University of Konstanz
Department of Economics

Identification of SVAR Models by Combining Sign Restrictions With External Instruments

Robin Braun & Ralf Brüggemann

Working Paper Series
2020-01

https://ideas.repec.org/s/knz/dpteco.html

Identification of SVAR Models by Combining Sign Restrictions With External Instruments*

Robin Braun[†] Ralf Brüggemann[‡]

May 15, 2020

Abstract

We discuss combining sign restrictions with information in external instruments (proxy variables) to identify structural vector autoregressive (SVAR) models. In one setting, we assume the availability of valid external instruments. Sign restrictions may then be used to identify further orthogonal shocks, or as an additional information on the shocks identified by the external instruments. In the latter case, the additional restrictions may be overidentifying and checked against the data. In a second setting, we assume that proxy variables are only 'plausibly exogenous'. In this case, various inequality restrictions based e.g. on correlations or variance contributions can be used for set-identification. This can be combined with conventional sign restrictions to further narrow down the set of admissible models. For our B-model type Proxy SVAR setup, we develop Bayesian inference and discuss the computation of Bayes factors to check overidentifying restrictions. We illustrate the usefulness of our methodology in estimating the effects of oil market and monetary policy shocks.

Keywords: Structural vector autoregressive model, sign restrictions, external instruments, Proxy VAR

JEL classification: C32, C11, E32, E52

*We thank participants of the Konstanz-Tübingen-Hohenheim Econometrics Seminar, the CFE-CMStatistics 2017, the 2017 annual meeting of the German Statistical Society as well as seminar participants at University of Melbourne, Monash University, National University of Córdoba, Universität Duisburg-Essen, and FAU Universität Erlangen-Nürnberg for useful comments. We also thank Gökhan Ider for excellent research assistance. Part of this research has been conducted when the second author was visiting Monash University, Australia. Financial support of the Graduate School of Decision Sciences (GSDS) at the University of Konstanz and the German Science Foundation (grant number: BR 2941/3-1) is gratefully acknowledged.

†Robin Braun: Bank of England, Centre for Macroeconomics, email: robin.braun@bankofengland.co.uk

‡Corresponding author: Ralf Brüggemann: University of Konstanz, Department of Economics, Box 129, 78457 Konstanz, Germany, email: ralf.brueggemann@uni-konstanz.de

1 Introduction

Structural vector autoregressive (SVAR) models have become a popular tool in applied macroeconomics for investigating the importance of different structural shocks by impulse responses, forecast error variance or historical decompositions. A key challenge in SVAR analysis is the identification of structural shocks and a number of different approaches have been used in the literature. This includes traditional short- and long-run (exclusion) restrictions, identification by some form of heteroskedasticity, sign restrictions on impulse responses, and the use of external instruments.[1] In this paper, we discuss how to combine the latter two, i.e. we discuss how to identify SVAR models by combining sign restrictions with information in time series that act as proxy or external instrumental variables for the structural shocks of interest. We argue that combining both approaches can be useful in many situations to sharpen identification and mitigate some drawbacks that may occur when using either sign restrictions or external instruments only.

Sign restrictions have been introduced into the SVAR literature by Faust (1998), Canova & De Nicoló (2002) and Uhlig (2005) as an alternative to existing methods involving controversial short- and long-run exclusion restrictions on the effect of structural shocks. In their most common form, they are inequality restrictions imposed on contemporaneous or higher horizon responses to the structural shocks of interest. More broadly, they have been exploited to bound other (functions of) structural parameters of the model, e.g. elasticities or variance decompositions. In the context of monetary policy shocks, for instance, sign restrictions have been used to avoid the so-called 'price-puzzle' by restricting the response of the price level to be non-positive for a certain period after a contractionary monetary policy shock but leaving the response of interest (e.g. the response of output) unrestricted. Employing sign restrictions leads to set identification only. An important practical problem of sign restrictions is that they are often rather weak, resulting in a wide range of admissible models with impulse responses that are not very informative.

Identification by external instruments provides another popular alternative method for identifying structural shocks. While the underlying economic shock of interest is unobservable to the researcher, there may be related time series (proxy variables) available that act

[1] See e.g. Kilian & Lütkepohl (2017) for an overview of different SVAR models and their applications.

as instrumental variables (IV) for the unobserved structural shock. The resulting model is often called a Proxy VAR model and many papers have successfully exploited this identification strategy, including Stock & Watson (2012), Mertens & Ravn (2012), Gertler & Karadi (2015), Gerko & Rey (2017), Mertens & Montiel Olea (2018), Lakdawala (2019), Känzig (2019) and Peersman (2020). While conceptually appealing, the external IV approach has also potential drawbacks. First, the exogeneity of instruments is questionable in many applications (see e.g. the discussion in Ramey (2016) on the narrative measures of monetary policy shocks). Furthermore, even a proxy variable that is truly exogenous may be a weak instrument, which complicates reliable inference substantially (Montiel Olea et al. 2020).

In this paper, we contribute to the literature by discussing how to combine the proxy variable approach with sign restrictions. In a first setting, we assume that credibly exogenous instruments are available for some structural shocks of interest. In this case, sign restrictions can be used for two purposes. On the one hand, sign restrictions may identify additional shocks from the group of shocks that are orthogonal to those identified by the external instruments. On the other hand, sign restrictions can be imposed in addition to the IV conditions such that they are informative with respect to the shocks for which instruments are available. For instance, sign restrictions can be helpful to disentangle multiple shocks that are to be identified by external instruments (see e.g. Piffer & Podstawski (2017)). Moreover, if the external variables are only weakly informative, sign restrictions may offer an extra piece of identifying information. Additional sign restrictions on these shocks may be overidentifying. As one example, consider the case of one structural shock identified by one valid external instrument. Then, the respective structural parameters are point-identified by the IV conditions and additional sign restrictions on this shock are overidentifying. In this paper, we propose to use Bayes factors as a statistical tool to check the validity of the additional restrictions.

In a second setting, we assume the availability of relevant but not necessarily exogenous proxy variables. In the microeconometric literature, these are known as 'plausibly exogenous' instruments (Conley et al. 2012), a terminology which we adapt in the remainder of this paper. Relaxing the corresponding exogeneity conditions instantly leads to a loss of point identification. However, as in the microeconometric literature, inequality restrictions can be used to bound the amount of endogeneity and to obtain a set-identified model, which may still be informative about the underlying structural relationships. In our context, the researcher can choose different types of inequality restrictions. We propose to incorporate various forms of prior information on the relation of the structural shock and the 'plausibly exogenous' proxy variables. For example, this may come from a stance on the sign of their correlation with the shocks or from using lower bounds on variance contributions. Furthermore, these restrictions can be easily combined with conventional sign restrictions on the responses of variables to further reduce the set of admissible models. One key advantage of this combination approach is that the additional information in the proxy variables helps to avoid wide and uninformative impulse response intervals often observed from SVARs using sign restrictions only.

To conduct inference, we rely on a unified econometric framework, a Bayesian SVAR model augmented by equations for the proxy variables. We formulate priors and posteriors for a B-model type SVAR, i.e. we use a model where the reduced form errors u_t are modeled as a linear function of the structural shocks ε_t, that is $u_t = B\varepsilon_t$.[2] We summarize the posterior distribution of the structural parameters by Markov Chain Monte Carlo (MCMC) methods. In order to sample from the conditional distribution of the structural parameters B, we use an Accept Reject Metropolis Hastings (AR-MH) algorithm (Chib & Greenberg 1995). The AR-MH explores the set-identified posterior efficiently exploiting the importance distribution developed in Arias et al. (2018, 2019). Finally, we discuss estimation of Bayes factors, which provides a formal statistical tool to check overidentifying restrictions against the data.

Our paper is related to an emerging literature that has discussed some form of combining sign restrictions with external instruments specifically, or non-model information more broadly. Related to our first setting, Cesa-Bianchi & Sokol (2017) discuss identification of an additional shock by sign restrictions in a Proxy SVAR. Furthermore, Jarociński & Karadi (2020) uses sign restrictions to disentangle two shocks assumed to be correlated

[2]The basic structure of our Proxy SVAR is of the same form as the one used in Angelini & Fanelli (2019). As explained in Section 2, this setup is labeled as a B-model SVAR in some parts of the literature (see e.g. Lütkepohl (2005, Chapter 9)).

with one external instrument. Both identification strategies can be handled within our inference framework, which provides a coherent way to formulate priors directly on the structural parameters of interest. Finally, Nguyen (2019) also suggests combining external instruments with sign restrictions on the corresponding structural parameters, and further uses Bayes factors to test these restrictions. In contrast to our paper, his approach relies on the framework of Baumeister & Hamilton (2015), and hence the formulation of prior distributions on structural parameters of $\mathrm{A} = B^{-1}$.

As mentioned above, our second setting is closely related to the microeconometric literature on plausibly exogenous instruments, which suggests to bound the degree of endogeneity to obtain a set-identified simultaneous equation model (Nevo & Rosen 2012, Conley et al. 2012). Another closely related paper is Ludvigson et al. (2020) introducing so-called 'external inequality constraints'. Essentially, their approach entails discarding models in which the shocks that are not or only loosely correlated with the proxy variables.[3] Our approach is more general with respect to important modeling aspects: For instance, we not only discuss discarding models based on correlations but also consider a variety of restrictions (for instance based on variance contributions) that may reflect different degrees of confidence a researcher has about the relation of the proxy and the structural shock of interest. Furthermore, we also propose methods that do not require the choice of a threshold on the correlation between the shocks and external variables. Finally, the Bayesian inference framework that we develop in this paper allows to properly take into account all sources of model and estimation uncertainty. There are several other papers that are more broadly related to the idea of exploiting non-model information to sharpen identification in set-identified SVAR models. For instance, Kilian & Murphy (2012) and Baumeister & Hamilton (2019) demonstrate how to incorporate microeconomic evidence on elasticities to sharpen the set of identified SVAR models. Moreover, Antolín-Díaz & Rubio-Ramírez (2018) and Zeev (2018) combine sign restrictions with narrative evidence on the sign of structural shocks and their historical decompositions.

Methodologically our paper also relates to the recent literature developing Bayesian Proxy VARs (Caldara & Herbst 2019, Arias, Rubio-Ramírez & Waggoner 2019, Giacomini

[3]See also Uhrin & Herwartz (2016) for a similar idea.

et al. 2019, Nguyen 2019).[4] In contrast to these papers, we consider inference in the B-model representation of the proxy-augmented SVAR model. As surveyed in Bruns & Piffer (2019), the B-model type SVAR is the most popular among applied researchers. Furthermore, we discuss the computation of Bayes factors to test the plausibility of overidentifying restrictions in B-models. To the best of our knowledge, both modeling aspects are new to the literature. Finally, in contrast to the aforementioned papers, we consider independent prior distributions on the reduced form autoregressive coefficients. While this comes with the need to rely on Markov Chain Monte Carlo methods for efficient inference, it allows to impose a wider spectrum of prior information including asymmetric priors across equations. This includes the original Minnesota prior of Litterman (1986) as well as a variety of hierarchical shrinkage priors (Koop et al. 2010).

We illustrate the usefulness of our method in two empirical applications. In the first application, we revisit a benchmark SVAR model for the global market of crude oil, which is typically identified by a combination of sign restrictions and elasticity constraints (Kilian & Murphy 2014, Baumeister & Hamilton 2019). We demonstrate how additional identifying information from external instruments can be used to check competing oil supply elasticity constraints from the literature against the data. In our second application, we use a combination of sign restrictions and additional information in the Romer & Romer (2004) narrative measure to identify the effects of monetary policy shocks in the United States. Given that the literature has doubts about the exogeneity of this narrative shock, we illustrate our second combination approach, which entails relaxing the exogeneity condition. Using our method, we no longer find puzzling results and, at the same time, more informative impulse response functions.

The remainder of the paper is structured as follows. Section 2 introduces the econometric modeling framework, discusses identifying restrictions, Bayesian inference as well as the estimation of Bayes factors. Section 3 illustrates the suggested methods in applications to oil market shocks and US monetary policy shocks. Section 4 summarizes and concludes.

[4]See also Drautzburg (2016) for an earlier contribution.

2 Methodology

In the following, we introduce the econometric framework used to identify SVARs by combining information from sign restrictions and proxy variables. We start by describing a SVAR model augmented by equations for the dynamics of the external proxy variables in Section 2.1. Section 2.2 discusses combining sign restrictions with proxy variables when the exogeneity restrictions of the external variables are assumed to be valid. In Section 2.3, we discuss the setting where we exploit information from proxy variables that are just 'plausibly exogenous', i.e. we cover proxy variables whose variation may partly reflect endogenous components. Bayesian inference in the model is discussed in Section 2.4 and Section 2.5 discusses the estimation of Bayes Factors.

