Block-recursive non-Gaussian structural vector autoregressions: Identification, Efficiency, and Moment Selection*

Sascha Alexander Keweloh**, Stephan Hetzenecker***

February 18, 2022

We combine block-recursive restrictions with higher-order moment conditions to identify and estimate non-Gaussian structural vector autoregressions. For a given block-recursive structure, we derive a set of identifying moment conditions based on the assumption of uncorrelated shocks across blocks and mean independent shocks within the blocks. We then obtain overidentifying moment conditions from the assumption of independent shocks and show that these conditions can decrease the asymptotic variance of the estimator. In particular, we derive conditions under which the frequently applied estimator based on the Cholesky decomposition is inefficient. We use a LASSO-type GMM estimator to select the relevant and valid overidentifying moment conditions in a data-driven way. A Monte Carlo experiment illustrates the improved performance of the proposed estimator. In the empirical illustration, we take advantage of the block-recursive framework to analyze the impact of speculative shocks in the oil market.

JEL Codes: C32, E52, E44

Keywords: SVAR, identification, non-Gaussianity, block-recursive, oil market, expectations

* The authors gratefully acknowledge the computing time provided on the Linux HPC cluster at Technical University Dortmund (LiDO3), partially funded in the course of the Large-Scale Equipment Initiative by the German Research Foundation (DFG) as project 271512359.

** TU Dortmund University / RGS Econ, E-Mail: sascha.keweloh@tu-dortmund.de, Postal Address: TU Dortmund, Department of Economics, D-44221 Dortmund. Financial support from the German Science Foundation, DFG - SFB 823, is gratefully acknowledged.

*** University Duisburg-Essen / RGS Econ, E-Mail: stephan.hetzenecker@vwl.uni-due.de, Postal Address: University Duisburg-Essen, Universitätsstraße 12, D-45117 Essen. Financial support by the Ruhr Graduate School in Economics is gratefully acknowledged.

1 Introduction

Identification of a structural vector autoregression (SVAR) requires to assume an a priori structure of the model. Traditionally, identification is based on imposing structure on the interaction of the variables, ideally derived from macroeconomic theory (e.g., short-run restrictions Sims (1980) or long-run restrictions Blanchard and Quah (1993)). However, uncontroversial theoretical restrictions are rare. More recently, data-driven approaches allow to identify the SVAR without imposing any restrictions on the interaction. Instead, identification is achieved by imposing structure on the stochastic properties of the shocks (e.g., time-varying volatility as discussed in Rigobon (2003), Lanne et al. (2010), Lütkepohl and Netšunajev (2017), and Lewis (2021) or non-Gaussian and independent shocks as discussed in Gouriéroux et al. (2017), Lanne et al. (2017), Lanne and Luoto (2021), Keweloh (2021b), and Guay (2021)).

Traditional identification approaches may appear unnecessarily restrictive compared to novel data-driven approaches. However, Olea et al. (2022) stress that these data-driven approaches rely on information in higher moments, while traditional approaches only rely on second moments. The data-driven approaches are sensitive to the imposed statistical properties on the higher moments, while the traditional approaches are not and hence, are robust to these statistical properties. Additionally, they argue that using economic theory for identification is a feature and not a handicap and conclude that traditional identification approaches remain relevant.

We agree with their reasoning and recognize the advantages of identification approaches based on economic theory. However, in many applications we can derive some, but not sufficiently many convincing restrictions from economic theory to ensure identification. Therefore, with a traditional purely restriction based approach, even the most plausible restrictions are worthless if there are not sufficiently many. We propose a Generalized Method of Moments (GMM) estimator that combines the traditional identification approach based on restrictions with the more recent data-driven approach based on non-Gaussianity. Our approach allows to impose a block-recursive structure, meaning that shocks in a given block only influence variables in the same block or blocks ordered below. The block-recursive structure seems plausible in many macroeconomic applications. Examples include applications analyzing (i) the interaction of macroeconomic and financial variables, where the former respond sluggishly while the latter respond quickly, or (ii) the interaction of small and large open economies, where large economies may have an immediate impact on small economics but not vice versa. Additionally, the block-recursive structure nests two important special cases: a recursive and an unrestricted SVAR.

Identification based on higher moments and non-Gaussian shocks oftentimes relies on the assumption of independent shocks which is criticized as too restrictive (see, e.g., Kilian and Lütkepohl (2017, Chapter 14)). Importantly, our identification result does not rely on independent shocks but is robust in the sense that it allows for various kinds of dependencies of the shocks. In particular, for a given block-recursive structure identification of the shocks within a given block is based on a small (subset) of cokurtosis conditions derived from mean independence of the shocks in the corresponding block.[1] Therefore, identification within a block follows from Lanne and Luoto (2021). Moreover, the impact of the shocks in one block on variables in another block is identified based only on covariance conditions and not on higher-order moment conditions and requires only uncorrelated shocks. Therefore, imposing a finer block-recursive structure reduces the dependency of identification on higher-order moment conditions.

