Data-Driven Inference on Sign Restrictions in Bayesian Structural Vector Autoregression *

Markku Lanne
University of Helsinki and CREATES

Jani Luoto
University of Helsinki

Abstract

Sign-identified structural vector autoregressive (SVAR) models have recently become popular. However, the conventional approach to sign restrictions only yields set identification, and implicitly assumes an informative prior distribution of the impulse responses whose influence does not vanish asymptotically. In other words, within the set the impulse responses are driven by the implicit prior, and the likelihood has no significance. In this paper, we introduce a Bayesian SVAR model where unique identification is achieved by statistical properties of the data. Our setup facilitates assuming a genuinely noninformative prior and thus learning from the data about the impulse responses. While the shocks are statistically identified, they carry no economic meaning as such, and we propose a procedure for labeling them by their probabilities of satisfying each of the given sign restrictions. The impulse responses of the identified economic shocks can subsequently be computed in a straightforward manner. Our approach is quite flexible in that it facilitates labeling only a subset of the sign-restricted shocks, and also concluding that none of the sign restrictions is plausible. We illustrate the methods by two empirical applications to U.S. macroeconomic data.

Keywords: Structural vector autoregression, independence, posterior model probability, monetary policy shock

JEL Classification: C32, C51, C52

*Financial support from the Academy of Finland is gratefully acknowledged. The first author also acknowledges financial support from CREATES (DNRF78) funded by the Danish National Research Foundation, while the second author is grateful for financial support from the Research Funds of the University of Helsinki.

1 Introduction

The structural vector autoregressive (SVAR) model is one of the prominent tools in empirical macroeconomics. While the reduced-form vector autoregression (VAR) is useful for describing the joint dynamics of a number of time series and forecasting, it is only when some structure is imposed upon it that interesting economic questions can be addressed. Typically structural VAR (SVAR) analysis involves tracing out the dynamic effects (impulse responses) of economic shocks on the variables included in the model, and these shocks are often identified by restricting their short-run or long-run effects (for a survey on SVAR models, see Kilian (2013)). In addition, identification by sign restrictions, put forth by Faust (1998), Canova and De Nicoló (2002), and Uhlig (2005), has become an increasingly popular way of computing impulse responses of economic shocks in the recent macroeconomic SVAR literature. Compared to approaches involving explicit parametric identification restrictions, sign restrictions are less stringent, yet manage to convey economic intuition. Therefore, they have a great appeal in empirical research, where a sufficient number of plausible parametric restrictions may be difficult to come by.

However, identification by sign restrictions also involves a number of problems. First, Baumeister and Hamilton (2015) recently showed that the conventional approach to imposing sign restrictions implicitly assumes an informative prior distribution of the impulse responses, whose influence does not vanish asymptotically. In other words, the impulse responses produced are only draws from the implicit prior distributions. Instead of the conventional procedure, they recommended making explicit the prior information used in the analysis. Second, as pointed out by Fry and Pagan (2011), sign restrictions fail to identify a unique model as a large number of models fit the data equally well, i.e., only set identification is achieved. From a Bayesian perspective, this need not be seen as a problem, but the uncertainty surrounding the impulse responses can be thought of as arising from having a limited set of data as well as from doubts about the true model structure, as Baumeister and Hamilton (2015) emphasize. Nevertheless, the lack of uniqueness impedes reporting the results of impulse response analysis, especially as the identified set may be very large. Third, it is not straightforward to assess the plausibility of given sign restrictions, which is in contrast to SVAR models identified by explicit parameter restrictions. Overidentifying restrictions in structural VAR models identified by parametric restrictions can be tested by usual classical tests or assessed by computing posterior model probabilities. Moreover, recent advances in the so-called statistical identification literature (see, e.g., Lanne et al. (2015), and the references therein) facilitate also testing exactly identifying economic parameter restrictions.

In this paper we address all three concerns mentioned above. Our starting point is the SVAR model where, following Lanne et al. (2015), unique identification is achieved by means of statistical properties of the data. Specifically, we assume that the components of the structural error vector are independent, each following some non-Gaussian univariate distribution. Lanne et al. (2015) considered the properties of the maximum likelihood estimator of this model, while we concentrate on Bayesian estimation that facilitates explicitly incorporating any prior information and computing posterior model probabilites. Because our model is uniquely identified instead of being just set identified, the posterior distributions of the impulse responses need not be driven by the priors, which facilitates learning about the impulse responses from the data. In addition, the model identification problem emphasized by Fry and Pagan (2011) is completely sidestepped.

Our uniquely identified SVAR model also facilitates assessment of sign restrictions. The posterior probabilities of the models restricted by them can be interpreted as the probabilities of those restrictions, and used to assess their plausibility. Subsequently those economic shocks that are found plausibly identified, can be given the economic interpretation related to the corresponding restrictions. It may also turn out that only a subset or none of the sign restrictions are in accordance with the data. The sign restrictions deemed unlikely are "rejected", and it is concluded that they are not useful in identifying the economic shocks in question. Impulse response and related analyses of the remaining shocks (if any) can then be conducted in a straightforward manner.

There are a few previous suggestions on how to assess the plausibility of sign restrictions. Straub and Peersman (2006) used the proportion of discarded models as an indicator of how well the New Keynesian model that had generated the restrictions, fit the data. This indicator is, however, ambiguous because a high rejection rate may just as well indicate sharp identification (the set of acceptable models is small) or an inefficient sampler as lack of fit. Piffer (2015) formalized this approach, but his procedure seems difficult to generalize beyond the bivariate VAR model. Baumeister and Hamilton (2015) illustrated how the effect of the tightness of priors on the posteriors yields information on the plausibility of the restrictions, but this approach is applicable only when the priors are explicitly spelled out. Furthermore, it does not work with the conventional approach to sign restrictions because in the absence of point identification the posterior will continue to be driven by these priors.

Closely related to our approach, Herwartz and Lütkepohl (2014), Lütkepohl and Netšunajev (2014), and Lanne et al. (2015) have recently informally assessed the conformity of sign restrictions with the data, and subsequently used the "accepted" restrictions in labeling the shocks of statistically identified SVAR models. In the first two papers, unique identification is achived by modeling the error term of the SVAR model as a Markov-switching process (see Lanne et al. (2010)). The idea of this approach is to check whether the impulse responses implied by the uniquely identified SVAR model satisfy the sign restrictions. If the restrictions are satisfied, the shocks can be labeled accordingly. The procedure put forth in this paper formalizes this approach and augments it by quantifying the likelihood of the sign restrictions.

The rest of the paper is organized as follows. In Section 2, we describe the SVAR model and discuss its identification along the lines of Lanne et al. (2015). Section 3 introduces the procedure for computing the probabilities of the sign restrictions and finding the plausible sign-identified shocks among all the statistically identified shocks. In Subsections 3.1 and 3.2, we propose stepwise procedures for the cases of a single and multiple sign-identified shocks, respectively. In Section 4 the computation of impulse responses and forecast error decompositions of the sign-identified shocks is discussed. Because we have a uniquely identified SVAR model, the model identification problem is avoided, which facilitates interpretation of forecast error variance decompositions. Morever, unlike the conventional sign identification approach, the effects of shocks of a given size can be examined. Section 5 contains two empirical applications that illustrate our procedure. In Subsection 5.1, we consider the identification of the monetary policy shock in Uhlig's (2005) SVAR model of the U.S. economy, and in Subsection 5.2 we revisit Peersman's (2005) fully identified macrocoeconomic SVAR model. Finally, Section 6 concludes.

2 Model

Consider the n-variate structural VAR(p) model

$$y_t = a + A_1 y_{t-1} + \cdots + A_p y_{t-p} + B\varepsilon_t, \tag{1}$$

where y_t is a vector of time series of interest, a is an intercept term, $A_1, \ldots, A_p$ are $n \times n$

coefficient matrices, and the matrix B summarizing the contemporaneous structural relations of the errors is assumed nonsingular. In order to facilitate identification of matrix B, we make two further assumptions. First, we assume that ε_t is a sequence of stationary random vectors with each component ε_{it}, $i = 1, \cdots, n$, being serially uncorrelated and having zero mean and finite positive variance. Second, it is assumed that the components ε_{it} are mutually independent, and at most one of them has a Gaussian marginal distribution. In the empirical applications in Section 5, we specifically assume that each component of the error vector individually follows a t distribution. Because a t-distributed random variable converges to a Gaussian as the number of degrees of freedom approaches infinity, this is more general than the usual (implicit) normality assumption and affords more flexibility (see, for example, Koop (2003, p. 126) for a more detailed discussion).