2.1 Augmented SVAR model

We consider a B-model type SVAR model (see e.g. (Lütkepohl 2005, Section 9.1)) given by

$$y_t = \nu + \sum_{i=1}^{p} A_i y_{t-i} + B\varepsilon_t, \quad \varepsilon_t \sim (0, I_n), \tag{2.1}$$

where $y_t = (y_{1t}, \dots, y_{nt})'$ is a $n \times 1$ vector of endogenous time series, ν is a $n \times 1$ vector of intercepts, and $A_i, i = 1, \dots, p$ are $n \times n$ matrices of autoregressive coefficients. The dynamics of the system is assumed to be driven by n structural shocks ε_t, where we assume that the elements of ε_t are orthogonal (contemporaneously uncorrelated) and are normalized to have unit variances. The $n \times n$ matrix B is the contemporaneous impact matrix and reflects the immediate responses of the variables y_t to the structural shocks ε_t. We assume stability of the VAR such that $\det A(z) = \det(I_K - A_1 z - \dots - A_p z^p) \neq 0$ for $|z| \leq 1$. This implies that the SVAR(p) has a MA(∞) representation given by $y_t = \mu_y + \sum_{j=0}^{\infty} \Xi_j B\varepsilon_{t-j} = \mu_y + \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j}$, where $\mu_y = \mathrm{E}(y_t)$ and the $n \times n$ coefficient matrices $\Theta_j = \Xi_j B$, are the structural impulse response functions (IRFs). The reduced form MA(∞) matrices Ξ_j can be computed recursively from $\Xi_j = \sum_{i=1}^{j} \Xi_{j-i} A_i$ with $\Xi_0 = I_n$.

Without additional restrictions this model is not identified. To see this, denote by $u_t = B\varepsilon_t$ the VAR forecast errors with corresponding reduced form covariance matrix $\mathrm{E}(u_t u_t') = \Sigma_u = BB'$. For any matrix $Q \in \mathcal{O}(n)$ of the orthogonal group $\mathcal{O}(n) = \{Q \in$

$\mathcal{R}^{n\times n} : QQ' = I_n\}$, an observationally equivalent model $\bar{B} = BQ$ can be obtained capable to generate the same reduced form dynamics.[5] Therefore, restrictions must be imposed on the structural impact matrix B in order to pin down a meaningful structural model.

In this paper, we focus on identification by combining sign restrictions with information in external variables. Let $m_t = (m_{1t}, \ldots, m_{kt})'$ be a $k \times 1$ vector of external variables designed to provide identifying information about a subset of $k < n$ structural shocks. Our econometric methods are based on augmenting the SVAR given in (2.1) by equations for m_t:

$$\underbrace{\begin{pmatrix} y_t \\ m_t \end{pmatrix}}_{\tilde{y}_t} = \underbrace{\begin{pmatrix} \nu \\ \nu_m \end{pmatrix}}_{\tilde{\nu}} + \sum_{i=1}^{p} \underbrace{\begin{pmatrix} A_i & 0_{n\times k} \\ \Gamma_{1i} & \Gamma_{2i} \end{pmatrix}}_{\tilde{A}_i} \underbrace{\begin{pmatrix} y_{t-i} \\ m_{t-i} \end{pmatrix}}_{\tilde{y}_{t-i}} + \underbrace{\begin{pmatrix} B & 0_{n\times k} \\ \Phi & \Sigma_\eta^{1/2} \end{pmatrix}}_{\tilde{B}} \underbrace{\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}}_{\tilde{\varepsilon}_t}, \qquad \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim (0, I_{n+k}). \tag{2.2}$$

As noted in Mertens & Ravn (2012), the additional equations have an intuitive measurement error interpretation. The k variables m_t are modeled as a linear function of lagged values of $\tilde{y}_t$, the structural errors ε_t, plus a zero mean measurement error η_t, which is assumed to be orthogonal to the structural shocks ε_t, i.e. $\eta_t \perp \varepsilon_t$. Γ_{1i}, Γ_{2i} and Φ are $k \times n$ coefficient matrices. Corresponding $n \times k$ blocks of zeros in the upper right parts of $\tilde{A}_i$ and $\tilde{B}$ ensure that m_t and the measurement error η_t are external to the model and have no implications for the dynamics of y_t. We also assume that $\tilde{B}$ has full rank, $\text{rk}(\tilde{B}) = n + k$, throughout the paper. Usually, proxy variables are designed to be unpredictable by lagged values of y_t and m_t, and do only contain contemporaneous information about ε_t. In this case, one can set $\Gamma_{1i} = \Gamma_{2i} = 0$, and the model shares the more natural representation introduced in Mertens & Ravn (2012). To keep notation simple, for the remainder of this section, we assume $\Gamma_{1i} = \Gamma_{2i} = 0$ holds, implying that the model reduces to

$$y_t = \nu + \sum_{i=1}^{p} A_i y_{t-i} + B\varepsilon_t, \tag{2.3}$$

$$m_t = \nu_m + \Phi\varepsilon_t + \Sigma_\eta^{1/2}\eta_t. \tag{2.4}$$

Without any further restrictions, it is straightforward to show that the augmented SVAR

[5] This yields the same reduced form covariance matrix $\Sigma_u = (BQ)(BQ)' = BQQ'B' = BB'$.

model is only identified up to orthogonal rotations of the form $\bar{B} = \tilde{B}Q$, where we now have $Q = \text{diag}(Q_1, Q_2)$, $Q_1Q_1' = I_n$ and $Q_2Q_2' = I_k$. The block structure of Q reflects the fact that the measurement error η_t is assumed to be orthogonal to the dynamics in y_t, implying a $n \times k$ block of zeros in the upper right part of $\tilde{B}$. Through restrictions on Φ, identifying information can be imposed to pin down values of B or equivalently, to narrow down values of Q_1.

Before discussing our proposed restrictions in more detail, we quickly relate the augmented model to inference frameworks discussed in Arias, Rubio-Ramírez & Waggoner (2019) and Caldara & Herbst (2019). Note that in both papers, the SVAR model is augmented by equations for external variables and inference is conducted in a Bayesian way. In the following, we will use the SVAR notation of Lütkepohl (2005). Then, our model could be seen as an augmented B-model where, abstracting from lags, we assume that $\tilde{y}_t = \tilde{B}\tilde{\varepsilon}_t$. In turn, Arias, Rubio-Ramírez & Waggoner (2019) discusses inference for a proxy-augmented SVAR of the form $\tilde{A}_0\tilde{y}_t = \tilde{\varepsilon}_t$ where $\tilde{A}_0 = \begin{pmatrix} A_0 & 0 \\ A_{02} & A_{03} \end{pmatrix}$ has a lower triangular block structure. Therefore, their model can be best thought of an augmented A-model in the terminology of Lütkepohl (2005). Finally, the model of Caldara & Herbst (2019) can be written as an AB-model, writing $\tilde{A}_0\tilde{y}_t = \tilde{B}\tilde{\varepsilon}_t$ and setting $\tilde{A}_0 = \begin{pmatrix} A_0 & 0 \\ 0 & I_k \end{pmatrix}$ and $\tilde{B} = \begin{pmatrix} I_n & 0 \\ \Phi & \Sigma_\eta^{\frac{1}{2}} \end{pmatrix}$. Which model representation is more useful in practice is highly application specific and depends on whether restrictions are more naturally imposed on B or B^{-1}. As surveyed by Bruns & Piffer (2019), the B-model is very popular in applied SVAR analysis. In particular, Bruns & Piffer (2019) look at publications in the top-five journals and the Journal of Monetary Economics. They find that out of all papers involving SVAR models, 76% use the B-model representation.

2.2 Combining sign restrictions with instrumental variables restrictions

We first discuss combining sign restrictions with instrumental variable (IV) restrictions, assuming that the external variables m_t are valid exogenous instruments. For this purpose, partition the structural shocks ε_t and the matrix Φ as

$$\varepsilon_t = [\underset{1\times q}{\varepsilon'_{1t}} : \underset{1\times(n-k-q)}{\varepsilon'_{2t}} : \underset{1\times k}{\varepsilon'_{3t}}]' \quad \text{and} \quad \Phi = [\underset{k\times q}{\phi_1} : \underset{k\times(n-q-k)}{\phi_2} : \underset{k\times k}{\phi_3}]. \tag{2.5}$$

Without loss of generality, assume that out of all n structural shocks, the researcher identifies the last k shocks (ε_{3t}) using k instrumental variables m_t. In our model, $\mathrm{E}(m_t\varepsilon'_t) = \Phi$ and using the partitioning in (2.5), we find

$$[\mathrm{E}(m_t\varepsilon'_{1t}) : \mathrm{E}(m_t\varepsilon'_{2t}) : \mathrm{E}(m_t\varepsilon'_{3t})] = [\phi_1 : \phi_2 : \phi_3].$$

The assumption that m_t are valid instruments for ε_{3t}, implies that m_t is correlated with ε_{3t} but uncorrelated with all other shocks in the system, that is $\mathrm{E}(m_t\varepsilon'_{1t}) = 0$ and $\mathrm{E}(m_t\varepsilon'_{2t}) = 0$. Consequently, the IV conditions imply

$$[\phi_1 : \phi_2] = 0_{k\times n-k}, \tag{2.6}$$

and

$$\phi_3 \neq 0, \quad \mathrm{rk}(\phi_3) = k, \tag{2.7}$$

where (2.6) and (2.7) are the exogeneity and relevance conditions, respectively. The amount of information in m_t can be conveniently quantified in terms of the reliability matrix $\Lambda = \Sigma_\eta^{-1}\phi_3\phi'_3$. For $k = 1$, this yields a scalar corresponding to the fraction of variance in the instrument m_t explained by ε_{3t}. For $k > 1$, the eigenvalues could be used for interpretation, yielding the corresponding fractions of variance explained by the principal components of m_t (Mertens & Ravn 2013). If $k = 1$, the scalar shock of interest (ε_{3t}) is point identified by the external instrument conditions, while for any $k > 1$, restrictions (2.7) and (2.6) do only partition the structural shocks into shocks ε_{3t} that correlate with the instruments, and shocks ε_{1t} and ε_{2t}, which are orthogonal to the instruments. Therefore, for $k > 1$, additional restrictions are necessary to identify disentangle the effects of each component in ε_{3t}.

Next, we discuss how additional sign restrictions can be useful in this setting. First, sign restrictions can be used to identify q additional shocks ε_{1t} within the same model. Let $B = [\underset{n\times q}{B_1} : \underset{n\times n-k-q}{B_2} : \underset{n\times k}{B_3}]$ be a partitioning of the impact matrix corresponding to (2.5). Consequently, the impact effects of shocks ε_{1t} are given in B_1. In this case, sign restrictions need to be imposed directly on B_1 or on functions of B_1 and $A_i, i = 1, \ldots, p$, such as impulse response functions, forecast error variance and historical decompositions. Within the unified framework of this paper, the additional identified shocks ε_{1t} are guaranteed to be orthogonal to those shocks identified by IV restrictions. Sign restrictions may also be imposed on ε_{3t}, the shocks identified by external instruments. There are two reasons why this can be useful. First, if more than one shock is to be identified by IV, sign restrictions can be imposed to further disentangle each component within ε_{3t}. For example, Piffer & Podstawski (2017) use sign restrictions on ϕ_3 when $k = 2$. Alternatively, they can be imposed on the impact matrix B_3 (see e.g. Bertsche (2019)). Second, if the external variables are only weakly informative, sign restrictions may offer additional identifying information. Then, by combing the external IV with 'internal' sign restrictions, the need for non-standard inference procedures can be mitigated. One interesting sign restriction in this direction has been proposed in Arias, Rubio-Ramírez & Waggoner (2019), who impose a lower bound on the eigenvalues of Λ, ruling out instrument irrelevance *a priori*. As we will demonstrate in the empirical application in Section 3.1, the combination of IV with sign restrictions can be very useful to further sharpen identification. Note that whenever the additional restrictions are overidentifying, they may also be checked against the data. In our Bayesian framework this can be done in form of Bayes factors and we discuss how to compute them in our framework in Section 2.5.

2.3 Combining sign restrictions with information of plausibly exogenous instruments

There are many situations in which the researcher is less confident about the exogeneity of the external variable. In the following, we discuss ways how to use these variables for identification even if they are not convincingly exogenous. In reference to the microeconometric literature, we adopt the terminology and call these proxy variables 'plausibly exogenous' (cf. Conley et al. (2012)). As discussed in Conley et al. (2012), relaxing the exogeneity constraints leads to a loss of point identification. However, knowledge on the relation between the shocks and proxy variables can still be exploited to bound the set of admissible models via appropriate inequality restrictions. Furthermore, these can be combined with more conventional sign restrictions on structural parameters of the SVARs to further narrow down the set of parameters.

For ease of exposition, we discuss the typical situation where we want to identify a single shock, say ε_{1t} ($q = 1$) or equivalently B_1, the first column of the impact matrix. Furthermore, assume that we have a scalar ($k = 1$) proxy variable m_t designed for ε_{1t}, which is only 'plausibly exogenous' such that the approach in Section 2.2 cannot be used in a credible way. In what follows, we suggest various restrictions that bound the relation between the proxy variable m_t and the structural shock of interest ε_{1t}. For simplicity, we outline the restrictions for the case with one proxy and one shock, although they can easily be adapted to more general settings. We propose the following restrictions on the relation between m_t and the structural shock ε_{1t}:

1. Retain all models where the correlation between m_t and the structural shock to be identified ε_{1t} is positive, i.e. keep models if

$$\mathrm{Corr}(m_t, \varepsilon_{1t}) = \frac{\mathrm{E}(m_t \varepsilon_{1t})}{\sqrt{\mathrm{Var}(m_t)}} > 0.$$

   From an economic point of view, this means that we are confident that the proxy variable m_t is at least positively correlated with the structural shock it has been designed for. Note that this restriction does little harm if the proxy is only loosely associated with the structural shock and might be interesting for variables, which are assumed to be only weakly informative for ε_{1t}.