However, if the shocks are independent, using only the set of identifying conditions, which is derived from mean independent shocks within blocks and uncorrelated shocks across blocks, can be inefficient. To demonstrate this, we prove that in a recursive SVAR with independent shocks the set of overidentifying higher-order moment conditions can contain additional information and allows to decrease the asymptotic variance of the GMM estimator.[2] Efficient estimation requires to detect and select the valid and relevant overidentifying conditions. To this end, Lanne and Luoto (2021) suggest to calculate the information and moment selection criteria proposed by Andrews (1999) and Hall et al. (2007) for all possible combinations of moment conditions. However, they note that this approach becomes infeasible in higher-dimensional SVARs.

In a general GMM setup, Cheng and Liao (2015) propose a LASSO-type GMM estimator, hereafter referred to as the penalized GMM estimator (pGMM), which consistently selects only rele-

[1] A common critique to the assumption of independent shocks is that it does not allow for multiple shocks to be driven by the same volatility process. Thereby, it rules out a case which may be encountered for some macroeconomic shocks. However, mean independent shocks and, in particular, the set of cokurtosis conditions used for identification allow for these kinds of dependencies.

[2] Note that this is not trivial. For example, in a linear regression model $y_t = \beta_1 x_t + \epsilon_t$ the GMM estimator with the moment condition $E[x_t \epsilon_t] = 0$ is identified and efficient under (conditional) homoscedastic errors. Therefore, including additional higher-order moment conditions like $E[x_t^2 \epsilon_t] = 0$ does not decrease the asymptotic variance of the GMM estimator even if the shocks or variables are non-Gaussian.

vant and valid overidentifying conditions in a data-driven way. We apply the pGMM estimator to the block-recursive SVAR to exploit potential efficiency gains from overidentifying moment conditions. Our block-recursive SVAR pGMM estimator is consistent, asymptotically normal and as efficient as the asymptotically efficient block-recursive SVAR GMM estimator, including all valid and relevant overidentifying moment conditions. Importantly, these properties also hold if there are invalid overidentifying moment condition which could arise due to dependent structural shocks. Additionally, the pGMM estimator refrains from selecting valid but redundant overidentifying conditions which would neither increase nor decrease the asymptotic variance of the estimator but lead to imprecise estimates in small samples due to a many moments problem.

Guay (2021) also proposes to combine restrictions with non-Gaussian identification. In particular, he tests which shocks of the SVAR are identified based on non-Gaussianity and subsequently, his approach only uses restrictions to identify the remaining part of the SVAR. In this approach, if all shocks are non-Gaussian, no restrictions have to be used and the SVAR can be estimated solely by higher-order moment conditions. Consequently, the identification approach relies as heavily on non-Gaussianity as possible and as little on restrictions as necessary. In contrast to that, our identification approach relies as much as possible on economically justified restrictions and on non-Gaussianity only when needed. To be precise, the more block-recursiveness restrictions the researcher imposes, the less identification depends on higher order-moment conditions.

We conduct two Monte Carlo experiments. In the first one, we demonstrate that the performance of a purely data-driven estimator based on non-Gaussianity deteriorates substantially with both a decreasing sample size and an increasing model size. However, exploiting the block-recursive order can mitigate this performance decline. In the second Monte Carlo experiment, we illustrate that the pGMM estimator successfully selects relevant moment conditions and increases the finite sample performance compared to other block-recursive SVAR estimators for a given block-recursive structure.

We use the block-recursive SVAR pGMM estimator to analyze the impact of oil supply and oil demand shocks, including speculative oil supply and demand shocks, on the oil price. In his seminal work, Kilian (2009) highlights that it is necessary to distinguish between oil supply and demand shocks rather than including solely an oil price shock in the SVAR for the oil market. However, oil prices are not only affected by supply and demand shocks, but also by speculative shocks causing shifts in the expectations of forward-looking traders (see, e.g., Baumeister and Kilian (2016)). In particular, new oil production technologies, anticipated wars, or news about oil discoveries or about the (future) state of the economy can shift expectations of future oil supply and future oil demand. The studies of Kilian and Murphy (2014), Juvenal and Petrella (2015), Byrne et al. (2019), and Moussa and Thomas (2021) extend the original oil market SVAR from Kilian (2009) to include speculative shocks. We contribute to this literature by explicitly distinguishing between speculative supply and speculative demand shocks.

The remainder of the paper is organized as follows: Section 2 reviews the SVAR and different identification schemes. Section 3 introduces the block-recursive SVAR. Section 4 derives identifying and overidentifying moment conditions in a block-recursive SVAR, analyzes which of the overidentifying conditions are redundant or relevant in a recursive SVAR, and describes the pGMM estimator. In Section 5, we present the Monte Carlo experiments. In Section 6, we use the proposed block-recursive estimator to analyze the impact of flow and speculative supply and demand shocks in the oil market. Section 7 concludes.