If the process y_t is stable, i.e.,

$$\det\left(I_n - A_1 z - \cdots - A_p z^p\right) \neq 0, \ |z| \leq 1 \ \ (z \in \mathbb{C}),$$

the SVAR(p) model (1) has a moving average representation

$$y_t = \mu + \sum_{j=0}^{\infty} \Psi_j B \varepsilon_{t-j}, \tag{2}$$

where μ is the unconditional expectation of y_t, Ψ_0 is the identity matrix, and $\Psi_j, j = 1, 2, \ldots$, are obtained recursively as $\Psi_j = \sum_{l=1}^{j} \Psi_{l-j} A_j$. The kth column of $\Psi_j B \equiv \Theta_j$, $j = 0, 1, \ldots$, contains the impulse responses of the kth structural shock ε_{it}, $i = 1, \ldots, n$, and it is these impulse reponses that are the main object of interest in SVAR analysis. An integrated VAR(p) process does not satisfy the stability condition above, and hence, has no moving average representation. Nevertheless, the impulse responses are also in this case given by the same recursion (see Lütkepohl (2005, Section 6.7)).

Under the non-Gaussianity and indepenence assumptions on the error term ε_t above, matrix B is unique apart from permutation and scalings of its columns (see Proposition 1 and its proof in Lanne et al. (2015)). In other words, the model remains the same after changing the order of the columns of B or multiplying them by constants once the shocks ε_{it} are reordered and scaled accordingly. Hence, the structural shocks and their impulse responses are uniquely identified, but despite this 'statistical' identification, the shocks cannot be labeled or given any economic interpretation without additional restrictions. Recently, Lanne et al. (2015) showed how conventional short-run and long-run identifying restrictions can be tested in this framework, and if not rejected, used for economic identification. Following Herwartz and Lütkepohl (2014), and Lütkepohl and Netšunajev (2014), they also demonstrated how sign restrictions can be informally used in this model to facilitate economic identification. In Section 3 below, we formalize their approach and show how to assess the plausibility of given sign restrictions in a SVAR model under our identifying restrictions.

3 Identification by Sign Restrictions

Because under the non-Gaussianity and independence assumptions the impact matrix B in (1) is uniquely identified (apart from permutation and scaling of its columns), also the structural shocks and their impulse reponses are identified. The only difference between the models corresponding to the different permutations of the columns of B is the ordering of the shocks. Therefore, a model with any fixed permutation facilitates quantifying the plausibility of any sign restrictions, and it is important to ensure that the entire analysis is based on the same permutation. We first compute the (posterior) probabilities of each combination of the shocks to be those identified by the sign restrictions. Subsequently, the impulse reponses (and forecast error variance decompositions) of the shocks that are found likely to satisfy the sign restrictions (if any), can be readily computed, and since they are linked to the given restrictions, their interpretation is straightforward. In this section, we concentrate on assessing the sign restrictions, while in the next section, we discuss the computation of the impulse responses and forecast error variance decompositions.

We set out with the case of a single structural shock identified by sign restrictions that restrict only the impact effect, which is probably the most common case in the empirical literature, and then proceed to the more general case of restrictions on the first $q+1$ impulse responses. Subsection 3.2 covers the case of multiple structural shocks.

3.1 Identifying a Single Structural Shock

Suppose we are interested in finding the impulse reponses of a single shock, whose impact effects on J of the variables in y_t are given. This might be, say, the monetary policy shock with a non-positive impact effect on prices and nonborrowed reserves and a non-negative impact effect on the Federal funds rate (cf. the empirical application in Section 5.1). Let us collect these sign restrictions in a $J \times n$ matrix R, whose elements equal 1, -1, or 0, and define a set Q such that

$$Q = \{\theta_{0k} : R\theta_{0k} \geq 0_{J\times 1}\}, \tag{3}$$

where θ_{0k} is the kth column of Θ_0, or equivalently, of the impact matrix B, corresponding to shock ε_{kt}. The set Q thus consists of the columns of B that satisfy the sign restrictions. Although we are after a single shock, Q may contain multiple columns of B, or it may be empty, i.e., there may be more than one shock or no shock satisfying the sign restrictions. This is in contrast to the conventional approach in the sign restriction literature, where a single shock satisfying the restrictions, by construction, is found.

Since our assumptions only identify B up to permutation of its columns, any (or none) of the n components of ε_t can be the structural shock satisfying the sign restrictions. In order to assess the plausibility of one of the shocks being the shock of interest, we compute for each shock ε_{kt}, $k = 1, \ldots, n$, the conditional probability of being this shock (conditional on the vector of data, $\mathbf{y}$, obtained by stacking y_t for $t = 1, \ldots, T$),

$$\Pr\left(\theta_{0k} \in Q, \theta_{0,m\neq k} \in Q^c \,|\mathbf{y}\right), \tag{4}$$

where Q^c denotes the complement of Q, and $m \in \{1, \ldots, n\}$. For each $k \in \{1, \ldots, n\}$, the quantity (4) can be interpreted as the posterior probability of the restricted SVAR model, where the sign restrictions embodied in R are imposed on the kth column of B only (cf. Koop (2003, p. 81) in the context of the linear regression model). Among the n models, we expect those that satisfy the restrictions in the (true) data-generating process (DGP) (i.e., the models for which $\theta_{0k} \in Q$ in the DGP) to have high posterior probabilities, with the likeliest model reaching the maximum. Thus, by ranking the restricted SVAR models according to their posterior probabilities, we are able to locate the likeliest model and, hence, the shock that is the likeliest to be the structural shock of interest.

It is, of course, possible that none of the shocks satisfies the sign restrictions. The probability of the SVAR model where the sign restrictions are violated for all $k \in \{1, \ldots, n\}$, can be readily computed by subtracting the sum of the probabilities in (4) from one. If this probability lies close to unity, the data do not lend support to the sign restrictions, and therefore, they cannot be used to identify the structural shock of interest. In other words, the sign restrictions are “rejected”.

An ambiguous situation arises if more than one of the shocks satisfy the restrictions with (almost) equal high probability. In this case, additional information is needed to discriminate between the plausible shocks. It may come in the form of quantitative information about the likely magnitude of the impulse responses (see, for example, Kilian and Murphy (2009)). It is also rather common practice in the literature to add sign restrictions on longer lags in the impulse responses beyond the impact effect, to which our framework also lends itself in a straightforward manner. To that end, in order to impose the same restrictions on a single shock embodied in matrix R on Θ_j, $j = 0, 1, \ldots, q$, we redefine the set Q as

$$Q = \left\{ \theta_k : (I_{q+1} \otimes R)\, \theta_k \geq 0_{J(q+1)\times 1} \right\}, \tag{5}$$

where θ_k denotes the kth column (corresponding to shock ε_{kt}) of $\Theta = \left[\Theta_0', \ldots, \Theta_q'\right]'$, a matrix consisting of the first $q + 1$ structural impulse responses. Analogously to (4), the conditional probability of ε_{kt} being the structural shock of interest is then defined as

$$\Pr\left(\theta_k \in Q, \theta_{m\neq k} \in Q^c \,|\mathbf{y}\right), \tag{6}$$

and the analysis proceeds as in the case of restrictions on Θ_0 only.

Although less common in the empirical literature, our framework also accommodates different sign restrictions on different lags (see, for instance, Inoue and Kilian (2013), who imposed an additional sign restriction on the sixth lag in Uhlig's (2005) SVAR model, discussed also in Section 5.1, when identifying the monetary policy shock). In that case, we would simply replace the matrix $I_{q+1} \otimes R$ in (5) by a block-diagonal matrix $diag\left(R^0, \ldots, R^q\right)$, where R^j incorporates the restrictions on the jth lag. If not all impulse responses with lags up to q, but only lags belonging to some set L are restricted, then this block diagonal matrix has only the R^j matrices with $j \in L$ on its main diagonal, and the matrix Θ is adjusted accordingly.

The proposed procedure can be summarized as follows:

Step 1. Estimate the joint posterior distribution of the parameters of the unrestricted SVAR model (1), and compute the posterior distribution of the reduced-form impulse response matrices Ψ_j, $j \in L$. If sign restrictions are imposed on all the $q + 1$ first lags of the impulse response function, $L = \{0, 1, \ldots, q\}$.

Step 2. Given the posterior output of B from Step 1, rearrange the columns of each B to ensure that all the posterior draws of B represent the same permutation. This is accomplished by first computing the transformed matrices $\widetilde{B}$, whose each column has Euclidean norm one, and then finding a permutation matrix P for which $C =$

$\widetilde{B}P = (c_{ij})$, satisfies $|c_{ii}| > |c_{ij}|$ for all $i < j$.[1] Then, for each B and Ψ_j, the uniquely identified structural impulse responses are given by $\Theta_j = \Psi_j BPD$, $j \in L$, where D is a diagonal matrix with elements equal to either 1 of -1 that transforms the diagonal elements of BP positive.