2. Retain all models where the correlation between m_t and the structural shock exceeds a threshold $\bar{c}$, i.e. keep models if

$$\mathrm{Corr}(m_t, \varepsilon_{1t}) = \frac{\mathrm{E}(m_t \varepsilon_{1t})}{\sqrt{\mathrm{Var}(m_t)}} > \bar{c}_1.$$

This restriction has been applied by Ludvigson et al. (2020). Of course, it is more restrictive than only a sign restriction on the correlation in that it rules out more models from the set of admissible SVARs. However, in our view, choosing $\bar{c}_1$ is difficult and hard to justify in practice.

3. Retain all models for which the variance of m_t explained by the structural shock ε_{1t} exceeds a threshold $\bar{c}_2$. To implement this restriction, recall that with $k = 1$ the measurement error equation for m_t is

$$m_t = \nu_m + \Phi \varepsilon_t + \sigma_\eta \eta_t, \qquad \eta_t \sim (0, 1),$$

where Φ is a $1 \times n$ vector. Since the regressors ε_t are orthogonal by assumption, the contribution of the jth structural shock to the variance of the proxy m_t is $\omega_j = \phi_j^2 / \mathrm{Var}(m_t)$, where ϕ_j is jth element of Φ. Therefore, one would keep models whenever

$$\omega_1 > \bar{c}_2. \tag{2.8}$$

To circumvent the problem of choosing the threshold value $\bar{c}_2$, we also suggest the following alternatives:

4. Keep only models where the identified shock of interest ε_{1t} shows a larger correlation with the proxy m_t than any other shock in the system, i.e. keep models if

$$\mathrm{Corr}(m_t, \varepsilon_{1t}) > \mathrm{Corr}(m_t, \varepsilon_{jt}), \qquad j = 2, \ldots, n. \tag{2.9}$$

5. Keep only models in which the identified shock of interest ε_{1t} explains more of the

variation in m_t than any other orthogonal component of the system, i.e. keep models if

$$\omega_1 > \omega_j, \qquad j = 2, \ldots, n. \tag{2.10}$$

In words, this discards all SVAR models where other shocks explain more variation of m_t than the one to be identified (ε_{1t}).

6. Keep only models in which the identified shock of interest ε_{1t} explains more of the variation in m_t than all other shock in the system together, i.e. keep models if

$$\omega_1 > \sum_{j=2}^{n} \omega_j. \tag{2.11}$$

Note that some of the suggestions that we make require identification of the other orthogonal components $\varepsilon_{jt}, j = 2, \ldots, n$ in the model. If they are not explicitly identified by economic restrictions, they may be thought of as arbitrary rotations of the underlying economic driving forces. As discussed in Section 2.4, we specify a uniform prior over the set of admissible model, rendering all rotations of $\varepsilon_{jt}, j = 2, \ldots, n$ equally likely. Note that restrictions 1.), 2.) and 6.) are invariant to orthogonal rotations of the remaining shocks and therefore, their identification (and hence the prior) does not matter. To see this for restriction 6.), consider the equation

$$m_t = \nu_m + \phi_1 \varepsilon_{1t} + \phi_2 \varepsilon_{2t} + \sigma_\eta \eta_t,$$

where ϕ_1 is 1×1, ϕ_2 is $1 \times n-1$ and ε_{2t} contains the $n-1$ other shocks of the system. Also consider the alternatively identified shocks $\bar{\varepsilon}_{2t} = Q_2' \varepsilon_{2t}$ with corresponding regression coefficients $\bar{\phi}_2 = \phi_2 Q_2$ where $Q_2 Q_2' = I_{n-1}$. Then, it holds

$$\sum_{j=2}^{n} \omega_j = \frac{\phi_2 \phi_2'}{\phi_1^2 + \phi_2 \phi_2' + \sigma_\eta^2} = \frac{\bar{\phi}_2 \bar{\phi}_2'}{\phi_1^2 + \bar{\phi}_2 \bar{\phi}_2' + \sigma_\eta^2}.$$

Applied researchers need to choose one particular way of exploiting the proxy variable from the menu above. As with the choice of alternative identification schemes in SVARs, this choice needs to be made by the researcher against the background of the particular application. For instance, in some applications, researchers may have a good understanding of reasonable values for threshold values. If no such information is available, then researchers may revert to methods 4.) to 6.).

We highlight that any of the restrictions outlined above may be easily combined with conventional sign restrictions on structural parameters of the model, e.g. on the effects of structural shocks or model implied elasticities. As we demonstrate in our empirical applications (Section 3.2), a combination with conventional sign restrictions can be a promising identification strategy if the latter alone are not strong enough to yield informative results.

2.4 Bayesian inference

In the following, we discuss how to conduct Bayesian inference for the augmented B-model type SVAR subject to the sign and zero restrictions discussed previously. Let $\tilde{A} = [\tilde{\nu}, \tilde{A}_1, \ldots, \tilde{A}_p]$, $\tilde{Y} = [\tilde{y}_1, \ldots, \tilde{y}_T]'$ and $X = [x_1, \ldots, x_T]'$ where $x_t = [1, \tilde{y}'_{t-1}, \ldots, \tilde{y}'_{t-p}]'$. Furthermore, let S_a and S_b be full rank selection matrices of zeros and ones such that $\alpha = S_a \text{vec}(\tilde{A})$ and $\beta = S_b \text{vec}(\tilde{B})$ are the nonzero free elements in $\tilde{A}$ and $\tilde{B}$. We work with a standard Gaussian likelihood for the data. Given known presample values $\tilde{y}_0, \tilde{y}_{-1}, \ldots, \tilde{y}_{-p+1}$, the likelihood of the augmented model is:

$$p(\tilde{Y}|\alpha, \beta) = (2\pi)^{-\frac{(n+k)T}{2}} |\tilde{B}\tilde{B}'|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\text{tr}(\tilde{B}^{-1'}\tilde{B}^{-1}(\tilde{Y} - X\tilde{A})(\tilde{Y} - X\tilde{A})')\right). \quad (2.12)$$

As the Gaussian likelihood is fully characterized by the first two moments, the likelihood is invariant to certain rotations of $\tilde{B}$. If no exogeneity restrictions are imposed, using $\tilde{B}$ or $\tilde{B}^* = \tilde{B}Q$ with $Q = \text{diag}(Q_1, Q_2)$ (see Section 2.1) gives the same likelihood value. In case we assume valid exogeneity assumptions, any rotation needs to satisfy $\tilde{\Phi} = \Phi Q_1 = [0_{k \times n-k}, \tilde{\phi}_3]$ as to ensure that the zero block in Φ is maintained.

With respect to the prior, we specify independent distributions for β and α. Specifically, for the autoregressive coefficients we assume a normal prior given by $p(\alpha; \alpha_0, V_\alpha) \sim \mathcal{N}(\alpha_0, V_\alpha)$. While this choice allows the user to include a wide range of prior information, normality implies conditional conjugacy and hence ensures straightforward treatment within Markov Chain Monte Carlo (MCMC) methods. Compared to the Bayesian proxy

SVARs of Caldara & Herbst (2019) and Arias, Rubio-Ramírez & Waggoner (2019), who exploit the fully conjugate prior on the A-model parameters pioneered in Sims & Zha (1998), the independent prior allows to potentially incorporate a wider spectrum of prior information that can be asymmetric across equations. These include the original Minnesota prior of Litterman (1986) and various popular hierarchical shrinkage priors (Koop et al. 2010). As discussed in Sims & Zha (1998) (Section 5.2), any asymmetric prior on the reduced form coefficients would be very difficult to handle in their framework.

For the unique elements β of the augmented impact matrix $\tilde{B}$ we set the prior density to

$$p(\beta; v_0, S_0) \propto |\det(\tilde{B})|^{-(v_0+(n+k))} \exp\left(-\frac{1}{2}\text{tr}\left(S_0\left(\tilde{B}\tilde{B}'\right)^{-1}\right)\right), \tag{2.13}$$

where v_0 is a scalar and S_0 a positive definite matrix of size $n + k \times n + k$.

For $k = 0$, $\tilde{B}$ is an unrestricted full rank matrix, and the prior is equivalent to other priors used in sign-restricted SVAR models. To see this, decompose the impact matrix as $\tilde{B} = PQ$ where P is a lower triangular matrix with positive diagonal elements and Q an orthonormal matrix ($QQ' = I_{n+k}$). Then, this prior choice induces that $\Sigma = PP' \sim i\mathcal{W}(v_0, S_0)$ follows an inverse Wishart and a Uniform (Haar) distribution for Q (Muirhead 1982), which is exactly the prior used in Uhlig (2005). Furthermore, our prior is equal to the prior specified in Arias et al. (2018) when both $p = 0$ and $k = 0$. Then, a simple change of variables for $\tilde{\text{A}} = \tilde{B}^{-1}$ yields the Jacobian of transformation $|\det(\tilde{\text{A}})|^{-2(n+k)}$ and the prior density of their paper $p(\tilde{\text{A}}; v_0, S_0) \propto |\det(\tilde{\text{A}})|^{v_0-(n+k)} \exp\left(-\frac{1}{2}\text{tr}\left(\tilde{\text{A}}' S_0 \tilde{\text{A}}\right)\right)$.

The conjugacy of $p(\beta)$ assures that akin to the likelihood, the prior is uniform over the set of admissible models. Hence, any identification of the model does only come from the restrictions discussed in Section 2.3 and 2.2, rather than from an inadvertent choice of the prior. At this point, we highlight that the algorithms considered in this paper are general enough to handle other choices of $p(\beta)$. For example, if the researcher wants to impose additional identifying information in terms of Bayesian priors of the forms recently discussed in Baumeister & Hamilton (2015), the density in equation (2.13) can be replaced accordingly.

The posterior distribution of the model parameters is proportional to the product of the priors and the likelihood:

$$p(\alpha, \beta|\tilde{Y}) = \frac{p(\tilde{Y}|\alpha, \beta)p(\alpha)p(\beta)}{p(\tilde{Y})}, \tag{2.14}$$

where the normalizing constant is $p(\tilde{Y}) = \int p(\tilde{Y}|\alpha, \beta)p(\alpha)p(\beta)d\alpha d\beta$. Given that the posterior is of no known form, we summarize posterior moments using Markov Chain Monte Carlo (MCMC) methods. In the following we will describe the MCMC algorithm in more detail.

We start with some notation. Denote by $\theta = \{\alpha, \beta\}$ the set of SVAR parameters, and by θ_{-x} the set of parameters excluding x. Setting arbitrary initial values $\theta^{(0)} = \{\alpha^{(0)}, \beta^{(0)}\}$, the proposed MCMC generates draws $\theta^{(i)}, i = 1, \ldots, M$ from the posterior by iteratively drawing from the following conditional distributions:

1. Draw $\alpha^{(i)}$ from $p\left(\alpha|\theta_{-\alpha}, \tilde{Y}\right) \sim \mathcal{N}(\bar{\alpha}, \bar{V}_\alpha)$ where mean and variance are:

$$\bar{V}_\alpha^{-1} = V_\alpha^{-1} + S_a((BB')^{-1} \otimes X'X)S_a',$$
$$\bar{\alpha} = \bar{V}_\alpha \left(V_\alpha^{-1} + S_a \text{vec}(X'\tilde{Y}(BB')^{-1})\right).$$

2. Draw $\beta^{(i)}$ from $p\left(\beta|\theta_{-\beta}, \tilde{Y}\right)$ which is proportional to:

$$p\left(\beta|\theta_{-\beta}, \tilde{Y}\right) \propto |\tilde{B}\tilde{B}'|^{-\frac{T+v_0+(n+k)}{2}} \exp\left(-\frac{1}{2}\text{tr}(\tilde{B}^{-1'}\tilde{B}^{-1}(S_0 + \tilde{U}\tilde{U}'))\right),$$

where $\tilde{U} = \tilde{Y} - X\tilde{A}$. Since the conditional distribution is of no known form, we rely on an Accept Reject Metropolis Hastings (AR-MH) step (Tierney 1994, Chib & Greenberg 1995). For a given proposal distribution $p^\star\left(\beta|\theta_{-\beta}, \tilde{Y}\right)$, which we discuss at a later point, the AR-MH algorithm involves two steps:

(a) *Accept-reject step*: Generate a candidate $\beta^\star \sim p^\star\left(\beta|\theta_{-\beta}, \tilde{Y}\right)$ and accept it with probability

$$\alpha_{\text{AR}}(\beta^\star) = \min\left\{1, \frac{p\left(\beta^\star|\theta_{-\beta}, \tilde{Y}\right)}{c_{AR} \times p^\star\left(\beta^\star|\theta_{-\beta}, \tilde{Y}\right)}\right\},$$

which is repeated until a draw is accepted.