2 Overview SVAR

This section briefly recalls the identification problem and common identification approaches for SVAR models. A detailed overview can be found in Kilian and Lütkepohl (2017). Consider the SVAR

$$y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + B_0 \varepsilon_t, \tag{1}$$

with parameter matrices $A_1, ..., A_p \in \mathbb{R}^{n \times n}$, an invertible matrix B_0, an n-dimensional vector of time series $y_t = [y_{1,t}, ..., y_{n,t}]'$ and an n-dimensional vector of i.i.d. structural shocks $\varepsilon_t = [\varepsilon_{1,t}, ..., \varepsilon_{n,t}]'$ with mean zero and unit variance.

W.l.o.g. we focus on the simultaneous interaction of the SVAR given by

$$u_t = B_0 \varepsilon_t, \tag{2}$$

with the reduced form shocks $u_t = y_t - A_1 y_{t-1} - ... - A_p y_{t-p}$, which can be estimated consistently by OLS. The reduced form shocks are an unknown mixture B_0 of the unknown structural shocks ε_t. So far, neither the mixing matrix B_0 nor the structural shocks ε_t are identified. To see this, define the unmixed innovations $e(B)$ as the innovations obtained by unmixing the reduced form shocks with some matrix B

$$e_t(B) := B^{-1} u_t. \tag{3}$$

Note that for $B = B_0$, the unmixed innovations are equal to the structural shocks ε_t, i.e., $e_t(B_0) = \varepsilon_t$. Additionally, given an estimate $\hat{B}$ of B_0 we refer to $e_t(\hat{B})$ as the estimated structural shocks. The true structural shocks ε_t and the true mixing matrix B_0 are unknown and without imposing further structure, we cannot verify whether the mixing matrix B and the unmixed innovations $e_t(B)$ are equal to the true mixing matrix B_0 and the true structural shocks ε_t.

To identify B_0 and the shocks ε_t, the researcher has to impose structure on the SVAR. The structure can be specified in two ways: We may

(i) impose more structure on the interaction of the shocks (see Sims (1980) for short-run restrictions, Blanchard (1989) for long-run restrictions, and Uhlig (2005) for sign restrictions),

(ii) impose more structure on the stochastic properties of the structural shock (see Lanne et al. (2010) for time-varying volatility or Gouriéroux et al. (2017), Lanne et al. (2017), Lanne and Luoto (2021) Keweloh (2021b), and Guay (2021) for non-Gaussian shocks).

Imposing structure on the stochastic properties of the shocks can be used to derive conditions for the unmixed innovations, while imposing structure on the interaction narrows the space of possible mixing matrices used to unmix the reduced form shocks.

In applied work, the probably most frequently imposed structure are uncorrelated structural shocks (meaning $\varepsilon_{i,t}$ is restricted to be uncorrelated with $\varepsilon_{j,t}$ for $i \neq j$) and a recursive interaction (meaning restricting B_0 such that $b_{ij} = 0$ for $i < j$ where b_{ij} denotes the element at row i and column j of B_0). Uncorrelated shocks with unit variance can be used to derive $(n+1)n/2$ (co-)variance conditions from $I = E[\varepsilon_t \varepsilon_t'] \overset{!}{=} E[e_t(B) e_t(B)']$. A recursive interaction implies that $n(n-1)/2$ parameters of B_0 are known a priori, leaving only $(n+1)n/2$ unknown parameters in the mixing matrix B. It is then straightforward to show that, if the remaining $(n+1)n/2$ parameters of the restricted B matrix generate unmixed innovations $e_t(B)$ which satisfy the $(n+1)n/2$ (co-)variance conditions, the matrix B has to be equal to B_0 and, hence, the unmixed innovations are equal to the structural shocks, meaning the SVAR is identified.[3]

However, economic theory rarely allows to derive the required $n(n-1)/2$ parameter restrictions to ensure identification. More recently, identification methods based on non-Gaussian and independent shocks have been put forward in the literature (see Gouriéroux et al. (2017), Lanne et al. (2017), Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021)). These identification schemes do not require to impose any restrictions on the impact of the shocks, in particular on the matrix B_0. Instead, the researcher has to impose structure on the stochastic properties of the shocks. If the structural shocks are not only mutually uncorrelated but mutually independent, we can derive additional moment conditions. For example, independent and mean zero shocks imply that all entries of coskewness matrices $E\left[\varepsilon_t \varepsilon_t' \varepsilon_{i,t}\right]$ for $i = 1, \ldots, n$ are zero except for the ith diagonal element, which contains the (unknown) skewness of the shock $\varepsilon_{i,t}$. Hence, we can exploit that the mixing matrix B has to generate unmixed innovations, which satisfy the coskewness moment conditions derived from $E\left[\varepsilon_t \varepsilon_t' \varepsilon_{i,t}\right] \overset{!}{=} E\left[e_t(B) e_t(B)' e_{i,t}(B)\right]$. Similarly, we can use that the mixing matrix B has to generate unmixed innovations which satisfy the cokurtosis moment conditions derived from $E\left[\varepsilon_t \varepsilon_t' \varepsilon_{i,t} \varepsilon_{j,t}\right] \overset{!}{=} E\left[e_t(B) e_t(B)' e_{i,t}(B) e_{j,t}(B)\right]$.