Step 3. Calculate the probabilities in (6) (or (4)) for all $k \in \{1, \ldots, n\}$ using the posterior distribution of the identified structural impulse responses. If the sum of these probabilities lies close to zero, or, in other words, if the posterior probability of the SVAR model satisfying none of the sign restrictions is high, conclude that the data are not compatible with the restrictions, and they cannot thus be used for identification. Otherwise label the structural shock of interest according to the calculated posterior probabilities.

3.2 Identifying Multiple Structural Shocks

The procedure introduced in Subsection 3.1, generalizes in a straightforward manner to the case of $g > 1$ structural shocks, each of which is restricted by J_i, $i = 1, \ldots, g$, sign restrictions. Instead of a single R matrix, we then have g $J_i \times n$ matrices R_i, each embodying the J_i restrictions, and the set

$$Q_i = \left\{\theta_k : (I_{q+1} \otimes R_i)\, \theta_k \geq 0_{J_i(q+1)\times 1}\right\}, \tag{7}$$

contains the columns of the matrix of impulse responses Θ that satisfy the ith sign restrictions.

Analogously to the case of a single shock, computing the posterior probability of the g shocks identified by the sign restrictions calls for going though all combinations of the columns of Θ. For example, the posterior probability of the restricted SVAR model in which the sign restrictions concern two structural shocks ($g = 2$) is given by

$$\Pr\left(\theta_k \in Q_1, \theta_l \in Q_2, \theta_{m\neq k,l} \in Q_2^c \,|\mathbf{y}\right) \quad \text{for } k, l \in \{1, \ldots, n\}, k \neq l, \tag{8}$$

[1]Note that the procedure does not quarantee that all the permuted posterior draws of B come from the same permutation. If the procedure completely failed to keep the chosen permutation between the posterior draws, the resulting posterior probability estimates in (6) (or (4)) would be equal for all the elements of ε_t. However, in large samples, the procedure never fails (in probability) because the posterior variances of the elements of B decrease with sample size (see Baumeister and Hamilton (2015), for a detailed discussion concerning the asymptotic properties of Bayesian inference in SVARs).

where Q_2^c is the complement of the union $Q_1 \cup Q_2$. In this case, we have $n(n-1)$ different SVAR models to go through. For fixed k and l, (8) is the posterior probability of ε_{kt} and ε_{lt} being the two structural shocks, and the sum of these probabilities over all combinations of k and l subtracted from one can be interpreted as the posterior probability of the sign restrictions not being satisfied.

In general, we have $n!$ permutations of the columns of Θ, on which the g restrictions can be placed. However, once the positions of the g shocks have been fixed, the ordering of the remaining unrestricted columns is irrelevant for the assessment of the plausibility of the restrictions. Because there are $(n-g)!$ permutations of these columns, the total number of restricted SVAR models that contain the g shocks in fixed positions is $n!/(n-g)!$. This suggests that the posterior probabilities of the restricted SVAR models, such as those in (8), can be evaluated by first calculating the probabilities of the $n!$ SVAR models where the g restrictions are imposed on any g columns of Θ, and then marginalizing over each set of the $(n-g)!$ models where they are imposed on same g columns of Θ to obtain the probabilities of the $n!/(n-g)!$ models.

Formally, all $n!$ possible permutations of the columns of Θ can be obtained as the products ΘP^s for $s \in \{1, \ldots, n!\}$, where P^s is an $n \times n$ permutation matrix. The probability that the first g columns of ΘP^s satisfy the g sign restrictions can be expressed as

$$\Pr\left(\theta_1^s \in Q_1, \ldots, \theta_g^s \in Q_g, \theta_{m\in\{g+1,\ldots,n\}}^s \in \mathbf{Q}_g^c \left|\mathbf{y}\right.\right), \text{ for } s \in \{1, \ldots, n!\} \tag{9}$$

where Q_g^c is the complement of the union $Q_1 \cup \cdots \cup Q_g$.[2] It can be readily checked that the quantities in (9) are the posterior probabilities for all the $n!$ restricted SVAR models. As pointed out above, the probabilities of each of the restricted $n!/(n-g)!$ SVAR models of interest are then obtained by summing the probabilities of the $(n-g)!$ models in which the g sign restrictions are imposed on the same g columns of Θ.

Each of these $n!/(n-g)!$ models represents one ordering of the g structural shocks in the vector ε_t. Thus, by ranking them by their posterior model probabilities, we are able to single out in probability the likeliest ordering of the g structural shocks of interest. If several orderings turn out equally likely, additional information is needed to discriminate between the corresponding structural models. Analogously to the single shock case, the probability of the sign restrictions failing to identify all g shocks, i.e., the probability of

[2]Notice that when all n shocks are identified, (9) reduces to $\Pr\left(\theta_1^s \in Q_1, \ldots, \theta_g^s \in Q_g \left|\mathbf{y}\right.\right)$, for $s \in \{1, \ldots, n!\}$.

the SVAR model where all the sign restrictions are violated for all $s \in \{1, \ldots, n!\}$, can be calculated by subtracting the sum of the probabilities in (9) from one. Again, if this probability lies close to one, there is little evidence in favor of the sign restrictions in the data, and they can be "rejected".

For notational convenience, we concentrate on the case of the same sign restrictions on each of the g shocks at lags from 0 to q. However, as in the case of a single structural shock, the approach generalizes in a straightforwad manner to the case where impulse responses of all shocks are not restricted at all lags, or the restrictions on (some of) the shocks are different across the lags. Peersman's (2005) study discussed in Section 5.2 provides an example of the latter situation: the sign restrictions on two of the four variables included are binding only instantaneously ($q = 0$), while for the two remaining variables, the restrictions are binding up to four quarters ($q = 4$).

In the case of multiple structural shocks of interest, there are several ways to proceed in checking for identification. We recommend the following procedure that is illustrated by means of an empirical application in Section 5.2. A straightforward alternative is, of course, to compute of the posterior probabilities of the $n!/(n-g)!$ SVAR models directly, and label the shocks according to the likeliest one, provided any model is deemed plausible. However, this procedure has the downside that it overlooks the cases where only part of the shocks are identififed.

Step 1. Follow Steps 1 and 2 described in Subsection 3.1, to obtain the identified structural impulse responses.

Step 2. Based on the posterior distribution of the identified structural impulse responses Θ, calculate the probabilities given in (6) (or (4)) for all $Q_1, \ldots, Q_g$ individually. If the sum of the probabilities based on Q_i is close to zero, remove the ith sign restriction ($i \in \{1, \ldots, g\}$). As already discussed in Subsection 3.1, such sign restrictions are "rejected" by the data and, therefore, cannot be used to identify the structural shock of interest.

Step 3. In order to label the shocks based on the probabilities given in (9), find the likeliest SVAR model identified by the remaining sets Q_i. If the posterior probability of the sign restrictions failing to identify all the remaining shocks is high, proceed by recursively removing the weakly identified shocks, i.e., those with respect to which more than one ordering is almost equally likely, until unique labeling is reached with high probability. If the probability of failing to identify the remaining shocks is close to one for all combinations of the sign restrictions, conclude that the data are not compatible with the sign restrictions, and they are thus not useful in identification.

4 Impulse Response and Forecast Error Variance Decomposition Analysis

As already discussed in the Introduction, with sign-identified structural VAR models, reporting the results of impulse response analysis can be problematic due to the so-called model identification problem. In other words, the sign restrictions are consistent with a wide range of statistical models, and the conventional approach thus generates a large set of impulse reponses, each pertaining to a different model. Therefore, pointwise medians of the impulse responses along with error bands are typically reported. Inoue and Kilian (2013) recently considered the shortcomings of these quantities when the model is identified by sign restrictions, and pointed out that the pointwise error bands fail to convey the true uncertainty of the impulse response functions. As a solution, they derived the joint posterior density of the impulse responses and recommended reporting their mode and $100(1-\alpha)\%$ highest posterior density (HPD) credible set.

As our model produces unique impulse response functions, given the size and sign of the shocks, it is not hampered by the model identification problem. Hence, conventional pointwise posterior median impulse responses and error bands could be entertained in a straightforward manner. It is however, well known that, while frequently applied, these may also yield misleading conclusions. Therefore, we recommend employing an extension of the joint posterior density approach of Inoue and Kilian (2013) also in our setup. It is also worth emphasizing that our approach facilitates analyzing the effects of shocks of a given size, and thus answering questions like "what would be the responses to a 25 basis point interest rate shock". As pointed out by Fry and Pagan (2011), inter alia, the conventional approaches to identification by sign restrictions do not lend themselves to such questions, and, therefore, only quantities like peak responses are typically reported.