(b) *Metropolis-Hastings step*: Accept the proposal $\beta^\star$ with probability $\alpha_{\text{MH}}(\beta^{(i-1)}|\beta^\star)$. Let $\mathcal{D}(\beta) = \left\{\beta : p\left(\beta|\theta_{-\beta}, \tilde{Y}\right) \leq c_{AR} \times p^\star\left(\beta|\theta_{-\beta}, \tilde{Y}\right)\right\}$ and $\mathcal{D}^C(\beta)$ its complement. Then:

$$\alpha_{\text{MH}}(\beta^{(i-1)}|\beta^\star) = \begin{cases} 1 & \text{if } \beta^{(i-1)} \in \mathcal{D}(\beta) \\ \dfrac{c_{AR} \times p^\star\left(\beta^\star|\theta_{-\beta}, \tilde{Y}\right)}{p\left(\beta^\star|\theta_{-\beta}, \tilde{Y}\right)} & \text{if } \beta^{(i-1)} \in \mathcal{D}^C(\beta), \beta^\star \in \mathcal{D}(\beta) \\ \dfrac{p\left(\beta^\star|\theta_{-\beta}, \tilde{Y}\right) p^\star\left(\beta^{(i-1)}|\theta_{-\beta}, \tilde{Y}\right)}{p\left(\beta^{(i-1)}|\theta_{-\beta}, \tilde{Y}\right) p^\star\left(\beta^\star|\theta_{-\beta}, \tilde{Y}\right)} & \text{if } \beta^{(i-1)}, \beta^\star \in \mathcal{D}^C(\beta) \end{cases}$$

The constant c_{AR} in the AR-MH step can be tuned to trade off the efficiency of the AR step against the acceptance probability in the MH step.[6] We iteratively tune this constant over a preliminary run of the MCMC as to capture twice the average ratio between target and proposal distribution. For the applications considered in this paper, this resulted in a reasonable trade-off between AR and MH steps, yielding acceptance probabilities of the latter in the range of 85%-99%.

The success of the AR-MH step depends critically on the design of the proposal distribution $p^\star\left(\beta|\theta_{-\beta}, \tilde{Y}\right)$. In Appendix A, we outline in detail a proposal distribution which relies on the methodology developed in Arias et al. (2018, 2019) to efficiently explore the conditional distribution of the set-identified parameters in $\tilde{B}$. Briefly summarized, the proposal involves drawing a candidate $\beta^\star = S_b \text{vec}(\tilde{B}^\star)$ for $\tilde{B}^\star = \text{chol}(\Sigma)Q$ by drawing $\Sigma \sim i\mathcal{W}(v, S)$ from an inverse Wishart with shape parameter S and degrees of freedom v, and $Q = \text{diag}(Q_1, Q_2)$ from a uniform distribution of Q_1 and Q_2 subject to the zero and sign restrictions discussed in Section 2.2 and 2.3. In order to capture the shape of the conditional distribution, we set $v = T + v_0$ and $S = \tilde{U}\tilde{U}' + S_0$. To evaluate the importance density of a candidate draw $\beta^\star$, we use numerical derivatives which account for the change of variables underlying the transformation of random variables Σ, Q to β.

[6] To see this, note that for increasing values of c_{AR}, the MH acceptance probability eventually approaches one given that any $\beta^\star \in \mathcal{D}$. However, at the same time the performance of the AR step deteriorates, as more and more draws are necessary until a draw is accepted.

After some burn in period, the algorithm is used to generate a large number of draws of the posterior distribution of θ. Those draws are then used in a standard fashion to summarize posterior quantities numerically.

2.5 Bayes factors

As outlined in Section 2.2, it can be useful to impose overidentifying sign restrictions in SVARs identified by external instruments. In this section, we discuss the use of Bayes factors as an econometric tool that quantifies the statistical support of such overidentifying restrictions. The use of Bayes factors as a tool to test overidentifying restrictions in SVARs is not completely new to the literature, see e.g. Woźniak & Droumaguet (2015), Lütkepohl & Woźniak (2018), Lanne & Luoto (2020) and Nguyen (2019). Let $p(\tilde{Y}|M_1)$ and $p(\tilde{Y}|M_2)$ the marginal likelihoods for two SVAR models (M_1 and M_2) identified by different sets of restrictions. Then, the Bayes factor is given by $\mathrm{BF}_{21} = p(\tilde{Y}|M_2)/p(\tilde{Y}|M_1)$ and directly quantifies the posterior odds of model M_2 over M_1. We refer readers to the comprehensive treatment in Kass & Raftery (1995) for more details on interpreting the size of Bayes factors.

In the following, we will demonstrate how Bayes factors can be computed in a straightforward way from the MCMC output of the less restrictive model. Throughout the following, denote by M_1 the SVAR model with less restrictive identifying restrictions. In this paper, we assume that the prior of the overidentified model M_2 can be factored as:

$$p(\theta|M_2) = \frac{p_2(\theta)p(\theta|M_1)}{\int p_2(\theta)p(\theta|M_1)d\theta} = \frac{p_2(\theta)p(\theta|M_1)}{c_\theta}. \tag{2.15}$$

Therefore, $p_2(\theta)$ is any *additional* identifying information imposed on the top of those assumed by the less restrictive model M_1. For the overidentifying sign restrictions that we aim to test, $p_2(\theta)$ simply takes the form of a uniform distribution over the restricted parameters space $\mathcal{S} \in \Theta$, that is $p_2(\theta) \propto \mathbf{1}(\theta \in \mathcal{S})$. But we note that more generally, $p_2(\theta)$ can be any probability density function designed to provide additional identifying information on the structural parameters. For example, those can take the form proposed in Baumeister & Hamilton (2019), who impose truncated t-distributions on model implied elasticities.

Note that for all priors of form (2.15), the posterior can be factored in an equivalent way:

$$
\begin{aligned}
p(\theta|M_2, \tilde{Y}) &\propto p_2(\theta)p(\theta|M_1)p(\tilde{Y}|\theta), \\
&\propto p_2(\theta)p(\theta|M_1, \tilde{Y}),
\end{aligned}
$$

such that

$$
p(\theta|M_2, \tilde{Y}) = \frac{p_2(\theta)p(\theta|M_1, \tilde{Y})}{\int p_2(\theta)p(\theta|M_1, \tilde{Y})d\theta} = \frac{p_2(\theta)p(\theta|M_1, \tilde{Y})}{c_{\theta|\tilde{Y}}} \tag{2.16}
$$

Under the prior (2.15) for M_2, the Bayes factor can be simplified considerably. First, note that using Bayes theorem and the fact that the models have the same parameters θ, we find

$$
\frac{p(\tilde{Y}|M_2)}{p(\tilde{Y}|M_1)} = \frac{p(\tilde{Y}|\theta)p(\theta|M_2)/p(\theta|M_2, \tilde{Y})}{p(\tilde{Y}|\theta)p(\theta|M_1)/p(\theta|M_1, \tilde{Y})} = \frac{p(\theta|M_2)p(\theta|M_1, \tilde{Y})}{p(\theta|M_1)p(\theta|M_2, \tilde{Y})}.
$$

Using expressions of equation (2.15) and (2.16) for prior and posterior of M_2 respectively, the Bayes factor simplifies to:

$$
\mathrm{BF}_{21} = \frac{p(\theta|M_2)p(\theta|\tilde{Y}, M_1)}{p(\theta|M_1)p(\theta|\tilde{Y}, M_2)} = \frac{\left(p(\theta|M_1)p_2(\theta)c_\theta^{-1}\right)p(\theta|\tilde{Y}, M_1)}{p(\theta|M_1)\left(p(\theta|M_1, \tilde{Y})p_2(\theta)c_{\theta|\tilde{Y}}^{-1}\right)} = \frac{c_{\theta|\tilde{Y}}}{c_\theta}.
$$

Furthermore, note that $c_{\theta|\tilde{Y}}$ and c_θ can be expressed as expectations of $p_2(\theta)$ over prior and posterior distribution respectively:

$$
\frac{c_{\theta|\tilde{Y}}}{c_\theta} = \frac{\int p_2(\theta)p(\theta|M_1, \tilde{Y})d\theta}{\int p_2(\theta)p(\theta|M_1)d\theta} = \frac{\mathrm{E}_{\theta|\tilde{Y}}[p_2(\theta)]}{\mathrm{E}_\theta[p_2(\theta)]}.
$$

This makes it straightforward to estimate $c_{\theta|\tilde{Y}}$ and c_θ using draws from the prior and posterior respectively of the less restrictive model M_1. In particular, we may use the simulation consistent averages $\hat{c}_{\theta|\tilde{Y}} = 1/J_1 \sum_{j=1}^{J_1} p_2(\theta^{(j)})$ for $\theta^{(j)} \sim p(\theta|M_1, \tilde{Y})$ and $\hat{c}_\theta = 1/J_2 \sum_{i=1}^{J_2} p_2(\theta^i)$ for $\theta^{(i)} \sim p(\theta|M_1)$. By a standard limiting theorem, the Monte Carlo estimators $\hat{c}_{\theta|\tilde{Y}}$ and $\hat{c}_\theta$ are asymptotically normal:

$$\sqrt{T}\begin{pmatrix}\hat{c}_{\theta|\tilde{Y}} - c_{\theta|\tilde{Y}} \\ \hat{c}_\theta - c_\theta\end{pmatrix} \rightarrow \mathcal{N}\left(0, \begin{pmatrix}\sigma^2_{\theta|\tilde{Y}} & 0 \\ 0 & \sigma^2_\theta\end{pmatrix}\right)$$

and by using the 'Delta' method, we find

$$\sqrt{T}\left(\frac{\hat{c}_{\theta|\tilde{Y}}}{\hat{c}_\theta} - \frac{c_{\theta|\tilde{Y}}}{c_\theta}\right) \rightarrow \mathcal{N}\left(0, \frac{\sigma^2_\theta}{c^2_{\theta|\tilde{Y}}} + \frac{\sigma^2_{\theta|\tilde{Y}} \cdot c^2_\theta}{c^4_{\theta|\tilde{Y}}}\right).$$

This last result can be used to obtain an approximate standard error for $\widehat{\mathrm{BF}}_{21} = \hat{c}_{\theta|\tilde{Y}}/\hat{c}_\theta$, where in practice, we replace the unknown quantities in the variance by corresponding estimates. Note that for $\sigma^2_{\theta|\tilde{Y}}$ and σ^2_θ, a Newey-West estimator (Newey & West 1987) is used in order to account for the autocorrelation inherent in simulating draws from the prior and posterior by MCMC methods.

3 Empirical applications

We demonstrate the usefulness of our methodological framework in two empirical applications. In Section 3.1, we use a combination of sign and instrumental variables restrictions to identify supply and demand shocks that drive oil prices. In that application, we rely on an instrument that is credibly exogenous and therefore, we use the type of restrictions described in Section 2.2. In Section 3.2, we analyze the effects of monetary policy shocks on macroeconomic and financial variables. For identification, we rely on the restrictions introduced in Section 2.3 that exploit identifying information from a 'plausibly exogenous' instrument in combination with conventional sign restrictions on structural parameters.

3.1 The importance of oil supply shocks for driving oil prices

Following the work of Kilian (2009), SVAR models have been extensively used to disentangle oil price movements into supply and demand shocks (see e.g. Kilian & Murphy (2012), Kilian & Murphy (2014), Herrera & Rangaraju (2020), Baumeister & Hamilton (2019), and Caldara et al. (2019)). Despite the large number of studies, there is still no consensus on the relative importance of oil supply and demand shocks. One question debated in the literature is how important oil supply shocks are for the determination of oil prices. In this line of the literature, the importance of oil supply shocks is measured by a forecast error variance decomposition (FEVD), i.e. the fraction of forecast error variance accounted for by the supply shock. Interestingly, the literature suggests a fairly wide range of FEVD estimates. Empirical results show a range between just above 0% to more than 40% depending on the underlying identification strategy. What drives these remarkably different estimates? To identify demand and supply shocks, most of the papers mentioned above use a combination of sign restrictions together with additional restrictions on the implied short-term oil supply elasticity. They essentially differ in how much the elasticity is restricted. In particular, studies that have imposed very small short-term supply elasticities for identification, arrive at estimates close to the lower bound (Kilian & Murphy 2012, 2014, Herrera & Rangaraju 2020). On the other hand, when larger supply elasticities are imposed for identification, one may find supply shocks to be equally important as demand shocks (Baumeister & Hamilton 2019, Caldara et al. 2019). Thus, in these models the relative importance of oil supply shocks depends to a large extent on the elasticity restrictions imposed.

In the following, we illustrate how our proposed method from Section 2.2 can be helpful to revisit the importance of oil market shocks. In a first step, we avoid the use of potentially controversial elasticity constraints by combining sign restrictions on impulse response functions with restrictions implied by using the oil supply shock of Kilian (2008) as an external instrument. Furthermore, as the resulting model is informative about oil supply elasticities, we can use Bayes factors to check different elasticity constraints used in the literature against the data.

We identify the shocks of interest in a VAR following recent specifications for the global oil market (see e.g. Känzig (2019) and Baumeister & Hamilton (2019), henceforth: BH19). In this model, the variables in the VAR are

$$y_t = (\text{prod}_t, \text{rea}_t, \text{rpo}_t, \text{i}_t)',$$

where prod_t is the log of world oil production, rea_t is a proxy for real economic activity, where we choose the industrial production index of Baumeister & Hamilton (2019). Furthermore, rpo_t is the real price of oil and i_t is the log of crude inventories, as computed in Kilian & Murphy (2014) and seasonally adjusted via the X-13ARIMA-SEATS program of the Census Bureau. We include $p = 13$ lags as the variables are included in levels (BH19 use p=12 but with variables in first differences).