3 Imposing structure in a SVAR

This section introduces the framework of the block-recursive SVAR. First, we discuss various structures of the interaction of the shocks allowed in this framework and then, assumptions on the stochastic properties of the shocks.

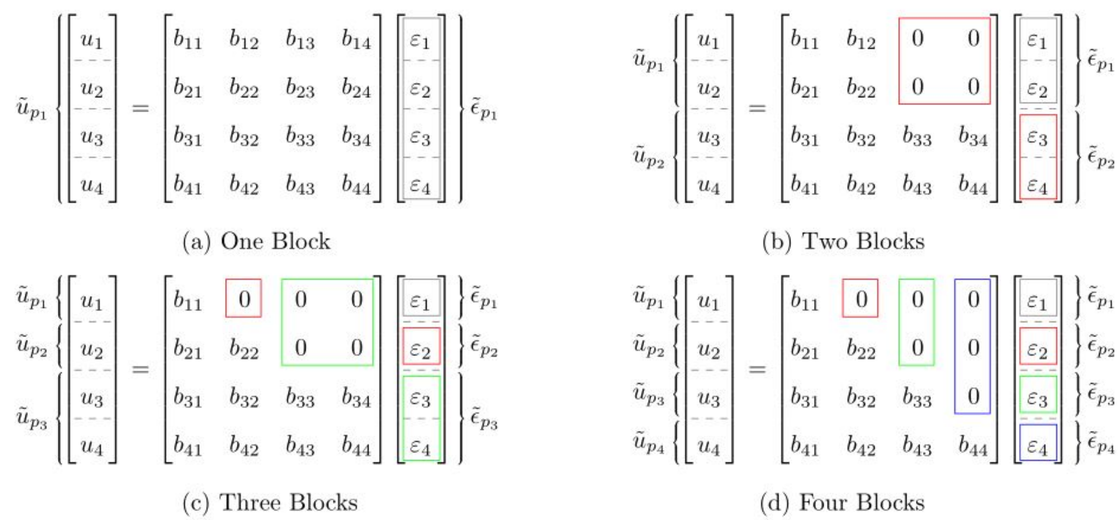
3.1 Imposing structure on the interaction of shocks

Traditionally, identification of a SVAR is based on the structure imposed on the interaction of the shocks (see Section 2). These restriction based approaches require restrictions on the interaction

[3]Note that this GMM approach is equivalent to the the frequently used estimator obtained by applying the Cholesky decomposition to the variance-covariance matrix of the reduced form shocks.

of the shocks to ensure identification, e.g., a recursive structure. The reasoning behind a recursive structure is oftentimes the prejudice that some variables, e.g., some macroeconomic variables like inflation, tend to move slowly, while other variables, e.g. financial variables like stock prices, react faster. However, in practice this intuitive reasoning oftentimes allows to order only some, but not all variables recursively. This motivates us to consider the block-recursive SVAR, meaning that the structural shocks are ordered in blocks of consecutive shocks and each structural shock can simultaneously affect all variables in the same block and in blocks ordered below but not variables in blocks ordered above.[4] Figure 1 shows different block-recursive structures in a SVAR with four variables. The examples show that a block-recursive structure generalizes the unrestricted

Figure 1: Examples of Different Block-Recursive SVAR Models.



Note: The figure illustrate how the the block structure can be defined by the structural shocks and our definition of $\tilde{\varepsilon}_{p_i}$ and $\tilde{u}_{p_i}$, $i = 1, \ldots, m$.

SVAR and the fully-recursive SVAR and includes both as extreme cases.

We now introduce the notation for the block-recursive SVAR. Suppose that the structural shocks can be ordered into $m \leq n$ blocks of consecutive shocks. Let the indices $p_1 = 1 < p_2 < \ldots < p_m \leq n$ denote the beginning of a new block and for a given block p_i let $\tilde{\varepsilon}_{p_i,t}$ and $\tilde{u}_{p_i,t}$ denote

[4]Zha (1999) derives identifying restrictions for the block-recursive SVAR. The author restricts not only the simultaneous interaction, but also the lagged interaction. Our proposed block-recursive structure affects only the simultaneous interaction, while the lagged interaction remains unrestricted.