The posterior density of the structural impulse responses implied by our model can be derived in a straightforward manner. For notational simplicity, let us ignore deterministic terms, and collect the coefficient matrices of model (1) in matrix $\widetilde{\mathbf{A}} = [A_1 \cdots A_p]$. Because the model is exactly identified (statistically), there is a one-to-one mapping between the first $p+1$ structural impulse responses $\widetilde{\Theta} = \left[B', (\Psi_1 B)', \ldots, (\Psi_p B)'\right]'$ and $\left[B, \widetilde{\mathbf{A}}\right]$, and the nonlinear function $\widetilde{\Theta} = f\left(B, \widetilde{\mathbf{A}}\right)$ is known. By a change of variable, the posterior density of the first $p+1$ structural impulse responses $\widetilde{\Theta} = \left[B', (\Psi_1 B)', \ldots, (\Psi_p B)'\right]'$ can thus be written as

$$
\begin{aligned}
p\left(\widetilde{\Theta}\,|\mathbf{y}\right) &= \left|\frac{\partial\left[\operatorname{vec}(B)', \operatorname{vec}(\widetilde{\mathbf{A}})'\right]}{\partial \operatorname{vec}\left(\widetilde{\Theta}\right)}\right| p\left(B, \widetilde{\mathbf{A}}\,|\mathbf{y}\right) \\
&= \left(\left|\frac{\partial \operatorname{vec}\left(\widetilde{\Theta}\right)}{\partial\left[\operatorname{vec}(B)', \operatorname{vec}(\widetilde{\mathbf{A}})'\right]}\right|\right)^{-1} p\left(B, \widetilde{\mathbf{A}}\,|\mathbf{y}\right) \\
&= |B|^{-np}\, p\left(B, \widetilde{\mathbf{A}}\,|\mathbf{y}\right). \qquad (10)
\end{aligned}
$$

where $p\left(B, \widetilde{\mathbf{A}}\,|\mathbf{y}\right)$ is the joint posterior density of B and $\widetilde{\mathbf{A}}$, and the second equality follows by the inverse function theorem. In the terminology of Inoue and Kilian (2013), the model corresponding to a draw $(B, \widetilde{\mathbf{A}})$ from the posterior distribution of the parameters of the SVAR model that maximizes (10) is the *modal* model that produces the mode of the structural impulse responses. Note that there is a one-to-one mapping from the joint posterior density of B and $\widetilde{\mathbf{A}}$ to only the $p+1$ first impulse responses that thus govern the selection of the modal model.

In addition to the mode, it is useful to have a measure of the uncertainty surrounding the impulse responses, and, following Inoue and Kilian (2013), we define the $100(1-\alpha)\%$ HPD credible set of the first $p+1$ impulse responses as

$$S = \left\{\widetilde{\boldsymbol{\Theta}} : p\left(\widetilde{\boldsymbol{\Theta}}\,|\mathbf{y}\right) \geq c_\alpha\right\},$$

where c_α is the largest constant such that $\Pr(S) \geq 1-\alpha$. We then report the impulse responses up to some prespecified horizon of models belonging to this set, in addition to those of the modes. In the empirical literature it is customary to set α equal to 0.32, i.e., to report the 68% credible sets. As Inoue and Kilian (2013) pointed out, there is no reason for these crible sets to be dense, but they will typically exhibit a "shot-gun" pattern.

As far as the forecast error variance decompositions are concerned, they can be calculated in a standard fashion using the mode of the structural impulse responses as defined above (see, for example, Lütkepohl (2005, Chapter 2.3)). Typically forecast error variance decompositions based on poinwise median impulse responsens are reported for sign-identified models. However, as pointed out by Fry and Pagan (2011), they have the problem that they are based on uncorrelated shocks, and may therefore be difficult to interpret because the contributions of all shocks need not sum to unity for all variables. In contrast, since our model is uniquely identified, this problem is avoided.

5 Empirical Illustrations

We illustrate the methods by means of two empirical applications, both involving the U.S. economy. The first one, discussed in Subsection 5.1 is concerned with the computationally most straightforward case of only one shock identified by sign restrictions. In particular, we focus on the monetary policy shock in Uhlig's (2005) model. Our second application in Subsection 5.2, in turn, involves Peersman's (2005) model with multiple sign-identified shocks, whose plausibility we set out to assess.

In both applications, we assume that the ith independent component of the error vector of the structural VAR model (1) follows a univiariate (standardized) Student's t distribution with λ_i degrees of freecom. This deviates from the Bayesian SVAR literature, where the error vector is typically assumed multivariate normal with a diagonal covariance matrix. It is important to realize that our distributional assumption encompasses the Gaussian case because a t-distributed random variable approaches Gaussianity as the number of degrees of freedom goes to infinity. It is also because of this property of the t distribution that the estimates of λ_i, $i = 1, \ldots, n$, indicate the strength of identification (recall that B is uniquely identified only under non-Gaussianity of at least $n-1$ components). According to the results, the error distributions indeed seem to be fat-tailed. In particular, the posterior means of the degree-of-freedom parameters lie between 2.1 and 6.1.

For each degree-of-freedom parameter λ_i, we assume an exponential prior distribution with mean 5 and variance 25. As to the error impact matrix B, we operate on its inverse $\text{vec}(B^{-1}) \equiv \mathbf{b}$, and assume a Gaussian prior distribution, i.e., $\mathbf{b} \sim N(\underline{\mathbf{b}}, \underline{V}_{\mathbf{b}})$. The reported results are based on the special case of $\underline{V}_{\mathbf{b}}^{-1} = c_{\mathbf{b}} I_{n^2}$ with $c_{\mathbf{b}} = 0$, which results in an uninformative (improper) prior for B^{-1}, $p(B^{-1}) \propto 1$. However, the results remain intact irrespective of the (reasonably) informative priors used. Notice that because the SVAR model is point-identified, the improper prior can be used (i.e., the inference can be based solely on the data). This is in contrast to the conventional approach in the sign restriction literature, where the models are only set-identified, and the posterior of the structural parameters within the identified set is proportional to the prior (see Baumeister and Hamilton (2015)). As a result, only under an informative prior does there exist a well-defined posterior for B.

Finally, we collect the deterministic terms and coefficient matrices of model (1) in matrix $\mathbf{A} = \left[a, A_1', \ldots, A_p'\right]'$, and assume a multivariate normal prior for $\text{vec}\left(\mathbf{A}\right) \equiv \mathbf{a}$, $\mathbf{a} \sim N\left(\underline{\mathbf{a}}, \underline{V}_{\mathbf{a}}\right)$. We set $\underline{\mathbf{a}}$ and $\underline{V}_{\mathbf{a}}$ at 0 and $10000^2 I_{pn^2+n}$, respectively. As a robustness check, we also entertained a number of informative priors for $\text{vec}\left(\mathbf{A}\right)$, including close variants of that proposed by Baumeister and Hamilton (2015), and found the results intact irrespective of the priors used. We defer more detailed discussion on the priors and the technical aspects of estimation to the appendix.

5.1 Single Monetary Policy Shock

Uhlig (2005) studied the effects of the U.S. monetary policy shock in a six-variable structural VAR(12) model with no intercept that we take as given. The monthly time series included in the model are the interpolated real GDP, the interpolated GDP deflator, a commodity price index, total reserves, nonborrowed reserves and the federal funds rate, and, for comparability, the sample period is 1965:1-2003:12 as in Uhlig (2005). Save the federal funds rate, all variables are expressed in logs.[3]

Following Uhlig (2005), we identify only the monetary policy shock. The sign restrictions from his Assumption A.1. state that the first six impulse responses of this shock to prices and nonborrowed reserves are nonpositive and to the federal funds rate nonnegative (i.e., $q = 5$ in the notation of Section 3). However, we start by checking the validity of the signs of the impact effects only ($q = 0$), and comment later on the case of restrictions on multiple lags. In all cases, when the variables are included in vector y_t in the order given above, the 4×6 matrix R in (3) or (5) equals

[3]See Uhlig (2005) for a more detailed description of the data set. The data were downloaded from the Estima website at `https://estima.com/resources_indx.shtml`.

$$\begin{pmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

As the first step, we compute the probabilities (4) of each of the six columns of B satisfying the sign restrictions. The sum of these probabilities is 0.29%, lending overall support to the sign restrictions on the impact effects. As a matter of fact, only one shock takes a high posterior probability, almost 0.29%, while the probabilities of the other shocks are virtually zero. Thus the monetary policy shock can be considered uniquely identified in probability. We also checked Uhlig's (2005) original sign restrictions on the six first impulse responses, and the probability of none of the shocks satisfying them turned out to be virtually one. In other words, the data do not lend support to such a highly restrictive model, which should not be too surprising given that each additional restriction presumably narrows down the set of admissible models.

