 provides an illustration for dynamic shock weights under the identification scheme of Uhlig (2005), studied in Section 3. Given the displayed extremely fast decay of weights, it is *nearly* sufficient to just look at the coefficients in the first entry P_0, as done for most of the plots displayed in Sections 3 to 5.

A.3 Shock volatility and Bayesian posteriors

I first provide a formalization of the limit results displayed in panel (b) of Figure 3.

Proposition A.1. *In the sign-identified three-equation model, impose the uniform Haar prior μ over the set of orthogonal matrices $O(n_x)$. Denote by $\lambda = \lambda(\sigma^m)$ the induced posterior probability of a negative output response as a function of the monetary policy shock volatility σ^m. Then, if $\phi_y > 0$,*

$$\lim_{\sigma^m \to 0} \lambda(\sigma) = 0$$
$$\lim_{\sigma^m \to \infty} \lambda(\sigma) = 1$$

Proof. The boundaries of the identified set of rotation vectors p are defined as the solution to the following optimization problems:

$$\max\backslash\min_{p\in S(3)} \quad \frac{1}{1+\phi_y+\phi_\pi\kappa}\left(\sigma^d p_{md}+\phi_\pi\sigma^s p_{ms}-\sigma^m p_{mm}\right)$$

such that

$$\begin{aligned} \kappa\sigma^d p_{md}-(1+\phi_y)\sigma^s p_{ms}-\kappa\sigma^m p_{mm} &\leq 0 \\ (\phi_y+\phi_\pi\kappa)\sigma^d p_{md}-\phi_\pi\sigma^s p_{ms}+\sigma^m p_{mm} &\geq 0 \end{aligned}$$

where $S(n)$ denotes the n-dimensional unit sphere. Now let $\sigma^d, \sigma^s \to 0$, and fix σ^m.[24] It is then easy to see that

$$\lim_{\sigma^d,\sigma^s\to 0} \overline{IS}_{y,m,0} = 0$$

The Haar prior assigns strictly positive mass to the sign-identified set, since $p_{mm} \in [0,1]$ and the weights on the demand and supply shocks are unrestricted. But the upper bound is only attained for a weight of 0 on the monetary policy shock, which maps into a measure-0 subset of the unit sphere under the Haar prior.

Next, as $\sigma^m \to 0$, the identifying restrictions simplify to

$$\begin{aligned} \kappa\sigma^d p_{md}-(1+\phi_y)\sigma^s p_{ms} &\leq 0 \\ (\phi_y+\phi_\pi\kappa)\sigma^d p_{md}-\phi_\pi\sigma^s p_{ms} &\geq 0 \end{aligned}$$

Thus p_{md} and p_{ms} must both be of the same sign. Re-arranging the inequalities, we see that, if $\phi_y > 0$, then p_{md}, p_{ms} must both be positive and, whenever $p_{md}, p_{ms} > 0$, we require that

$$\frac{p_{md}}{p_{ms}} \in \left[\frac{\phi_\pi}{\phi_y+\phi_\pi\kappa}, \frac{1+\phi_y}{\kappa}\right] \times \frac{\sigma^s}{\sigma^d}$$

where the interval has strictly positive length precisely because $\phi_y > 0$. With the relative demand and supply weights in this interval, and p_{mm} unrestricted, we see again that the posterior mass of the identified set is strictly positive. Since $p_{md}, p_{ms} \geq 0$, we conclude that the lower bound of the identified set is 0, attained with a weight of 1 on the monetary policy shock. Again this maps into a measure-0 subset of the unit sphere under the uniform Haar

[24] With $\sigma^m \to \infty$, the upper bound would remain finite and strictly positive, but the weight on the monetary policy shock corresponding to any strictly positive output response would also be vanishingly small, so the argument is unchanged.

prior. □

Proposition A.1 provides a formal rationalization of the results in Paustian (2007) and Canova & Paustian (2011): For more volatile structural shocks, the Haar-implied posterior probability of a correctly signed impulse response increases.

However, it is also straightforward to use the constructive logic of the proof to find alternative measures $\tilde{\mu}$ that, for any *finite* set of variances $\sigma = (\sigma^d, \sigma^s, \sigma^m)$, give arbitrary posterior probabilities for either positive or negative output responses. By the discussion in Section 3.1, the identified set of the static three-equation model always contains strictly positive and strictly negative output responses. For example, for a given set of variances $\sigma = (\sigma^d, \sigma^s, \sigma^m)$, let $\mathcal{P}^+$ denote the set of weight vectors that give a strictly positive output response. For any alternative set of volatilities $\tilde{\sigma} = (\tilde{\sigma}^d, \tilde{\sigma}^s, \tilde{\sigma}^m)$, and for any $p \in \mathcal{P}^+$, let $\tilde{p}_i^* = p_i \times \frac{\sigma^i}{\tilde{\sigma}^i}$, $i \in (d, s, m)$, and $\tilde{p} \equiv \frac{\tilde{p}^*}{||\tilde{p}^*||}$. All weight vectors $\tilde{p}$ in the thus defined set $\tilde{P}^+$ give strictly positive output responses in the rescaled model. It remains to simply pick a measure $\tilde{\mu}$ that assigns arbitrarily large prior probability to $\tilde{\mathcal{P}}^+$.[25]

[25] Since my results rely on statements about *relative* shock volatilities, it is unsurprising that the conclusions of Section 3 apply just as well to structural models estimated on different sample periods. For example, in a variant of Smets & Wouters (2007) estimated on pre-Great Moderation samples, all shocks are more volatile, but again demand and supply shocks are more prominent than monetary policy shocks, so the shape of posterior distributions over identified sets is largely unchanged.