the vectors of all structural and reduced form shocks in the ith block, such that

$$\tilde{\varepsilon}_{p_i,t} := \left[\varepsilon_{p_i,t}, \varepsilon_{p_i+1,t}, \ldots, \varepsilon_{p_{i+1}-1,t}\right]' \quad \text{and} \quad \tilde{u}_{p_i,t} := \left[u_{p_i,t}, u_{p_i+1,t}, \ldots, u_{p_{i+1}-1,t}\right]', \tag{4}$$

where $p_{m+1} := n + 1$ for ease of notation. Moreover, let l_i denote the number of shocks in block i for $i = 1, ..., m$. The vector of all structural shocks ε_t can then be decomposed into the m blocks $\varepsilon_t = [\tilde{\varepsilon}'_{p_1,t}, \ldots, \tilde{\varepsilon}'_{p_m,t}]'$ and the reduced form shocks can be decomposed analogously into $u_t = [\tilde{u}'_{p_1,t}, \ldots, \tilde{u}'_{p_m,t}]'$. The SVAR is block-recursive with $m \leq n$ blocks with $p_1 = 1 < p_2 < \ldots < p_m \leq n$, if shocks in the ith block have no simultaneous impact on reduced form shocks in blocks j with $j < i$ such that for $i = 1, \ldots, m$

$$b_{ql} = 0, \text{ for } l \geq p_i \text{ and } q < p_i. \tag{5}$$

Any block-recursive structure can be described by the following assumption.

Assumption 1. *(Block-recursive interaction.)*
For $m \leq n$ blocks with $p_1 = 1 < p_2 < \ldots < p_m \leq n$ and $q, l = 1, ..., n$ let $B_0 \in \mathbb{B}_{brec} \equiv \mathbb{B}_{brec}(p_1, \ldots, p_m) := \{B \in \mathbb{B} | \, b_{ql} = 0 \text{ if } \exists p_i \in \{p_1, ..., p_m\} \text{ with } l \geq p_i \text{ and } q < p_i\}$.

3.2 Imposing structure on the stochastic properties of shocks

Imposing structure according to Assumption 1 on the interaction is not sufficient to ensure identification and further assumptions on the dependence and potential non-Gaussianity of the shocks are required. In the following, we discuss different structures imposed on the mutual dependencies of the shocks.

Almost all identification approaches at least assume uncorrelated structural shocks such that $E[\varepsilon_{i,t}\varepsilon_{j,t}] = E[\varepsilon_{i,t}] E[\varepsilon_{j,t}]$ for $i \neq j$.[5] Uncorrelated shocks are justified by the idea that a given structural shock contains no information on other structural shocks, e.g., a structural monetary policy shock should not depend on other structural shocks. In general, imposing uncorrelated structural shocks does not rule out that the structural shocks are dependent. If they are dependent, the interpretation of the estimated SVAR via impulse response functions can be

[5] Proxy-variable identification approaches are different and instead assume that structural shocks are uncorrelated with an external proxy variable (see, e.g., Stock and Watson (2012), or Mertens and Ravn (2013)).

misleading. For example, consider the two random variables $\varepsilon_1 \sim \mathcal{N}(0,1)$ and $\varepsilon_2 = \varepsilon_1^2 - 1$ such that both random variables are uncorrelated, but dependent. Policy analysis based on impulse response functions typically uses the ceteris paribus assumption that only a single shock varies, while the other shocks remain unchanged. In the example above, both shocks are uncorrelated, but nevertheless always move simultaneously. Therefore, uncorrelated structural shocks are not sufficient to guarantee that the ceteris paribus assumption holds.

A more rigorous implementation of the idea that a given shock contains no information on other shocks is to assume independent shocks such that $E[h(\varepsilon_{i,t})g(\varepsilon_{j,t})] = E[h(\varepsilon_{i,t})]\,E[g(\varepsilon_{j,t})]$ for $i \neq j$ and any bounded, measurable functions $g(\cdot)$ and $h(\cdot)$. If shocks are independent, a structural shock cannot contain any information on any other structural shock. Therefore, independent structural shocks justify the ceteris paribus interpretation used in policy analysis based on impulse response functions. However, several authors argue that the assumption of independent structural shocks is too strong (cf. Kilian and Lütkepohl (2017, Chapter 14), Lanne and Luoto (2021), Lanne et al. (2021), or Olea et al. (2021)). In particular, independence of the shocks implies that also the volatility processes of the shocks are independent, which may be too restrictive for some macroeconomic applications. For example, suppose that $\tilde{\varepsilon}_{1,t}$ and $\tilde{\varepsilon}_{2,t}$ are drawn independently of each other and represent unscaled structural shocks. Moreover, in each period an additional volatility shock v_t is drawn independently of the other shocks and the structural shocks are given by $\varepsilon_{1,t} = \tilde{\varepsilon}_{1,t} v_t$ and $\varepsilon_{2,t} = \tilde{\varepsilon}_{2,t} v_t$. These structural shocks are uncorrelated, but dependent since the variance of one shock contains information on the variance of the other shock.

A compromise between the two extreme cases of uncorrelated and independent shocks is the assumption of mean independent shocks, such that $E[\varepsilon_{i,t} g(\varepsilon_{j,t})] = E[\varepsilon_{i,t}]\,E[g(\varepsilon_{j,t})]$ for $i \neq j$ with a bounded, measurable function $g(\cdot)$. If shocks are mean independent, a structural shock cannot contain any information about the mean of other structural shocks. Mean independent shocks can justify the ceteris paribus assumption used in impulse response analysis and at the same time allow for dependent volatility processes. In particular, the two shocks $\varepsilon_{1,t} = \tilde{\varepsilon}_{1,t} v_t$ and $\varepsilon_{2,t} = \tilde{\varepsilon}_{2,t} v_t$ defined above are mean independent since a given shock contains no information on the mean of the other shock.