The modes of the impulse responses of a 25 basis point contractionary monetary policy shock along with their 68% joint regions of high posterior density are depicted in Figure 1. Compared to the results of Inoue and Kilian (2013) based on Uhlig's (2005) original sign restrictions our impulse responses seem more precisely estimated.[4] As to the impulse responses of the monetary policy shock on the real GDP that Uhlig (2005) was mostly interested in, its mode is slightly positive only in the first three months and then turns persistently negative, which is intuitively appealing. While there still remains uncertainty about the effects of monetary policy in that the 68% HPD credible set contains positive output responses, it is the negative values that dominate, in contrast to what Uhlig (2005) and Inoue and Kilian (2013) found.This is likely to follow from the fact that our model is exactly identified, whereas sign restrictions alone only reach set identification (cf. the corresponding results of Inoue and Kilian (2013) based on sign and recursive restrictions).

Finally, in Table 1, we report the forecast error variance decomposition of the monetary policy shock at a number of horizons.[5] As discussed in Section 4, the forecast

[4]Inoue and Kilian (2013) computed the 68% HPD credible set in the same way as we did, which facilitates convenient comparison. However, it must be kept in mind that it is not possible to fix the size of the shock in conventional SVAR models identified by sign restrictions.

[5]The forecast error variance decompositions are based on the modes of the impulse responses, and

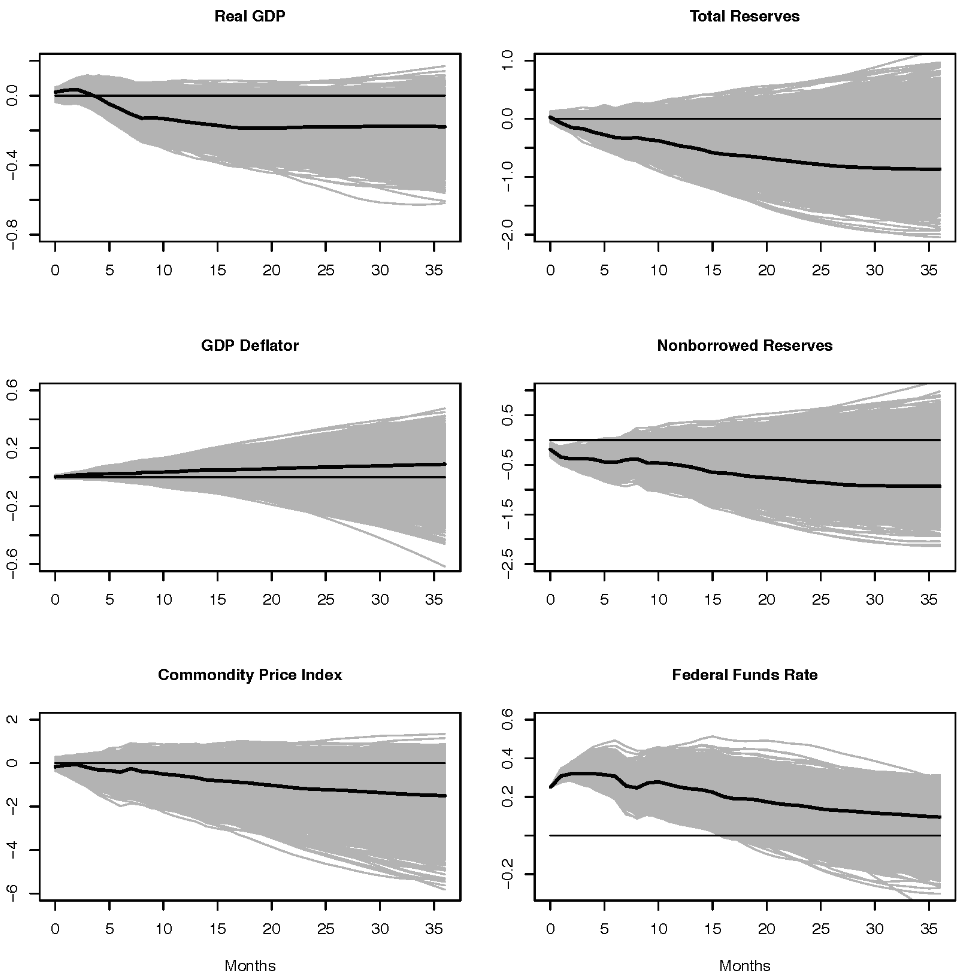

Figure 1: Impulse responses of a 25 basis point contractionary monetary policy shock. The black lines depict the modes of the responses, and the shaded areas are the 68% joint regions of high posterior density.

error variance decomposition is problematic in the case of sign restrictions because of the model identification problem that our approach avoids. In line with Uhlig's (2005) findings, the monetary policy shock accounts for a very small fraction of the forecast error variance of real output albeit this fraction increases slowly with the horizon. However, while Uhlig (2005) found the fraction of the forecast error variance of the federal funds rate accounted for by the monetary policy shock after six months negligible, our results somewhat surprisingly suggest that it is of great importance also at longer horizons.

the variances of the Student's t structural shocks are set at the posterior means of $\lambda_i/(\lambda_i - 2)$ for $i \in \{1, \dots n\}$.

Table 1: Forecast error variance decomposition of the monetary policy shock in Uhlig's (2005) model.

	Horizon (months)					
Variable	1	2	6	12	24	36
Real GDP	0.004	0.005	0.005	0.029	0.063	0.080
GDP Deflator	0.000	0.000	0.002	0.003	0.005	0.006
Commondity Price Index	0.211	0.052	0.038	0.077	0.257	0.375
Total Reserves	0.008	0.021	0.106	0.174	0.209	0.194
Nonborrowed Reserves	0.029	0.062	0.124	0.165	0.192	0.189
Federal Funds Rate	0.998	0.977	0.863	0.691	0.496	0.451

The figures are the proportions of the forecast error variance at each horizon accounted for by the monetary policy shock based on the modes of the impulse responses.

5.2 Multiple Economic Shocks

In this section, we demonstrate how the validity of sign restrictions can be checked in the case of multiple identified shocks. We revisit Peersman's (2005) structural VAR model for the U.S. economy. The model has four variables: the first difference of log oil price (Δoil_t), output growth (Δg_t), consumer inflation (Δp_t), and the three-month nominal interest rate (s_t). The data are quarterly, and the sample period is 1980:Q1–2002:Q2.[6] We follow Peersman (2005) and estimate a VAR(3) model with a linear trend and a constant.

Peersman (2005) identified four shocks that have the effects with signs given in Table 2. For instance, following a contractionary monetary policy shock that has a non-negative effect on the interest rate, there is no increase in prices, in accordance with Uhlig's (2005) identification in Section 5.1. Collecting the variables in the vector $y_t = (\Delta oil_t, \Delta g_t, \Delta p_t, s_t)'$, the R matrices embodying the sign restrictions identifying the oil price, supply and demand shocks in (7) can be written as

$$R_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, R_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \text{ and } R_4 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

[6] The data were downloaded from the Journal of Applied Econometrics Data Archive at: `http://qed.econ.queensu.ca/jae/2005-v20.2/peersman/`.

respectively, while the matrix R_3 corresponding to the demand shock is a 4×4 identity matrix. Peersman (2005) set the time period over which the restrictions are binding four quarters ($q = 4$) for output and consumer prices, while he only restricted the contemporaneous response of the oil price and the nominal interest rate ($q = 0$). However, we restrict only the impact effects (i.e., the case of $q = 0$), and in line with Peersman (2005), find that the value of q has little influence.[7]

Table 2: Sign restrictions for the four shocks in Peersman's (2005) model.

Shock	oil_t	g_t	p_t	s_t
Oil price	+	−	+	+
Supply	?	+	−	−
Demand	+	+	+	+
Monetary policy	−	−	−	+

The + and − signs indicate that the effect can be nonnegative or nonpositive, respectively, while the ? sign stands for an unspecified effect.

As the first step, we assess the plausibility of each of the shocks separately. The sums of the posterior probabilities of the oil price, supply, demand and monetary policy shocks equal 0.56, 0.19, 0.53 and 0.01, respectively. As these can be interpreted as the posterior probabilities of each set of sign restrictions related to one shock, we conclude that the data lend support to the first three shocks, while the monetary policy shock is "rejected".

The sum of the probabilities in (9) for the remaining three shocks turns out to be 0.08, indicating that there is some, but not very strong support in favor the sign restrictions identifying them. We find two potential labelings of the shocks with nonnegligible posterior probabilities (9), 0.05 and 0.03, and both identify the same shocks as the oil price and supply shocks, while the demand shock is different in each case. The impulse response functions in Figure 2 and the forecast error decompositions in Table 3 are based on the former labeling. There we call the remaining shock the monetary policy shock, but given the results above, its effects should be interpreted with caution.

[7] The results based on Peersman's (2005) identification are not reported to save space, but they are available upon request.

Given the small posterior probability, 0.05, of the identification of the three shocks, it is of interest to check the identification of the oil price and supply shocks only. The probability of their being identified (i.e., the sum of the probabilities (8) of any two shocks satisfying the related sign restrictions) equals 0.27, and the probability of the particular labeling found above is 0.26, with all other combinations taking negligible probabilities. Hence, there is relatively strong evidence in favor of the identification of at least these two shocks.