In our analysis, we include monthly data starting in 1958M01 up to 2017M12. Here, we use sample information until 1983M12 to train a prior distribution based on ordinary least square quantities (OLS). Thereby, we avoid specifying a subjective prior distribution, which may influence Bayes factors used to check the elasticity restrictions later in the analysis. Furthermore, given that oil prices were regulated before 1974, we broadly follow Baumeister & Hamilton (2019) in inflating the variance in the prior distribution by a factor of ten, and thereby heavily discount the information we draw from the training sample.[7]

In the structural analysis, we follow Kilian & Murphy (2014) in identifying three out of the four shocks in the SVAR. In particular, we identify an oil supply shock denoted as ε_t^s, and aggregated demand shock ε_t^{ad} and an oil-specific demand shock ε_t^{od}. For details on the economic interpretation of these shocks, we refer the reader to Kilian & Murphy (2014). Identification is achieved by (a combination of) the following restrictions:

1. Sign restrictions (SR):

$$\begin{pmatrix} u_t^{\Delta\text{prod}} \\ u_t^{\text{rea}} \\ u_t^{\text{rpo}} \\ u_t^{\Delta\text{i}} \end{pmatrix} = \begin{pmatrix} - & + & + & * \\ - & + & - & * \\ + & + & + & * \\ * & * & + & * \end{pmatrix} \begin{pmatrix} \varepsilon_t^s \\ \varepsilon_t^{ad} \\ \varepsilon_t^{od} \\ \varepsilon_{4t} \end{pmatrix}.$$

2. Instrumental variable (IV) constraints: $\text{E}[\varepsilon_t^s m_t] \neq 0$, while $\text{E}[\varepsilon_t^{ad} m_t] = \text{E}[\varepsilon_t^{od} m_t] = \text{E}[\varepsilon_t^4 m_t] = 0$.

3. Elasticity constraints. Denote the supply elasticities by $\eta_1 = B_{12}/B_{32}$ and $\eta_2 = B_{13}/B_{33}$. We then use two different restrictions suggested in the literature:

[7]Specifically, we set $\alpha_0 = \hat{\alpha}_0$ and $V_\alpha = 10 \cdot \hat{V}_0$, where $\hat{\alpha}_0$ are OLS estimates of α and $\hat{V}_0$ is the covariance estimate of $\hat{\alpha}$ based on OLS. Furthermore, we set $v_0 = T_0/10$ and $S_0 = 10 T_0 \hat{\Sigma}$ where T_0 is the sample size of the training sample and $\hat{\Sigma}$ the sample covariance matrix of the (training sample) OLS residuals augmented by the instrument m_t.

(a) HR20: Herrera & Rangaraju (2020) set an upper bound of 0.04 that covers all the recent micro-economic estimates computed by Anderson et al. (2018), Bjørnland et al. (2017) and Newell & Prest (2019). Hence: $p(\eta_{1/2} \leq 0.04) = 1$ and 0 else. This upper bound is therefore a little above the 0.0258 imposed in Kilian & Murphy (2014).

(b) BH19: Baumeister & Hamilton (2019) impose identifying information on the oil supply elasticity in form of a positively truncated t-distribution with mode at 0.1, scale parameter equal to 0.2 and 3 degrees of freedom, $\eta_{1/2} \sim t(0.1, 0.2, 3)$. Compared to (a), this attaches high probability mass to much larger values.

The impact sign restrictions in 1.) are those from Kilian & Murphy (2014). The restrictions in 2.) reflect the relevance and exogeneity restrictions of the instrumental variable m_t, indicating that m_t is correlated with the oil supply shocks ε_t^s but uncorrelated with all other shocks in the system. As an instrument m_t, we make use of the exogenous supply shock series as proposed in Kilian (2008), which reflects unexpected oil supply disruptions caused e.g. by geopolitical turmoils and wars. For our analysis, we have recomputed Kilian's monthly oil supply shock series from oil production data and extended it to match our estimation sample. The extended series includes shocks related to the Libyan civil war and militia attacks during 2011 and 2013. We give a detailed description on how we have constructed the time series and a plot in Appendix B. Finally, the elasticity restrictions that have been used in the literature are summarized in 3.).

We discuss the main results from our empirical analysis next. First, we illustrate in the top row of Figure 1 that using different elasticity restrictions in conjunction with of sign restrictions implies fairly different results with respect to the importance of oil market shocks. The figure contains contributions of identified supply and demand shocks to the forecast error variance of oil prices. The solid line with the shaded area gives median estimates with 68% posterior credibility sets for a model identified with sign restrictions together with elasticity constraints from Baumeister & Hamilton (2019)[combining 1. + 3(b)]. The dashed lines give corresponding estimates for a model identified by sign restriction together with elasticity constraints in Herrera & Rangaraju (2020) [combining 1. + 3(a)]. The 68% posterior credibility set of the former model (BH19) implies a much larger contribution of

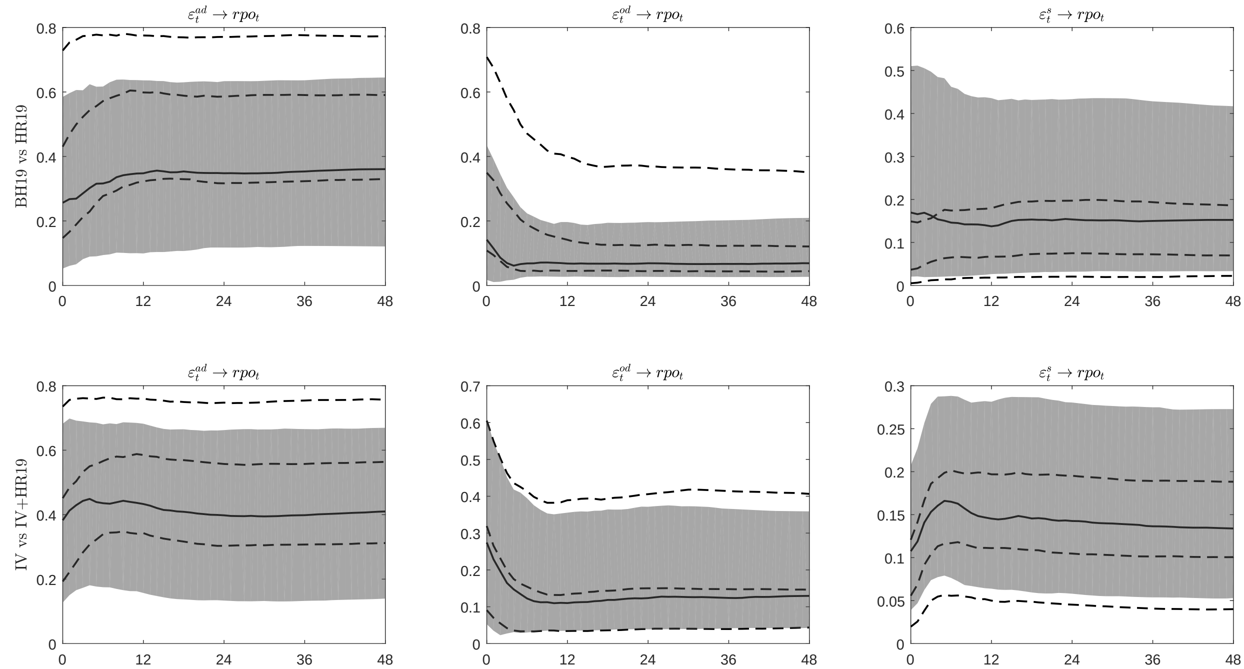

Figure 1: Posterior median with 68% credibility region for the forecast error variance decomposition (FEVD) of the real price of oil. Top: sign restricted SVARs plus BH19 (solid line with shaded area) compared to sign restricted model plus HR20 (dashed). Bottom: sign restricted SVARs with IV restrictions (solid line with shaded area) compared to sign restricted model with IV and HR20 restrictions.

oil supply shocks (up to 50%), while the HR20 model with much tighter elasticity bounds yields substantially smaller contributions for the supply shock with values ranging between just above 0% to about 20%.

Next, we consider the results of a model identified by combining the sign restrictions 1.) with the IV restriction in 2.). This corresponds to the setting introduced in Section 2.2. We argue that this combination is helpful because using either of the two restrictions in isolation may not be useful in the current application. First, using only sign restrictions would result in large and uninformative credibility sets for the variance decomposition of the oil price (see e.g. Kilian & Murphy (2012)). Second, using solely the Kilian shock as an instrument for identifying the oil supply shock may be problematic too as it is found to be a weak instruments (see e.g. Montiel Olea et al. (2020)). Note that our combination model avoids using the potentially controversial elasticity restrictions.

The solid line with the shaded area in the bottom row of Figure 1 gives median estimates with 68% posterior credibility sets from this model. The results show that median variance

Table 1: Posterior distribution of supply elasticities and Bayes factors for overidentifying restrictions

Panel A: Posterior quantiles			
Elasticities	16%	50%	84%
η_1	0.017	0.049	0.110
η_2	0.013	0.040	0.099

Panel B: Bayes factors				
Restrictions	$\mathrm{E}_{\theta\mid\tilde{Y}}[p_2(\theta)]$	$\mathrm{E}_{\theta}[p_2(\theta)]$	$\widehat{\mathrm{BF}}$	s.e.
BH19	4.966	2.624	1.893	0.002
HR20	0.204	0.005	41.291	0.036

Bayes factors computed as described in section 2.5. For BH19, $p_2(\theta) : \eta_{1/2} \sim t(0.1, 0.2, 3)$ while for HR20 $p_2(\theta) : p(\eta_{1/2} \leq 0.04) = 1$ and 0 else.

contributions of oil supply shocks range between 10% to 15%, with the credibility set indicating that at most 27% of the variation in oil prices is explained by supply shocks. Compared to the models that use elasticity restrictions above, the relatively small variance contributions are closer to those of Herrera & Rangaraju (2020). We also note that the instrument is quite informative with respect to the oil supply elasticities. Panel A of Table 1 shows the 16%, 50% and 84% posterior quantiles for the model (SR+IV) implied oil supply elasticities. Median estimates of η_1 and η_2 are between 0.04 and 0.05. Note again, that this model does not restrict the elasticities explicitly. We find, however, that the median estimates are much lower than the value considered in BH19, and closer to the upper bound of HR20. Clearly, using the instrument on top of the sign restrictions also avoids having unreasonably large elasticities found when using sign restrictions alone (see again Kilian & Murphy (2012)).

Using our Bayesian setup, we can now use the Bayes factors introduced in Section 2.5 in the model identified by SR+IV restrictions to check whether additional elasticity constraints are supported by the data. Panel B of Table 1 reports the estimated Bayes factors for different models. First, consider comparing the SR+IV model against a model that additionally imposes the BH19 elasticity restrictions (SR+IV+BH19). The estimated Bayes factor is about 2.6, which indicates mild support for the imposed BH19 restriction.[8]

[8]We follow the guideline for interpreting Bayes factors in Section 3.2 of Kass & Raftery (1995).

In contrast, for the HR20 elasticity restriction (SR+IV+HR20) we find a Bayes factor of 41.3, which represents much stronger support. We may also use the information from Table 1 to compare the HR20 restrictions against the BH19 restrictions directly by simply taking the ratio of the Bayes factors, $\mathrm{BF}_{\mathrm{HR20}}/\mathrm{BF}_{\mathrm{BH10}} \approx 21.8$. This indicates strong support in favor of the restrictions in HR20 over those of BH19.[9]

Finally, given that the HR20 elasticity restrictions receive strong support by the data, we report results of the corresponding model that imposes the HR20 restrictions on top of sign and IV restrictions (SR+IV+HR20). The results are shown as dashed lines in the bottom row of Figure 1 and indicate that supply shocks now account between 5% and 20% of oil price variation. Compared with the SR+IV specification without elasticity constraints, the variance contributions of this model are slightly lower. At the same time, additional elasticity information decreases posterior uncertainty somewhat.

Our empirical results suggest that the HR20 elasticity restrictions are more strongly supported by the data. Using them in addition to sign and IV restrictions, suggests that oil supply shocks are less important for oil price determination than suggested in some parts of the literature. More generally, we have illustrated how our framework of combining IV and sign restrictions can be used to discriminate between competing additional constraints in a data-driven way.

3.2 The effects of monetary policy

The effects of monetary policy shocks on macroeconomic aggregates have been extensively studied using SVAR models (see Ramey (2016) for a recent review of the literature). In the early literature, surprises to monetary policy have been identified by using a Cholesky decomposition of the reduced form VAR covariance matrix, with the policy instrument ordered below the real variables, see e.g. Christiano et al. (1999). This identifying assumption implies that the central bank can respond instantaneously to movements in the real sector of the economy, while the real variables may only respond to the policy shock with a lag of

[9] We have also conducted a number of robustness checks with respect to the training sample and the discount factor. Using a variance factor of 20 instead of 10, leads to very similar Bayes factors as those reported in Table 1. Moreover, shortening the training sample to 1974M1-1983M12 changes the Bayes factors somewhat but does not alter the conclusion of favoring HR20 restrictions over BH19.

one period. Such an identification is in line with macroeconomic models subject to nominal rigidities (Christiano et al. 2005). However, if financial variables are included in the VAR analysis, the recursiveness assumption is unrealistic no matter of the ordering, since it can be assumed that both, monetary policy and financial markets respond immediately to all structural shocks in the system.