Imposing structure on the dependence of the structural shocks allows to derive moment conditions (see, e.g., Lanne and Luoto (2021), Keweloh (2021b), or Guay (2021)). For $i, j, k, l = 1, ..., n$ we define the following moment conditions:

$$\begin{aligned}
&\text{Variance:} && E[e(B)_{i,t}^2] - 1 = 0 && && (6)\\
&\text{Covariance:} && E[e(B)_{i,t} e(B)_{j,t}] = 0, && \text{for } i < j && (7)\\
&\text{Coskewness:} && E[e(B)_{i,t}^2 e(B)_{j,t}] = 0, && \text{for } i \neq j && (8)\\
& && E[e(B)_{i,t} e(B)_{j,t} e(B)_{k,t}] = 0, && \text{for } i < j < k && (9)\\
&\text{Cokurtosis:} && E[e(B)_{i,t}^3 e(B)_{j,t}] = 0, && \text{for } i \neq j && (10)\\
& && E[e(B)_{i,t}^2 e(B)_{j,t} e(B)_{k,t}] = 0, && \text{for } i \neq j,\ i \neq k,\ j < k && (11)\\
& && E[e(B)_{i,t} e(B)_{j,t} e(B)_{k,t} e(B)_{l,t}] = 0, && \text{for } i < j < k < l && (12)\\
& && E[e(B)_{i,t}^2 e(B)_{j,t}^2] - 1 = 0, && \text{for } i < j && (13)
\end{aligned}$$

The variance conditions in Equation (6) follow from the unit variance normalization. The remaining conditions are derived from different assumptions on the dependence of the structural shocks. In particular, uncorrelated structural shocks only imply the covariance conditions in Equation (7). Mean independent shocks additionally imply the coskewness conditions in Equation (8) and (9) and the cokurtosis conditions in Equation (10)-(12). In addition, the symmetric cokurtosis conditions in Equation (13) follow from independent shocks.

Moreover, note that if all structural shocks are Gaussian, the conditions in Equation (8)-(13) do not contain information beyond the information contained in the variance and covariance conditions.

4 Estimation of a block-recursive SVAR

In this section, we combine identification based on recursiveness restrictions and non-Gaussian shocks. First, for a given block-recursive structure we derive corresponding identifying asymmetric cokurtosis conditions based on mean independent shocks within the blocks. Importantly, identification is achieved without many higher-order moment conditions and holds under fairly general conditions on the dependencies of the shocks. Second, we show that additional overidentifying higher-order moment conditions implied by independent shocks can decrease the asymptotic variance of the estimator if the imposed structure is correct. Third, we propose to use a LASSO-type GMM estimator to select the valid and relevant overidentifying higher-order moment conditions in a data-driven way. Consistency of the estimator only relies on the identifying moment conditions and, thus, is robust to various kinds of dependencies of the shocks. Furthermore, it can exploit efficiency gains from valid and relevant overidentifying conditions and ignore noise from valid but redundant overidentifying conditions.

4.1 Identification

In this section, we show that identification in a block-recursive SVAR can be achieved by the variance and covariance conditions in Equation (6) and (7) and the asymmetric cokurtosis conditions in Equation (10) corresponding to innovations in the same block. The identification result is robust in the sense that it allows for various sorts of dependencies of the shocks. To be clear, shocks in different blocks only need to be uncorrelated and shocks in the same block only need to fulfill the asymmetric cokurtosis conditions.

Let $E[f_{\mathbf{2}}(B, u_t)] = 0$ contain all variance and covariance conditions in Equation (6) and (7) and let $E[f_{\mathbf{4_{p_k}}}(B, u_t)] = 0$ contain all asymmetric cokurtosis conditions from Equation (10) corresponding to shocks in block k, e.g., $E[e(B)^3_{i,t} e(B)_{j,t}] = 0$ for $i, j = p_k, ..., p_{k+1} - 1$ and $i \neq j$. We define the identifying moment conditions as

$$E[f_{\mathbf{N}}(B, u_t)] := E \begin{bmatrix} f_{\mathbf{2}}(B, u_t) \\ f_{\mathbf{4_{p_1}}}(B, u_t) \\ \vdots \\ f_{\mathbf{4_{p_m}}}(B, u_t) \end{bmatrix} = 0. \tag{14}$$

In the following, we simplify the notation for moment conditions, e.g., we write $E[f_{\mathbf{N}}(B, u_t)]$ instead of $E[f_{\mathbf{N}}(B, u_t)] = 0$. Note that the identifying moment conditions do not contain asymmetric cokurtosis conditions of shocks in different blocks, e.g., the moment conditions $E[e(B)^3_{i,t} e(B)_{j,t}]$ for shocks $e(B)_{i,t}$ and $e(B)_{j,t}$ in different blocks are not contained in $E[f_{\mathbf{N}}(B, u_t)]$. The conditions $E[f_{\mathbf{N}}(B, u_t)]$ can be justified by the following assumption.