The modes of the impulse responses to unit shocks along with their 68% credible sets are depicted in Figure 2. The effects of the oil price shock (with the strongest identification) satisfy Peersman's (2005) restrictions. There is also posterior evidence in favor or the effect of this shock on oil prices and consumer inflation, as the 68% credible sets of the associated responses do not include the zero line. Unfortunately, the rest of the responses are a posteriori insignificant. The oil price response to the supply shock that Peersman (2005) left unrestricted turns out slightly negative, and its output response is positive in line with his restrictions. However, the modes of the responses of prices and the interest rate to the supply shock are of the wrong sign, as are most of the responses to the other two shocks.

The very wide credible sets of the impulse responses compared to Peersman's (2005) results probably in part reflect the fact that we are reporting the 68% joint credible sets, while he reported the corresponding pointwise error bands of each individual impulse response function. The latter are likely to undermine the true uncertainty, as pointed out by Inoue and Kilian (2013).[8] Furthermore, while Peersman's (2005) results are strongly driven by the prior restrictions imposed on the structural parameters, our results are based solely on the data, and are also therefore likely to be wider. Nevertheless, our credible sets have the advantage that they, at least in principle, facilitate learning about the structural parameters from the data.

Insignificance of the impulse responses highlights the importance of our assessment procedure, as obviously mere visual inspection of the impulse responses would not be

[8]We also checked the results using the likelihood preserving normalization rule of Waggoner and Zha (2003) for the signs of the columns of B. The conclusion that most of the responses are a posteriori insignificant remains intact irrespective of the normalization rule used. It is also worth noting that the probabilities concerning the identification of the shocks are practically the same for the normalization rules used in this paper.

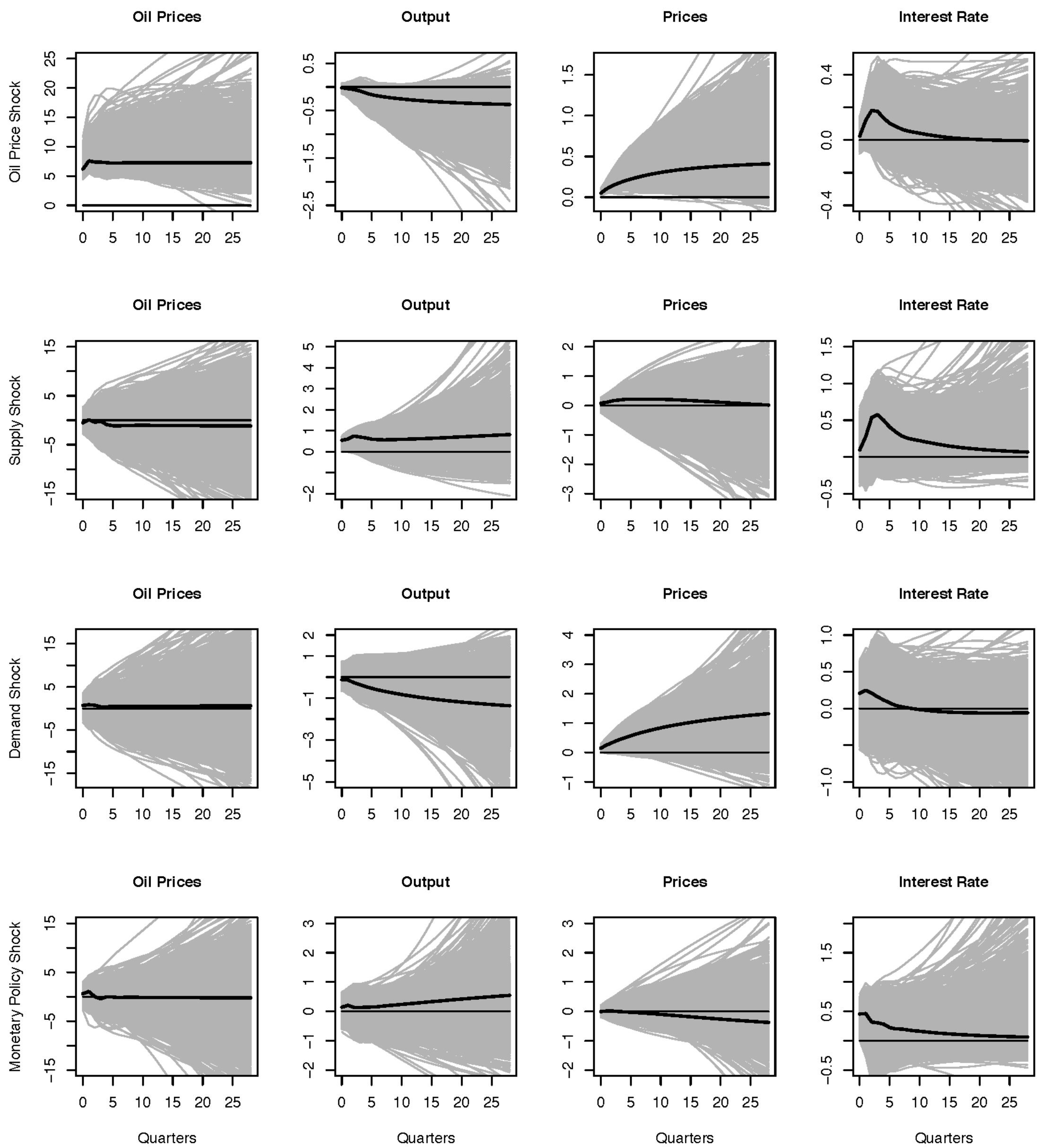

Figure 2: Impulse responses of the four shocks. The black lines depict the modes of the responses, and the shaded areas are the 68% joint regions of high posterior density.

helpful in labeling the shocks. The overall conclusion from the results of this procedure and the impulse responses seems to be that only the oil price shock can be safely be considered identified.

In Table 3, we report the contributions of the oil price shock to the forecast error variance of the different variables at selected horizons. They are based on the modes of the impulse responses. Not surprisingly, the oil price shock accounts for the bulk of the forecast error variance of oil price at all horizons, and it is also important for prices.

In contrast, its importance for output and interest rate is relatively minor, with the other shocks accounting for the majority of the forecast error variance, especially at short horizons.

Table 3: Forecast error variance decomposition of the oil price shock in Peersman's (2005) model.

	Horizon (quarters)					
Variable	1	2	4	8	12	16
Oil Price	0.993	0.991	0.995	0.996	0.996	0.996
Output	0.004	0.012	0.038	0.162	0.224	0.248
Prices	0.390	0.478	0.503	0.496	0.482	0.467
Interest Rate	0.006	0.070	0.176	0.183	0.173	0.165

The figures are the proportions of the forecast error variance at each horizon accounted for by the oil price shock based on the modes of the impulse responses.

Our general conclusion is that all of Peersman's (2005) sign restrictions are not supported by the data. While it seems that they quite succesfully identify the oil price shock and potentially also the supply shock, the shocks labeled as the demand and monetary policy shocks fail to accord with the data. Similar conclusion were reached by Herwartz and Lütkepohl (2014), who analyzed these restrictions in the Markov-switching SVAR model of Lanne et al. (2010). However, compared to that paper, our results are more useful in that we were able to single out (in probability) the shocks that do conform with the data.

6 Conclusion

We have introduced a new Bayesian procedure for checking sign restrictions in structural VAR models. The procedure is based on the structural VAR model where, following Lanne et al. (2005), non-Gaussian and mutually independent errors are assumed. Under these assumptions, the structural shocks, and, hence, their impulse responses are (locally) uniquely identified, which facilitates checking the validity of any set of sign restrictions in a straightforward manner. Our contribution is twofold. First, we introduce a Bayesian implementation of the SVAR model under the assumptions of Lanne et al. (2005). Second, and more importantly, we show how the plausibility of sign restrictions can be quantified, and our methods can thus be seen as a formalization of the approaches proposed in the previous statistical identification literature (see, in particular, Lütkepohl and Netšunajev (2014)).

The impulse responses and forecast error variance decompositions of the economic shocks that are found identified with high probability, can then be computed using any of the conventional methods put forth in the literature. Having a uniquely identified SVAR model brings about two great advantages. First, the computations are much simpler than typically in the sign identifiication literature. Second, we avoid the so-called model identification problem, which facilitates straightforward interpretation of forecast error variance decompositions and reporting the results of impulse response analysis.

We illustrated the new methods by two empirical applications. We found support for the sign restrictions employed by Uhlig (2005) to identify the monetary policy shock in his U.S. data set. Moreover, while there was great uncertainty about the impact of this shock on the real GDP, its effect was found negative after the first few quarters, which is intuitively appealing. In Peersman's (2005) U.S. macroeconomic model, we were able to convincingly identify only two of the four shocks that he suggested to identify by sign restrictions. Of these, the mode of the impulse responses of the oil price shock also satisfied all the sign restrictions.