Therefore, alternative identification schemes have emerged in recent years that avoid the recursiveness assumption. One strand of the literature uses sign restrictions, possibly combined with zero restrictions to set-identify the monetary policy SVAR. These restrictions are derived from conventional wisdom, such that a monetary policy tightening should be associated with an increase in the interest rates but not in consumer prices (Uhlig 2005, Faust 1998) or that the Fed tightens monetary policy stance in reaction to surprising increases in output and inflation (Arias, Caldara & Rubio-Ramírez 2019). Unfortunately, because of the implied set identification, this often leads to wide confidence intervals around impulse responses such that results are often uninformative with respect to financial variables.

An alternative branch of the literature uses narrative measures of monetary policy shocks for identification. Among the most prominent measures are shock series based on readings of Federal Open Market Committee (FOMC) minutes cleaned by Greenbook forecasts[10] for output and inflation (Romer & Romer 2004, Coibion 2012, Miranda-Agrippino et al. 2018) and factors based on changes in high frequency future prices around FOMC meetings (Faust et al. 2004, Gertler & Karadi 2015, Nakamura & Steinsson 2018). However, it is a very difficult task to construct convincing exogenous instruments for monetary policy. With respect to the Romer & Romer shock (henceforth R&R), the authors themselves state that their series is only 'relatively free of endogenous and anticipatory movement' (Romer & Romer 2004). To ensure against remaining endogeneity they exclude the possibility of a contemporaneous response of the macroeconomic variables to the narrative series. Furthermore, as demonstrated in Caldara & Herbst (2019), the FOMC responds not only to forecasts of output and inflation, but also responds to the information in credit spreads. This finding directly invalidates the use of the R&R residual as an external instrument to study the effects of monetary policy on financial markets. The exogeneity of instruments

[10] Greenbook forecasts are those published by the central bank in their FOMC minutes and therefore assumed to be in the information set of the central bank.

based on high frequency future data is also questionable. Ramey (2016), for example, finds that the main instrument of Gertler & Karadi (2015) suffers from a nonzero mean, significant autocorrelation and predictability by Greenbook forecasts. Furthermore, as highlighted in Nakamura & Steinsson (2018), it is difficult to ensure that the observed reactions in asset prices are due to news about monetary policy, and do not reflect superior information of the FOMC about future economic conditions.

As laid out in Section 2.3, our methodology provides a simple framework to exploit identifying information in proxy variables that are just 'plausibly exogenous', which we will combine with conventional sign restrictions. We start our analysis identifying a monetary policy shock based on Arias, Caldara & Rubio-Ramírez (2019) (ACR henceforth). To further narrow down the set of admissible models, we restrict the relation between the SVAR monetary policy shock and the R&R narrative shock. In particular, we impose the additional restriction that the monetary policy shock explains more variance of the narrative series than all other shocks together. As we will demonstrate, this sharpens identification of the set identified model and leads to more informative results, while at the same time avoids the potentially wrong assumption of exogeneity.

For our empirical study, we follow the specification of ACR and include the following variables into a monthly SVAR model: $y_t = (\text{gdp}_t, \text{def}_t, \text{cp}_t, \text{tr}_t, \text{nbr}_t, \text{ffr}_t)'$, where gdp_t is the real gross domestic product, def_t is the GDP deflator, cp_t is a commodity price index, tr_t are total reserves, nbr_t are non-borrowed reserves, and ffr_t is the federal funds rate. All variables are transformed to log times 100, except for ffr_t which is included in annualized percentages.[11] To ease comparison, we stick the original sample period that starts in 1965M1 and ends in 2007M12. We use $p = 12$ lags to account for sufficient dynamics of the time series vector. With respect to the narrative series, we use $m_t = \text{rr}_t$, the R&R narrative shock series updated by Wieland & Yang (2016). Furthermore, we set $\Gamma_{1i} = \Gamma_{2i} = 0$, excluding predictability of R&R by lagged values of $\tilde{y}_t$ and choose weakly informative prior distributions with $\alpha_0 = 0$, $V_\alpha = 1e07 \times I$, $v_0 = k + n + 1$ and $S_0 = I_{k+n}$.

To demonstrate the merits of our approach, we will compare the following identification schemes: a pure IV approach which assumes that the R&R shock is a valid instrument

[11] All time series were obtained from the replication files of Arias, Caldara & Rubio-Ramírez (2019). Note that gdp_t and def_t were interpolated based on US industrial production and CPI prices, respectively.

for monetary policy (R1), a combination of zero and sign restriction as considered in ACR (R2), and a combined identification scheme that relaxes the exogeneity assumption (R3).

In the first identification scheme (R1), we treat the R&R residuals as an exogenous instrument for the monetary policy shock ($\varepsilon_t^{mp} = \varepsilon_{1t}$). Therefore the identifying restrictions for R1 are given by $\text{E}[\varepsilon_t^{mp} m_t] \neq 0$ and $\text{E}[\varepsilon_{it} m_t] = 0, i \neq 1$, yielding the zero restrictions on Φ discussed in Section 2.2.

With respect to the sign restrictions (R2), ACR propose to identify ε_t^{mp} through a combination of zero and sign restrictions on the monetary policy rule implicit in the SVAR model. Rewriting the model as a simultaneous equation system, the systematic component of monetary policy is given by:

$$r_t = \xi_y u_t^{gdp} + \xi_\pi u_t^{def} + \xi_{cp} u_t^{cp} + \xi_{tr} u_t^{tr} + \xi_{nbr} u_t^{nbr} + \sigma_\xi \varepsilon_t^{mp}. \tag{3.1}$$

The coefficients can be backed out by $\xi_y = -a_{0,n1}^{-1} a_{0,11}$, $\xi_\pi = -a_{0,n1}^{-1} a_{0,12}$, $\xi_{cp} = -a_{0,n1}^{-1} a_{0,13}$, $\xi_{tr} = -a_{0,n1}^{-1} a_{0,14}$, $\xi_{nbr} = -a_{0,n1}^{-1} a_{0,15}$, and $\sigma_\xi = a_{0,n1}^{-1}$ where $a_{ij,0}$ are the elements of $A_0 = B^{-1}$. ACR impose the following combination of restrictions on equation (3.1): $\text{R2} : \{\xi_y > 0, \xi_\pi > 0, \xi_{tr} = 0, \xi_{nbr} = 0\}$, implying that the central bank increases the federal funds rate in response to positive surprises of output or prices, while it does not show systematic reactions towards surprises in monetary aggregates.[12]

Finally, the combined identification scheme (R3) is based on the same restrictions as considered in ACR (R2). In addition, we impose the following restriction on the relation between ε_t^{mp} and m_t: $\omega_1 > \sum_{j=2}^n \omega_j$, where ω_j is the contribution of the jth structural shock to the variance of m_t (see Section 2.3). This restriction implies that the monetary policy shock must explain more variation of m_t than all other structural shocks together. While this restriction is certainly weaker than exogeneity restrictions imposed by an IV approach, we still extract information by discarding monetary policy shocks that are only loosely related to the R&R narrative shock.

We start our analysis with a comparison of the posterior quantiles for the parameters governing the monetary policy rule (see Table 2). Our findings suggest that a Proxy SVAR

[12] Strictly speaking, these restrictions imply a set identified model based on a combination of zero and sign restrictions. Within our framework, this requires adjusting the proposal distribution of Appendix A to account for the additional zero restrictions, see Arias et al. (2018) for details.

Table 2: Posterior distribution for parameters of the policy rule

	R1			R2			R3		
quantile	ξ_y	ξ_π	ξ_{cp}	ξ_y	ξ_π	ξ_{cp}	ξ_y	ξ_π	ξ_{cp}
5%	-11.85	-38.91	-2.20	0.25	0.33	-0.61	0.15	0.07	-0.11
50%	0.99	1.56	0.03	0.84	2.44	-0.01	0.42	0.59	-0.00
95%	16.48	45.35	1.70	3.63	10.64	0.45	0.74	1.48	0.12

Posterior quantiles of the parameters governing the monetary policy rule

based on the R&R residual is very uninformative about the underlying parameters. In particular, 90% posterior confidence sets include implausible values in terms of magnitudes and sign, for both the reaction of the central bank towards real activity and prices. For the second model (R2), the restrictions imply plausible values for the sign of the underlying parameters. However, the 90% probability intervals still suggest very large values. For example, the 95% quantile of ξ_π would imply that in reaction to a 1% increase in prices, the central bank systematic reaction is to increase the federal funds rate by as much as 10 percentage points. Adding the additional restriction on the relation between the policy shock and the R&R residual substantially narrows down the probability intervals. Values between 0.15 and 0.75 for ξ_y and between 0.07 and 1.5 for ξ_π seem reasonable and are more in line with estimates in the DSGE literature.

Figure 2 reports responses to a monetary policy shock from SVARs identified with restrictions R1, R2 and R3. The top row shows results from the model identified by using the R&R series as an external instrument (R1). We observe puzzling results with a short-term increase in output together with a sharp and significant positive response in aggregate prices (a pronounced price puzzle). These responses are considered to be unreasonable and cast doubt on the credibility of the IV identification. In particular, these results may simply reflect that the R&R instrument is not truly exogenous. Consequently, any results from an analysis treating the R&R series as an exogenous IV should therefore be taken with great caution.

In contrast, when using the ACR zero/sign restrictions only (R2, second row of Figure 2), the puzzling results disappear but apparently lead to very wide error bounds that most often include the zero line. For most variables, model and estimation uncertainty is too large and therefore the responses are completely uninformative. The model does, however,

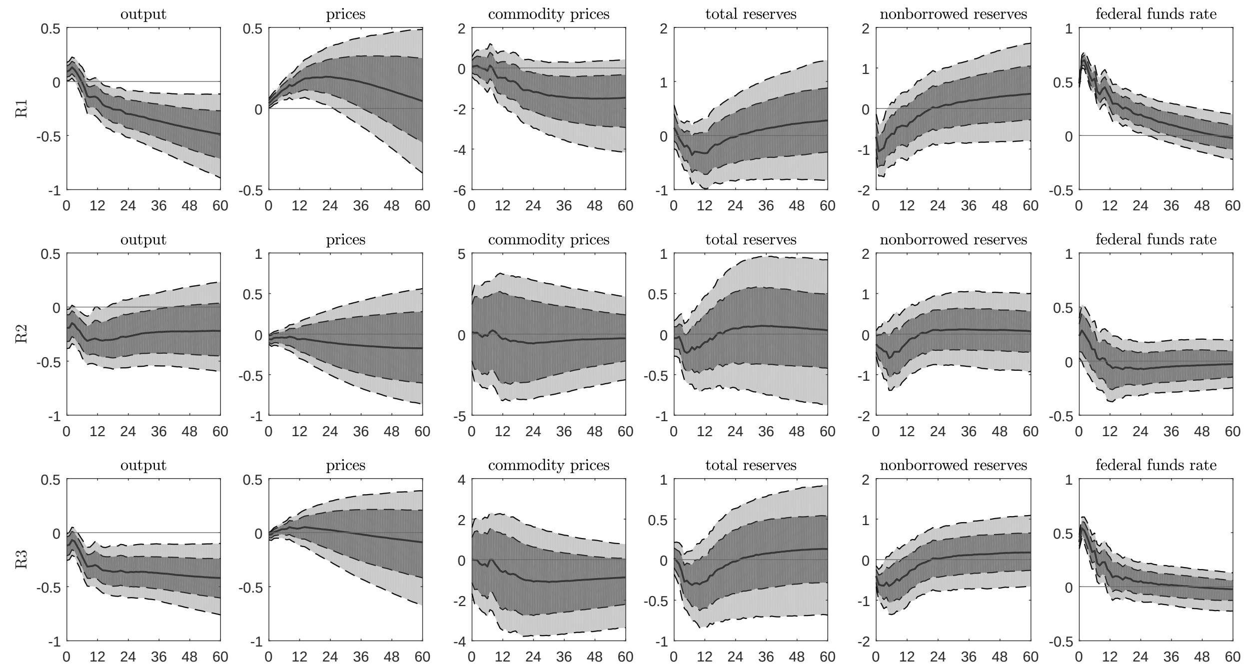

Figure 2: Impulse responses in the monetary policy SVAR obtained by using different identifying restrictions. Posterior median (solid line), 68% and 90% posterior credibility sets (dotted lines). Sample period: 1965M01-2007M12.

indicate a short-term drop in output. The bottom row of Figure 2 shows the results from using R3, the combination of sign restrictions together with assuming that the identified monetary policy shock explains more variation in the R&R series than all other shocks together (see above). We see that the combined approach leads to tighter credibility sets and therefore gives more informative results than using sign restrictions only (R2). In particular, we now find a clearly significant and permanent drop in output together with significant short-term drop in non-borrowed and total reserves.