Assumption 2. *(Block-recursive mean independence.)*
For $m \leq n$ blocks with $p_1 = 1 < p_2 < \ldots < p_m \leq n$,

(i) all shocks are uncorrelated, i.e., $E[\varepsilon_{i,t}\varepsilon_{j,t}] = 0$ for $i \neq j$.

(ii) all shocks within the same block are mean independent, i.e., $E[\varepsilon_{i,t}|\varepsilon_{-i,t}] = 0$ for $i \in \{p_k, p_k + 1, ..., p_{k+1} - 1\}$ and $-i = \{p_k, p_k + 1, ..., p_{k+1} - 1\} \backslash i$ for $k = 1, \ldots, m$.

The identifying moment conditions contain n variance conditions, $n(n-1)/2$ covariance conditions and $\sum_{k=1}^{m} l_k(l_k - 1)/2$ asymmetric cokurtosis conditions, where $l_k := p_{k+1} - p_k$ denotes the number of shocks in block k. Therefore, each additional specified block refines the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ such that they contain fewer higher-order moment conditions. In the extreme case when the SVAR is specified recursively, meaning each block contains only one variable, the identifying moment conditions contain no higher-order moment conditions. In the other extreme case of a single block containing all variables, the identifying moment conditions contain all $n(n-1)$ asymmetric cokurtosis conditions and are similar to the conditions proposed in Lanne and Luoto (2021).[6]

The following proposition shows that the identifying moment conditions are sufficient to locally identify the block-recursive SVAR.

Proposition 1. *(Identification in the block-recursive SVAR.)*
Let $u_t = B_0\varepsilon_t$ with $m \leq n$ blocks and $B_0 \in \mathbb{B}_{brec} \equiv \mathbb{B}_{brec}(p_1, \ldots, p_m)$ such that Assumption 1 holds. Moreover, suppose that Assumption 2 holds. If at most one structural shock in each block has zero excess kurtosis, the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)] = 0$ locally identify $B = B_0$ for $B \in \mathbb{B}_{brec}$.

Proof. The proof recursively applies the identification result from Lanne and Luoto (2021) and can be found in Appendix A.3. □

[6]Lanne and Luoto (2021) propose to select $n(n-1)/2$ asymmetric cokurtosis conditions, which is sufficient for local identification if none of the asymmetric conditions does include the third power of a Gaussian shock. They advocate to rely on a moment selection criterion to avoid including redundant conditions or conditions of Gaussian shocks. Additionally, Lanne and Luoto (2021) note that including all $n(n-1)$ asymmetric cokurtosis conditions ensures local identification even if conditions related to Gaussian shocks are included. We argue that the degree of overidentification remains reasonably small even if we include all asymmetric cokurtosis conditions and therefore, including redundant conditions can be expected to be rather harmless. For example, in a SVAR with four variables and no restrictions the identifying moment conditions consists of 22 conditions to identify 16 parameters. Thus, we suggest to use all asymmetric cokurtosis conditions in order to avoid the cumbersome process of selecting a subset of the conditions.

In Proposition 1 the impact of shocks on variables in different blocks is identified based on covariance conditions. The interaction of shocks on variables within the same block is identified based on asymmetric cokurtosis conditions and the local identification result of Lanne and Luoto (2021). Local identification means that the moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ identify B_0 in a small neighborhood of B_0 (see Hall (2005)). Importantly, the proposition also holds for different higher-order moment conditions ensuring identification within the blocks. For example, the identifying conditions $E[f_{\mathbf{N}}(B, u_t)]$ could contain all variance-covariance, coskewness and cokurtosis conditions implied by independent structural shocks for each block. In this case, global identification up to sign and permutation within each block follows from Keweloh (2021b).

Without further restrictions, data-driven approaches relying on non-Gaussian and independent shocks can only ensure identification up to sign and permutation. This means that the order and sign of the shocks in the impulse response functions is not identified. In practice, the researcher has to manually assign labels to the shocks. Restricting the solution to a given block-recursive structure simplifies the permutation or labeling problem. In particular, shocks can only be permuted inside blocks. For instance, in example (b) in Figure 1 shocks from the second block cannot be permuted into the first block since this violates the block-recursive structure. Therefore, specifying a finer block-recursive structure simplifies the labeling of the shocks.