Our procedure could be extended to checking the validity of and discriminating between dynamic stochastic general equilibrium (DSGE) models. Recently, Canova and Paustian (2011) suggested a two-step procedure where a set of robust (sign) restrictions implied by a DSGE model (or multiple competitive DSGE models) is first imposed to identify a SVAR model, and the plausibility of another set of restrictions is then checked in this identified model. It is crucial that the first-step restrictions hold in the data, and reliability of the procedure increases with the number of identified shocks in the first step. Hence, our procedure could be employed to find the maximal set of robust restrictions. Furthemore, the probabilites of the second-step restrictions could subsequently be computed conditional on the shocks identified in the first step.

Appendix

In this appendix we provide a Metropolis-within-Gibbs algorithm for the estimation of the posterior distribution of the parameters of the SVAR model given in (1).

Let us start by describing the conditional likelihood function. We assume that the ith component ($i \in \{1, \ldots, n\}$) of the error ε_t follows Student's t distribution with λ_i degrees of freedom. For computational convenience, we reparameterize ε_{it} as $h_{it}^{-1/2}\eta_{it}$, where η_{it} is a standard normal random variable, and $\lambda_i h_{it}$ follows the chi-square distribution with λ_i degrees of freedom. Then, $\varepsilon_t = H_t^{-1/2}\eta_t$, where η_t is a $(n \times 1)$ vector of standard normal random variables, and $H_t = diag\left(h_{1t}, \ldots h_{nt}\right)$. From (1), a change of variable yields

$$p\left(\mathbf{y}\left|\mathbf{A}, B, \mathbf{H}\right.\right) \propto \left|\det\left(B^{-1}\right)\right|^T \prod_{t=1}^{T} \left|H_t\right|^{1/2} \exp\left[-\frac{1}{2}\sum_{t=1}^{T} u_t' B^{-1\prime} H_t B^{-1} u_t\right], \tag{A.1}$$

where $\mathbf{A} = \left[a, A_1', \ldots, A_p'\right]'$, $\mathbf{H} = diag\left(h_{11}, \ldots h_{n1}, \ldots, h_{1T}, \ldots h_{nT}\right)$, $u_t = y_t - a - A_1 y_{t-1} - \cdots - A_p y_{t-p}$, and $\mathbf{y}$ is a $(Tn \times 1)$ vector obtained by stacking y_t for $t = 1, \ldots, T$.

We operate on B^{-1}, the inverse of B, and to facilitate unique identification, we make two additional assumptions. First, we restrict the parameter space of $B^{-1} = (c_{ij})$ such that $|c_{jj}| > |c_{ij}|$ for all $i > j$. Second, we assume that the diagonal elements of B^{-1} are positive. In practice these conditions are imposed by multiplying the conditional likelihood by an indicator function, which equals unity if B^{-1} belongs to the defined space, and zero otherwise. Notice that because the likelihood function (and therefore the posterior) is invariant with respect to permutation of the rows of B^{-1} (columns of B), we can reorder the columns of the restricted posterior B matrices produced by Markov chain Monte Carlo simulation without changing the posterior model probabilities (see Geweke (2006) for discussion).

A Gaussian prior distribution is assumed for vec$(B^{-1}) = \mathbf{b}$, $\mathbf{b} \sim N\left(\underline{\mathbf{b}}, \underline{V}_{\mathbf{b}}\right)$, and we simulate from the conditional posterior of B^{-1} by an accept-reject Metropolis-Hastings (ARMH) algorithm (see, for example, Chib and Greenberg (1995)). To obtain a good proposal density for B^{-1}, we approximate the log conditional likelihood by the second order Taylor expansion around some $\widetilde{\mathbf{b}}$:

$$\begin{aligned}\log p\left(\mathbf{y}\left|\mathbf{A}, B, \mathbf{H}\right.\right) \approx\ & \log p\left(\mathbf{y}\left|\mathbf{A}, \widetilde{\mathbf{b}}, \mathbf{H}\right.\right) \\ & + \left(\mathbf{b} - \widetilde{\mathbf{b}}\right)\mathbf{f} - 0.5\left(\mathbf{b} - \widetilde{\mathbf{b}}\right)'\mathbf{G}\left(\mathbf{b} - \widetilde{\mathbf{b}}\right),\end{aligned}$$

where $\mathbf{f}$ and $\mathbf{G}$ are the gradient and the negative Hessian of the log conditional likelihood evaluated at $\widetilde{\mathbf{b}}$, respectively. Combining the above with the prior density yields

$$\log p\left(\mathbf{b}\,|\mathbf{A},\mathbf{H},\mathbf{y}\right) \approx -\frac{1}{2}\left[\mathbf{b}'\left(\mathbf{G}+\underline{V}_{\mathbf{b}}^{-1}\right)\mathbf{b}-2\mathbf{b}'\left(\mathbf{f}+\mathbf{G}\widetilde{\mathbf{b}}+\underline{V}_{\mathbf{b}}^{-1}\underline{\mathbf{b}}\right)\right],$$

which is a (log) kernel of a multivariate normal density. We construct the Taylor expansion around the mode, $\widetilde{\mathbf{b}}=\widehat{\mathbf{b}}$. At first, we make no additional assumptions concerning the space of B^{-1}, and thus the (local) posterior mode can be quickly obtained by the Newton-Raphson method, using explicit formulae for $\mathbf{f}$ and $\mathbf{G}$, and the current draw of $\mathbf{b}$ as an initial point (see, for example, Chan (2015)).[9]

If the resulting local mode does not satisfy the restrictions stated below the likelihood function (A.1), we replace $\widehat{\mathbf{b}}$ and $\mathbf{G}$ (evaluated at $\widehat{\mathbf{b}}$) with $\overline{\mathbf{b}}=(I_n\otimes DP)\,\widehat{\mathbf{b}}$ and $\overline{\mathbf{G}}=\left(I_n\otimes(DP)^{-1\prime}\right)\mathbf{G}\left(I_n\otimes(DP)^{-1}\right)$, respectively, where P is the permutation matrix for which the matrix $P\widehat{B^{-1}}=(\widehat{c}_{ij})$ satisfies $|\widehat{c}_{jj}|>|\widehat{c}_{ij}|$ for all $i>j$, and D is a diagonal matrix with elements equal to either 1 of -1 that transforms the diagonal elements of $P\widehat{B^{-1}}$ positive.

Because the latter transformation may change the signs of the rows of $\widehat{B^{-1}}$, it may result in a value of the proposal density which is virtually zero at the current draw of $\mathbf{b}$ (causing the proposal, say $\mathbf{b}^*$, to be rejected). Therefore, to improve the performance of the sampler, we use the following mixture of two multivariate normal densities as a

[9]For the derivation of the explicit formulae for $\mathbf{f}$ and $\mathbf{G}$, it may be convenient to express the log conditional likelihood as

$$\begin{aligned}\log p\left(\mathbf{y}\,|\mathbf{A},B^{-1},\mathbf{H}\right) &\propto T\log\left|\det\left(B^{-1}\right)\right| \\ &\quad -\frac{1}{2}\mathrm{vec}(B^{-1})'\left(\sum_{t=1}^{T}u_tu_t'\otimes H_t\right)\mathrm{vec}\left(B^{-1}\right).\end{aligned}$$

Then, it can be readily checked that

$$\begin{aligned}\mathbf{f} &= \frac{\partial}{\partial\mathrm{vec}(B^{-1})}\log p\left(\mathbf{y}\,|\mathbf{A},B^{-1},\mathbf{H}\right)|_{\mathbf{b}=\widetilde{\mathbf{b}}} \\ &= T\mathrm{vec}\left(B'\right)-\left(\sum_{t=1}^{T}u_tu_t'\otimes H_t\right)\mathrm{vec}\left(B^{-1}\right), \\ \mathbf{G} &= -\frac{\partial^2}{\partial\mathrm{vec}(B^{-1})\mathrm{vec}(B^{-1})'}\log p\left(\mathbf{y}\,|\mathbf{A},B^{-1},\mathbf{H}\right)|_{\mathbf{b}=\widetilde{\mathbf{b}}} \\ &= TK_{nn}\left(B'\otimes B\right)+\left(\sum_{t=1}^{T}u_tu_t'\otimes H_t\right),\end{aligned}$$

where K_{nn} is a $\left(n^2\times n^2\right)$ commutation matrix (see, e.g., Lütkepohl (1996, p. 115)).

proposal density in the ARMH algorithm:

$$q(\mathbf{b}) = \frac{1}{2}\mathbf{N}\left(\mathbf{b}\left|\widehat{\mathbf{b}}, \mathbf{G}\right.\right) + \frac{1}{2}\mathbf{N}\left(\mathbf{b}\left|\overline{\mathbf{b}}, \overline{\mathbf{G}}\right.\right).$$

In our pratical implementations of the algorithm, only a few draws from $q(\mathbf{b})$ are typically required in the accept-reject (AR) steps. Furhermore, the Metropolis-Hastings (MH) acceptance rates tend to vary between 0.85% and 0.99%. The reported results are based on $\underline{V}_{\mathbf{b}}^{-1} = c_{\mathbf{b}} I_{n^2}$, where $c_{\mathbf{b}} = 0$, which results in an uninformative prior for B^{-1}, $p(B^{-1}) \propto 1$. As a robustness check, we also considered some informative priors, and found no change in the results.