Finally, we also demonstrate that our combination approach is particularly useful if interest is in the response of financial variables. To this end, we add one financial variable at a time to our monetary VAR and then compute responses to a monetary policy shock. The financial variables we consider are real stock prices, measured as the log of consumer price deflated S&P500 index, the mortgage spread, defined as difference between 30-year fixed rate mortgage average and the 10-year treasury yield, the commercial paper spread, defined as 3-month AA financial commercial paper rate minus the 3-months T-bill rate, and the 'excess bond premium' measure of credit market tightness developed by Gilchrist &

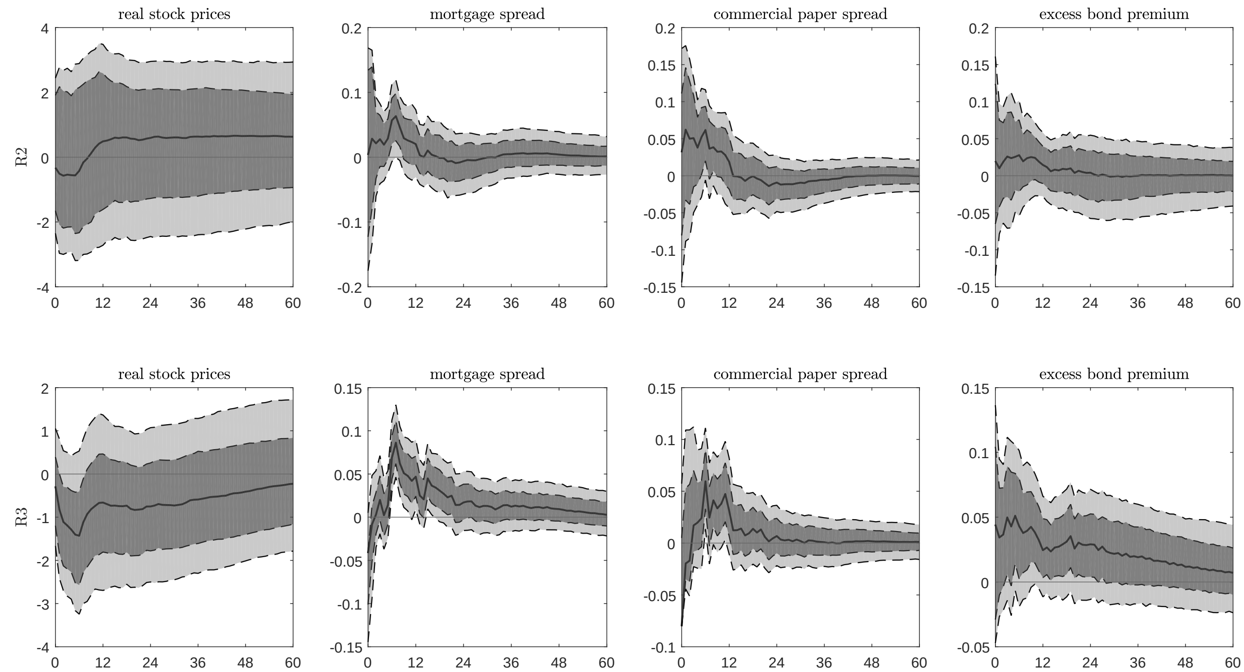

Figure 3: Impulse responses in the monetary policy SVAR augmented by one financial variable at a time. Posterior median (solid line), 68% and 90% posterior credibility sets (dotted lines). Sample periods: 1965M01-2007M12 (real stock prices, commercial paper spread), 1971M04-2007M12 (mortgage spreads), 1973M01-2007M12 (excess bond premium).

Zakrajšek (2012). Note that the mortgage spread and the EBP are only available starting from 1971M04 and 1973M01, respectively. Consequently, we have adjusted the sample period of models using these variables accordingly. We compare results from using the ACR zero/sign restrictions only (R2) against using the combination R3 and show the results in Figure 3.

As in the monetary SVAR without financial variables, we find the general pattern that posterior credibility sets are much tighter if we exploit information from the R&R series in addition to the sign restrictions (R3). This, in turn, leads to more informative impulse response patterns. For instance, using R2 (only ACR zero/sign restrictions) not much can be said on the response of stock prices and the excess bond premium to a monetary policy shock as the credibility sets are wide and include zero. In contrast, the picture is clearer when we use R3. In this model, real stock prices tend to fall and the excess bond premium responds positively, indicating tighter credit markets. The responses are significantly different from zero, at least if judged by the 68% posterior credibility sets.

Similarly, using R3 instead of R2 suggests a significant increase in the mortgage spread after about 6 months. Thus, using a combination as suggested in Section 2.3 is very useful here, since the way we exploit information from the R&R series does not require m_t to be exogenous. Instead, using the information in the R&R series only to discard models that imply shocks which are only loosely connected to the proxy series, gives results that could not have been obtained with the ACR zero/sign restrictions alone.

4 Conclusion

In this paper we discuss different ways of combining sign restrictions with information in time series that act as proxy or external instrumental variables for the identification of structural shocks in SVAR models. For this purpose, we employ a B-model type Proxy SVAR as the econometric modeling framework. In the first combination setting, we assume that valid instrumental variables are available for some of the shocks. Then additional sign restrictions can either be used to identify other shocks in the system or to further disentangle multiple shocks identified by valid external instruments. Furthermore, sign restrictions used on top of the IV restrictions, may be overidentifying and checked against the data. The second combination variant suggests ways to impose sign restrictions when the external proxy variables are only 'plausibly exogenous'. Here, we suggest to use inequality restrictions to bound the set of admissible models by discarding those that imply structural shocks without close relation to the external proxy time series. We have discussed various ways to characterize this relation e.g. based on correlations or variance contributions. In contrast to similar suggestions in the literature, we also include methods that avoid choosing thresholds, which can be advantageous in many empirical applications. These restrictions can be combined with conventional sign restrictions to narrow down the identified set.

Methodologically, we develop a Bayesian inference approach for the B-model type Proxy SVAR. To the best of our knowledge, this has not been discussed elsewhere in the literature. Given that B-model type SVARs are very popular in applied work, we see our approach as a useful addition to the Proxy VAR literature, specifically as our framework can accommodate more flexible priors. Our Bayesian inference involves efficient Markov Chain Monte Carlo methods, which incorporate the proposal distributions of Arias et al. (2018, 2019) via an Accept-Reject Metropolis Hasting (AR-MH) algorithm. Furthermore, we discuss the estimation of Bayes factors that can be used to check overidentifying restrictions against the data.

We also illustrate the usefulness of our method in two empirical applications. In the first application, we revisit a benchmark SVAR model for the global market of crude oil. We show how a model that combines IV and sign restrictions can be used to check various oil supply elasticity constraints from the literature against the data. In the second application, we identify the effects of monetary policy shocks in the United States by a combination of sign restrictions and information in the Romer & Romer (2004) narrative measure of a US monetary policy. Results based on our second combination approach, where we relax the exogeneity assumption, avoids puzzling results and leads to more informative responses in a number of financial variables.

Overall, the empirical illustrations provide new and useful empirical insights that could not have been obtained without our combination approach. Thus, our paper suggests that combining sign restrictions and external proxy variables for structural shock identification is a promising way to sharpen results from SVAR models.

References

Anderson, S. T., Kellogg, R. & Salant, S. W. (2018), 'Hotelling under pressure', *Journal of Political Economy* **126**(3), 984–1026.

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

Antolín-Díaz, J. & Rubio-Ramírez, J. F. (2018), 'Narrative sign restrictions for SVARs', *American Economic Review* **108**(10), 2802–29.

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

Arias, J. E., Rubio-Ramírez, J. F. & Waggoner, D. F. (2018), 'Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications', *Econometrica* **86**(2), 685–720.

Arias, J., Rubio-Ramírez, J. & Waggoner, D. F. (2019), 'Inference in Bayesian proxy-SVARs'.

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

Baumeister, C. & Hamilton, J. D. (2019), 'Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks', *American Economic Review* **109**(5), 1873–1910.

Bertsche, D. (2019), The effects of oil supply shocks on the macroeconomy: A Proxy FAVAR approach, Working Paper Series of the Department of Economics, University of Konstanz 2019-06, Department of Economics, University of Konstanz.

Bjørnland, H. C., Nordvik, F. M. & Rohrer, M. (2017), Supply flexibility in the shale patch: Evidence from North Dakota, Working paper, Norges Bank, Norway.

Bruns, M. & Piffer, M. (2019), Bayesian structural VAR models: a new approach for prior beliefs on impulse responses, Discussion paper, DIW Berlin.

Caldara, D., Cavallo, M. & Iacoviello, M. (2019), 'Oil price elasticities and oil price fluctuations', *Journal of Monetary Economics* **103**, 1–20.

Caldara, D. & Herbst, E. (2019), 'Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs', *American Economic Journal: Macroeconomics* **11**(1), 157–92.

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

Cesa-Bianchi, A. & Sokol, A. (2017), Financial shocks, credit spreads and the international credit channel, Working paper, Bank of England.

Chib, S. & Greenberg, E. (1995), 'Understanding the Metropolis–Hastings algorithm', *The American Statistician* **49**(4), 327–335.

Christiano, L. J., Eichenbaum, M. & Evans, C. L. (1999), Monetary policy shocks: What have we learned and to what end?, *in* J. B. Taylor & M. Woodford, eds, 'Handbook of Macroeconomics', Vol. 1 of *Handbook of Macroeconomics*, Elsevier, chapter 2, pp. 65–148.

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.

Conley, T. G., Hansen, C. B. & Rossi, P. E. (2012), 'Plausibly exogenous', *Review of Economics and Statistics* **94**(1), 260–272.

Drautzburg, T. (2016), A narrative approach to a fiscal DSGE model, Working Paper 16-11, Federal Reserve Board Philadelphia.

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

Faust, J., Swanson, E. T. & Wright, J. H. (2004), 'Identifying VARs based on high frequency futures data', *Journal of Monetary Economics* **51**(6), 1107–1131.

Gerko, E. & Rey, H. (2017), 'Monetary policy in the capitals of capital', *Journal of the European Economic Association* **15**(4), 721–745.

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

Giacomini, R., Kitagawa, T. & Read, M. (2019), Robust bayesian inference in proxy svars, Working Paper CWP23/19, Department of Economics, University College London.

Gilchrist, S. & Zakrajšek, E. (2012), 'Credit spreads and business cycle fluctuations', *American Economic Review* **102**(4), 1692–1720.

Herrera, A. M. & Rangaraju, S. K. (2020), 'The effect of oil supply shocks on u.s. economic activity: What have we learned?', *Journal of Applied Econometrics* **35**(2), 141–159.

Jarociński, M. & Karadi, P. (2020), 'Deconstructing monetary policy surprises—the role of information shocks', *American Economic Journal: Macroeconomics* **12**(2), 1–43.

Känzig, D. R. (2019), 'The macroeconomic effects of oil supply shocks: New evidence from opec announcements', *Available at SSRN 3185839* .

Kass, R. E. & Raftery, A. E. (1995), 'Bayes factors', *Journal of the American Statistical Association* **90**(430), 773–795.

Kilian, L. (2008), 'Exogenous oil supply shocks: How big are they and how much do they matter for the U.S. economy?', *The Review of Economics and Statistics* **90**(2), 216–240.

Kilian, L. (2009), 'Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market', *American Economic Review* **99**(3), 1053–1069.

Kilian, L. & Lütkepohl, H. (2017), *Structural Vector Autoregressive Analysis*, Cambridge University Press, in press.

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.

Kilian, L. & Murphy, D. P. (2014), 'The role of inventories and speculative trading in the global market for crude oil', *Journal of Applied Econometrics* **29**(3), 454–478.

Koop, G., Korobilis, D. et al. (2010), 'Bayesian multivariate time series methods for empirical macroeconomics', *Foundations and Trends® in Econometrics* **3**(4), 267–358.

Lakdawala, A. (2019), 'Decomposing the effects of monetary policy using an external instruments SVAR', *Journal of Applied Econometrics* **34**(6), 934–950.

Lanne, M. & Luoto, J. (2020), 'Identification of economic shocks by inequality constraints in Bayesian structural vector autoregression', *Oxford Bulletin of Economics and Statistics* **82**(2), 425–452.

Litterman, R. B. (1986), 'Forecasting with Bayesian vector autoregressions—five years of experience', *Journal of Business & Economic Statistics* **4**(1), 25–38.

Ludvigson, S. C., Ma, S. & Ng, S. (2020), Shock restricted structural vector-autoregressions, Working paper.

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer Science & Business Media.

Lütkepohl, H. & Woźniak, T. (2018), 'Bayesian Inference for Structural Vector Autoregressions Identified by Markov-Switching Heteroskedasticity'.

Mertens, K. & Montiel Olea, J. L. (2018), 'Marginal tax rates and income: New time series evidence', *The Quarterly Journal of Economics* **133**(4), 1803–1884.

Mertens, K. & Ravn, M. O. (2012), 'Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks', *American Economic Journal-Economic Policy* **4**(2), 145–181.

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.

Miranda-Agrippino, S., Ricco, G. et al. (2018), 'Identification with external instruments in structural vars under partial invertibility', *Manuscript, Bank of England* .

Montiel Olea, J. L., Stock, J. H. & Watson, M. W. (2020), 'Inference in SVARs identified with external instruments', *Journal of Econometrics* **forthcoming**.

Muirhead, R. J. (1982), *Aspects of Multivariate Statistical Theory*, Vol. 197, John Wiley & Sons.

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

Nevo, A. & Rosen, A. M. (2012), 'Identification with imperfect instruments', *Review of Economics and Statistics* **94**(3), 659–671.

Newell, R. G. & Prest, B. C. (2019), 'The unconventional oil supply boom: Aggregate price response from microdata', *Energy Economics* **40**(3), 1–30.

Newey, W. K. & West, K. D. (1987), 'A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance-matrix', *Econometrica* **55**(3), 703–708.

Nguyen, L. (2019), Bayesian inference in structural vector autoregression with sign restrictions and external instruments, Technical report.

Peersman, G. (2020), 'International food commodity prices and missing (dis)inflation in the euro area', *The Review of Economics and Statistics* **forthcoming**.

Piffer, M. & Podstawski, M. (2017), 'Identifying uncertainty shocks using the price of gold', *The Economic Journal* **128**(616), 3266–3284.

Ramey, V. (2016), Macroeconomic shocks and their propagation, *in* 'Handbook of Macroeconomics', Vol. 2, Elsevier, chapter 2, pp. 71–162.

Romer, C. D. & Romer, D. H. (2004), 'A new measure of monetary shocks: Derivation and implications', *American Economic Review* **94**(4), 1055–1084.

Sims, C. A. & Zha, T. (1998), 'Bayesian methods for dynamic multivariate models', *International Economic Review* **39**(4), 949–968.

Stock, J. H. & Watson, M. W. (2012), 'Disentangling the channels of the 2007-09 recession', *Brookings Papers on Economic Activity* **43**(1), 81–156.

Tierney, L. (1994), 'Markov chains for exploring posterior distributions', *The Annals of Statistics* **22**(4), 1701–1728.

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.

Uhrin, G. B. & Herwartz, H. (2016), Monetary policy shocks, set-identifying restrictions, and asset prices: A benchmarking approach for analyzing set-identified models, Technical report, University of Goettingen, Department of Economics.

Wieland, J. F. & Yang, M.-J. (2016), Financial dampening, Working Paper 22141, National Bureau of Economic Research.

Woźniak, T. & Droumaguet, M. (2015), Assessing monetary policy models: Bayesian inference for heteroskedastic structural VARs, Working Paper, University of Melbourne.

Zeev, N. B. (2018), 'What can we learn about news shocks from the late 1990s and early 2000s boom-bust period?', *Journal of Economic Dynamics and Control* **87**, 94–105.

A Proposal distribution used in the AR-MH algorithms

We describe the proposal distribution $p^\star(\beta; v, S)$ used by the AR-MH algorithm in Section 2.4 and 2.5 in more detail. We use the following notation. Let $\tilde{n} = n + k$, $e_{\tilde{n},j}$ be the jth column of $I_{\tilde{n}}$, $Q = \text{diag}(Q_1, Q_2)$ be a $\tilde{n} \times \tilde{n}$ orthogonal block diagonal matrix where Q_1 is orthogonal of size $n \times n$ and Q_2 orthogonal of size $k \times k$. Furthermore, Σ is a symmetric positive definite matrix dimension $\tilde{n}$. As mentioned in the main part of the paper, the structural impact matrix of the proxy-augmented SVAR, $\tilde{B}$, is parameterized as $\tilde{B} = \text{chol}(\Sigma)Q = PQ$ where $\text{chol}(\cdot)$ is the lower triangular Cholesky decomposition. If the external variable is assumed to be a valid instrument, we have specified zero restriction on $\tilde{B}$ as discussed in Section 2.2. We follow Arias, Rubio-Ramírez & Waggoner (2019) and denote the restrictions as:

$$J\tilde{B}e_{\tilde{n},j} = 0_{k\times 1} \text{ for } 1 \leq j \leq n-k, \tag{A.1}$$

$$JPQe_{\tilde{n},j} = JPL'Q_1e_{n,j} = 0_{k\times 1} \text{ for } 1 \leq j \leq n-k, \tag{A.2}$$

where $J = [0_{k\times n} : I_k]$ and $L = [I_n : 0_{n\times k}]$. That is, the exogeneity restrictions can be written as linear constraints on either $\tilde{B}$ or Q. Denote by $\tilde{z}_j$ the number of restrictions on the jth column of Q_1, which is k for $1 \leq j \leq n-k$ if the exogeneity constraints are imposed and 0 otherwise. Then, the proposal distribution in the AR-MH algorithm draws $\beta^\star$ by the following algorithm:

1. Draw $P = \text{chol}(\Sigma)$ where $\Sigma \sim i\mathcal{W}(v, S)$.

2. Generate $Q = \text{diag}(Q_1, Q_2)$ from a uniform distribution, subject to zero and sign restrictions, as in Arias, Rubio-Ramírez & Waggoner (2019):

   (a) For $1 \leq j \leq n$, draw $w_{1,j} = x_{1,j}/||x_{1,j}||$ with $x_{1,j} \sim \mathcal{N}(0, I_{n+1-j-\tilde{z}_j})$

   (b) For $1 \leq j \leq k$, draw $w_{2,j} = x_{2,j}/||x_{2,j}||$ with $x_{2,j} \sim \mathcal{N}(0, I_{k+1-j})$

   (c) Compute $Q_1 = [q_{1,1} : \cdots : q_{1,n}]$ recursively by setting $q_{1,j} = K_{1,j}w_{1,j}$, where $K_{1,j}$ is such that it forms a null space of the matrix $M_{1,j} = [q_{1,1} : \cdots : q_{1,j-1} : G(P)']'$

with $G(P) := JPL'$ and for $1 \leq j \leq n-k$. For $n-k+1 \leq j \leq n$, set $M_{1,j} = [q_{1,1} : \cdots : q_{1,j-1}]'$. This captures the exogeneity restrictions as in Section 2.2. If they do not hold (as discussed in Section 2.3), simply use $M_{1,j} = [q_{1,1} : \cdots : q_{1,j-1}]'$ for $1 \leq j \leq n$.

(d) Compute $Q_2 = [q_{2,1} : \cdots : q_{2,n}]$ recursively by setting $q_{2,j} = K_{2,j} w_{2,j}$ for $K_{2,j}$ such that it forms a null space of $M_{2,j} = [q_{1,1} : \cdots : q_{1,j-1}]'$ for $1 \leq j \leq k$.

(e) If the sign restrictions are satisfied, proceed. Otherwise, repeat step 2.

3. Set $\tilde{B}^\star = PQ$ and $\beta^\star = S_b \mathrm{vec}(\tilde{B}^\star)$.

Note that by construction $\Sigma = \tilde{B}\tilde{B}'$ and furthermore, $\tilde{B}$ will satisfy the desired zero block restrictions on the upper right part as well as on Φ if the exogeneity restrictions of equation (A.2) are imposed additionally.

In the following, we give the density implied by this proposal distribution. Denote the mapping $[w', \mathrm{vec}(\Sigma)']' \overset{f}{\to} \beta^\star$ and its inverse by $\beta^\star \overset{f^{-1}}{\to} [w', \mathrm{vec}(\Sigma)']'$, where $w = [w'_{1,1}, \ldots, w'_{1,n}, w'_{2,1}, \ldots, w'_{2,k}]'$. Then, a draw from $\beta^\star \sim p^\star(\beta; v, S)$ has density value:

$$p^\star(\beta; v, S) \propto \det(\tilde{B}^\star \tilde{B}^{\star\prime})^{-\frac{v+\tilde{n}+1}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}(S(\tilde{B}^\star \tilde{B}^{\star\prime})^{-1})\right) v_{f^{-1}}(\tilde{B}^\star), \tag{A.3}$$

where the first part comes from the inverse Wishart density of Σ, and $v_{f^{-1}}(\tilde{B})$ is the "volume element" as denoted in Arias et al. (2018), which accounts for the change in variables when transforming draws from Σ, Q to $\tilde{B}$. In our case, we have that following Theorem 2 of Arias et al. (2018):

$$v_{f^{-1}}(\tilde{B}) = |\det(J_{f^{-1}}(\tilde{B}) J_{f^{-1}}(\tilde{B})')|^{\frac{1}{2}}, \tag{A.4}$$

where $J_{f^{-1}}(\tilde{B})$ is the Jacobian of f^{-1} evaluated at $\tilde{B}$. Note that this holds only if S_b in $\beta^\star = S_b \mathrm{vec}(\tilde{B})$ is specified as to include all zero constraints, that is those on the upper right block of $\tilde{B}$, as well as those on Φ if exogeneity restrictions are specified as in equation (A.2).[13]

To ensure that the mappings f and f^{-1} are differentiable and one to one, we follow Appendix A.3 of Arias et al. (2018) to compute $K_{1,j}$ and $K_{2,j}$ by the QR decomposition using

[13] Note that otherwise, Theorem 3 of Arias et al. (2018) would apply.

the Gram Schmidt process. In order to evaluate the Jacobian, we use numerical derivatives of f^{-1}. Given that the dimension of $\beta^\star$ is usually relatively small, the computational costs are not very high.[14]

B Reconstructing and extending Kilian's oil supply shock

In Section 3.1, we have used a monthly series of oil supply shocks as in Kilian (2008). Only a quarterly time series for the 'exogenous' oil price shock from 1973Q2-2004Q3 is available on Lutz Kilian's homepage. A corresponding time series on the monthly frequency is not readily available and we would also like to use a more recent sample period. Therefore, we have reconstructed the monthly series shock series using updated oil production data from the US Energy Information Administration (see Monthly Energy Review, Table 11.1a and Table 11.b, https://www.eia.gov/totalenergy/data/monthly/index.php). As described in Kilian (2008) the construction is based on computing oil supply shortfalls based on counterfactual oil growth rates for countries that have been exposed to exogenous oil supply disruption caused e.g. by geopolitical turmoils and wars (see the Kilian paper for a precise description of the shock construction methodology). Reconstructing the series allows us to extend the shock measure to the sample 1973M02-2017M12 used in our paper. For this period, we have added to more exogenous events that affected oil production in Libya. The first event is related to the Libyan war in 2011, which led to a sharp drop of oil production. We start the counterfactual in March 2011 and it ends in April 2012. Since no other OPEC country was affected by the civil war, the benchmark of all OPEC countries' production minus Libya. The second event was triggered in May 2013 by a series of militia attacks that started the civil unrest. Consequently, we start a second counterfactual for Libya starting in that period. Using the information from the oil market reports, it is clear that Libya never managed to resolve the civil unrest with two rival governments in the country. For this reason the counterfactual continues until the end of our sample in 2017M12. For

[14]This is particularly an advantage over Arias, Rubio-Ramírez & Waggoner (2019), given that in their approach, the mappings underlying the Jacobian is of various magnitudes larger since they include the whole SVAR parameters, that is also the autoregressive parameters.

this second event, we have removed Iran from the benchmark group in the period May 2013 to December 2015, as Iran faced international sanctions that led to problems for the oil industry. For the time between May 2016 until the end of our sample, sanctions on Iran were less stringent due to a political deal and consequently, we have included Iran in the benchmark during this period. Starting in January 2016, we have also removed Venezuela from the benchmark as this country faced its own problems related to a political and economic crisis.

The resulting shock series is shown in Figure 4. Note that transforming our shock series to quarterly frequency and comparing it with the original Kilian quarterly shock series shows a correlation of about 0.995.

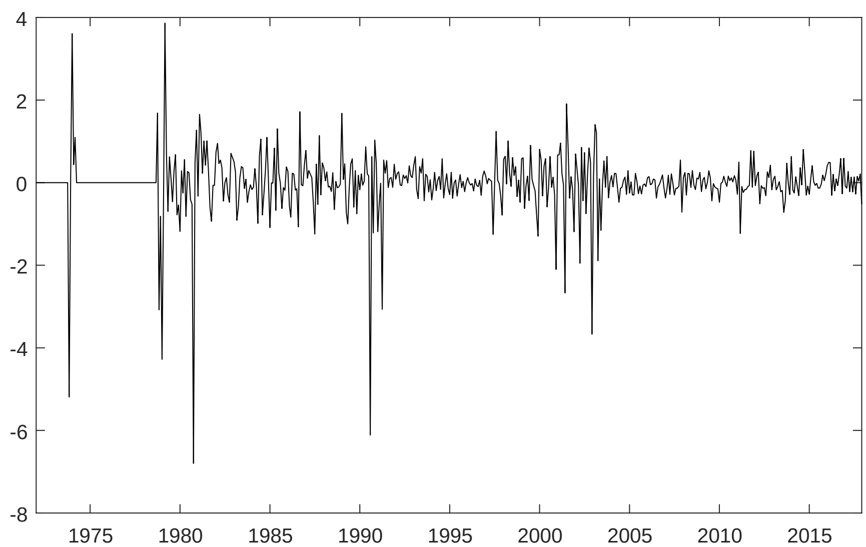

Figure 4: Exogenous oil production shortfall series as in Kilian (2008) (extended). Sample period: 1973M01 - 2017M12.

UNIVERSITY OF KONSTANZ
Department of Economics

Universitätsstraße 10
78464 Konstanz
Germany

Phone: +49 (0) 7531-88-3713
Fax: +49 (0) 7531-88-3130

www.wiwi.uni-konstanz.de/econdoc/working-paper-series/

University of Konstanz
Department of Economics