As far as the full conditional posterior of $\mathbf{A}$ is concerned, it is easy to check that the conditional likelihood (A.1) can also be expressed as

$$p(\mathbf{y}\,|\mathbf{A}, B, \mathbf{H}) \propto \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{X}\text{vec}(\mathbf{A}))' \mathbf{\Omega}(\mathbf{y} - \mathbf{X}\text{vec}(\mathbf{A}))\right],$$

where $\mathbf{X}$ is obtained by stacking $(I_n \otimes X_t)$ for $t = 1, \ldots, T$, $X_t = \left(1, y'_{t-1}, \ldots, y'_{t-p}\right)$, and $\Omega = (I_T \otimes B'^{-1}) H (I_T \otimes B^{-1})$. Assuming a multivariate normal prior for $\text{vec}(\mathbf{A}) = \mathbf{a}$, $\mathbf{a} \sim N(\underline{\mathbf{a}}, \underline{V}_{\mathbf{a}})$, we then obtain

$$\text{vec}(\mathbf{A})\,|B, \mathbf{H}, \mathbf{y} \sim \mathbf{N}\left(\overline{\mathbf{a}}, \overline{V}_{\mathbf{a}}\right),$$

where $\overline{V}_{\mathbf{a}}^{-1} = \underline{V}_{\mathbf{a}}^{-1} + X'\Omega X$, and $\overline{\mathbf{a}} = \overline{V}_{\mathbf{a}}\left(\underline{V}_{\mathbf{a}}^{-1}\underline{\mathbf{a}} + \mathbf{X}'\mathbf{\Omega}\mathbf{y}\right)$. The precision-based sampling method of Chan and Jeliazkov (2009) can be used to simulate draws from $N\left(\overline{\mathbf{a}}, \overline{V}_{\mathbf{a}}\right)$ efficiently. In this paper, we set $\overline{\mathbf{a}} = 0$ and $\overline{V}_{\mathbf{a}} = 10000^2 I_{pn^2+n}$. However, we also checked the results using some informative priors, including close variants of the prior proposed by Baumeister and Hamilton (2015). The results remained intact irrespective of the priors used.

We now turn to the sampling of the latent variables $\{h_{1t}, \ldots h_{nt}\}_{t=1}^{T}$. The log conditional likelihood is proportional to

$$\begin{aligned}
\log p(\mathbf{y}\,|\mathbf{A}, B, \mathbf{H}) &\propto \sum_{t=1}^{T} \log |H_t|^{1/2} - \frac{1}{2}\sum_{t=1}^{T} u'_t B^{-1\prime} H_t B^{-1} u_t \\
&= \sum_{t=1}^{T}\left[\sum_{i=1}^{n} \log h_{it}^{1/2} - \frac{1}{2}\varepsilon'_t H_t \varepsilon_t\right] \\
&= \sum_{t=1}^{T}\left[\sum_{i=1}^{n}\left(\log h_{it}^{1/2} - \frac{1}{2} h_{it}\varepsilon_{it}^2\right)\right],
\end{aligned}$$

where $\varepsilon_t = B^{-1}u_t$ and $u_t = y_t - a - A_1 y_{t-1} - \cdots - A_p y_{t-p}$. Recall that the (hierarchical) prior of each $\lambda_i h_{it}$ is the chi-squared distribution with λ_i degrees of freedom. Then, by multiplying $p(\mathbf{y}\,|\mathbf{A},B,\mathbf{H})$ by the product of the prior densities of h_{it} for $t \in \{1,\ldots,T\}$ and $i \in \{1,\ldots,n\}$ we obtain

$$p(h_{it}\,|\mathbf{A},B,\lambda,\mathbf{y}) \propto h_{it}^{(\lambda_i-1)/2} \exp\left\{-\left[\lambda_i + \varepsilon_{it}^2\right] h_{it}/2\right\},$$

where $\lambda = (\lambda_1,\ldots,\lambda_n)$. This implies that each h_{it} $(t \in \{1,\ldots,T\}, i \in \{1,\ldots,n\})$ can be sampled from the chi-square distribution as follows:

$$\left[\lambda_i + \varepsilon_{it}^2\right] h_{it}\,|\mathbf{A},B,\lambda,\mathbf{y} \sim \chi(\lambda_i + 1).$$

We assume an exponential prior distribution for each λ_i, $\lambda_i \sim Exp(\underline{\lambda}_i)$. From the hierarchical prior density of h_{it} $(t \in \{1,\ldots,T\})$ and the assumption $\lambda_i \sim Exp(\underline{\lambda}_i)$, it follows that the conditional posterior density of λ_i can be written as proportional to

$$\begin{aligned} p\left(\lambda_i \middle| \{h_{it}\}_{t=1}^{T}, \mathbf{y}\right) \;\propto\; & \left[2^{\lambda_i/2}\Gamma(\lambda_i/2)\right]^{-T} \lambda^{\lambda T/2} \left(\prod_{t=1}^{T} h_{it}^{(\lambda_i-2)/2}\right) \\ & \times \exp\left[-\left(\frac{1}{\underline{\lambda}_i} + \frac{1}{2}\sum_{t=1}^{T} h_{it}\right)\lambda_i\right]. \end{aligned}$$

It is the hierarchical prior structure in which each λ_i affects the data only through $\{h_{it}\}_{t=1}^{T}$ that lies behind this result. Following Geweke (2005), we simulate from the conditional posterior of the degree-of-freedom parameter λ_i using an independence-chain MH algorithm. As a candidate distribution of λ_i, we use the univariate normal distribution with mean equal to the mode of the log conditional posterior, and the precision parameter equal to the negative of the second derivative of the log posterior density evaluated at the mode.

References

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

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

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

Chan, J. (2015). The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. *Journal of Business and Economic Statistics* (forthcoming).

Chan, J. and I. Jeliazkov (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation* 1, 101-120.

Chib, S., and E. Greenberg (1995). Understranding the Metropolis-Hastings algorithm. *The American Statistician* 49, 327-335.

Faust, J. (1998). The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy* 49, 207–244.

Fry, R., and A. Pagan (2011). Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature* 49, 938–960.

Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics*. Wiley: Hoboken, New Jersey.

Geweke, J. (2007). Interpretation and inference in mixture models: Simple MCMC works. *Computational Statistics & Data Analysis* 51, 3529-3550.

Herwartz, H., and H. Lütkepohl (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks. *Journal of Econometrics* 183, 104–116.

Inoue, A., and L. Kilian (2013). Inference on impulse response functions in structural VAR models. *Journal of Econometrics* 177, 1–13.

Kilian, L. (2013). Structural vector autoregressions, in Hashimzade, N., and M.A. Thornton (eds.), *Handbook of Research Methods and Applications in Empirical Macroeconomics*, Edward Elgar: Cheltenham, 515–554.

Kilian, L., and D. Murphy (2009). Why agnostic sign restrictions are not enough: Understanding the dynamics of the oil market VAR models. *Journal of the European Economic Association* 10, 1166–1188.

Koop, G. (2003). *Bayesian Econometrics*. Wiley: Chichester.

Lanne, M., H. Lütkepohl, and K. Maciejowska (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control* 34, 121–131.

Lanne, M., M. Meitz, and P. Saikkonen (2015). Identification and estimation of non-Gaussian structural vector autoregressions. CREATES Research Paper 2015-16.

Lütkepohl, H. (1996). *Handbook of Matrices*. Wiley: Chichester.

Lütkepohl H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag: Berlin.

Lütkepohl H., and A. Netšunajev (2014). Disentangling demand and supply shocks in the crude oil market: How to check sign restrictions in structural VARs, *Journal ofApplied Econometrics* 29, 479–496.

Peersman, G. (2005). What caused the early millennium slowdown? Evidence based on vector autoregressions. *Journal of Applied Econometrics* 20, 185–207.

Straub, R., and G. Peersman (2006). Putting the New Keynesian model to a test. *International Monetary Fund Working Paper* 06/135.

Piffer, M. (2015). Assessing identifying restrictions in SVAR models. Unpublished manuscript. DIW Berlin.

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

Waggoner, D.,F. and T. Zha (2003). Likelihood preserving normalization in multiple equation models. *Journal of Econometrics* 114, 329–347.