Define the block-recursive SVAR GMM estimator which minimizes the variance, covariance and the asymmetric cokurtosis conditions over the set of block-recursive matrices as

$$\hat{B}_{\mathbf{N}} := \underset{B \in \mathbb{B}_{brec}}{\arg\min}\, g_{\mathbf{N}}(B)' W_{\mathbf{N}} g_{\mathbf{N}}(B), \tag{15}$$

with a suitable weighting matrix $W_{\mathbf{N}}$ and $g_{\mathbf{N}}(B) := 1/T \sum_{t=1}^{T} f_{\mathbf{N}}(B, u_t)$. Consistency and asymptotic normality follow from the identification result in Proposition 1 and standard assumptions including valid moment conditions implied by the dependence structure imposed in Assumption 2. That is,

$$\hat{B}_{\mathbf{N}} \xrightarrow{p} B_0 \tag{16}$$

$$\sqrt{T}\left(\left(vec\left(\hat{B}_{\mathbf{N}}\right) - vec\left(B_0\right)\right) \xrightarrow{d} \mathcal{N}\left(0, V_{\mathbf{N}}\right), \tag{17}$$

where the formula for the asymptotic variance, $V_{\mathbf{N}}$, is standard but lengthy and, therefore, deferred to Appendix A.1. Moreover, under standard assumptions the weighting matrix $W^*_{\mathbf{N}} := S^{-1}_{\mathbf{N}}$ with $S_{\mathbf{N}} := \lim_{T\to\infty} E[g_{\mathbf{N}}(B)g_{\mathbf{N}}(B)']$ leads to the estimator $\hat{B}^*_{\mathbf{N}}$ with lowest possible asymptotic variance (see, e.g., Hall (2005)).

In many applications, the researcher is only interested in some structural shocks. For this case, we derive a partial identification result under weaker assumptions.

Proposition 2. *(Partial identification in the block-recursive SVAR.)*
Let $u_t = B_0\varepsilon_t$ with $m \leq n$ blocks and $B_0 \in \mathbb{B}_{brec} \equiv \mathbb{B}_{brec}(p_1, \dots, p_m)$ such that Assumption 1 holds. Moreover, let $B_{i,0}$ denote the columns of B_0 representing impact of the structural shocks in the ith block. Let $\widetilde{\mathbb{B}}_{brec} := \mathbb{B}_{brec}(\tilde{p}_1, \dots, \tilde{p}_{\tilde{m}})$ denote a potentially different block-recursive interaction. Assume that there exists a block $\tilde{p}_j$ of $\widetilde{\mathbb{B}}_{brec}$ which contains the shocks of block p_i, i.e., there exits a j, $1 \leq j \leq \tilde{m}$, such that $\tilde{p}_j = p_i$ and $\tilde{p}_{j+1} = p_{i+1}$.
The moment conditions $E\begin{bmatrix} f_{\mathbf{2}}(B, u_t) \\ f_{\mathbf{4}_{\tilde{\mathbf{p}}_\mathbf{j}}}(B, u_t) \end{bmatrix} = 0$ locally identify $B_{i,0}$ for $B \in \widetilde{\mathbb{B}}_{brec}$ if the following conditions hold:

1. *The shocks ε_t are uncorrelated.*
2. *The asymmetric cokurtosis conditions of block $\tilde{p}_j$ hold.*
3. *At most one shock in block $\tilde{p}_j$ has zero excess kurtosis.*

Proof. The proof can be found in Appendix A.3. □

Proposition 2 reveals that we can identify a specific block of shocks by using only the second moments of all shocks and the asymmetric cokurtosis conditions of the shocks in the block of interest as long as the block of interest is specified correctly and contains at most one Gaussian shock. To see the advantages of the partial identification result, consider that we are only interested in the last two structural shocks in Figure 1 (b). In this example, Proposition 2 implies that the impact of the last two shocks is identified even if (i) the first and second shock are both Gaussian, (ii) the first and second shock do not satisfy the asymmetric cokurtosis conditions but are only uncorrelated, or (iii) the block-recursive structure is misspecified as the one displayed in

Figure 1 (c). Additionally, Proposition 2 implies that the moment conditions used in Proposition 1 identify the shocks in a block of interest if the block of interest is specified correctly, contains at most one Gaussian shock, and there exists a B such that the moment conditions are fulfilled. However, the B matrix can differ from B_0, except for the columns corresponding to the block of interest.

4.2 Overidentification and efficiency gains

In the previous section, we proposed a block-recursive SVAR GMM estimator, which uses only a (small) subset of asymmetric cokurtosis conditions, and provide an identification result which does not require independent shocks. However, the excluded set of coskewness and cokurtosis conditions can decrease the asymptotic variance of the estimator and hence, increase the efficiency of the estimator. In this section, we define the overidentified block-recursive SVAR GMM estimator which contains all coskewness and cokurtosis conditions implied by independent shocks. Additionally, we derive conditions for the redundancy and relevance of the overidentifying coskewness and cokurtosis conditions in a recursive SVAR with independent structural shocks.

Assumption 3. *(Independent shocks.)*
All shocks are independent, i.e., $\varepsilon_{i,t}$ *is independent of* $\varepsilon_{j,t}$ *for* $i \neq j$.

For a given block-recursive SVAR, define the overidentifying moment conditions as

$$E[f_{\mathbf{D}}(B, u_t)] = E \begin{bmatrix} f_{\mathbf{3}\setminus\mathbf{N}}(B, u_t) \\ f_{\mathbf{4}\setminus\mathbf{N}}(B, u_t) \end{bmatrix}, \tag{18}$$

where $E[f_{\mathbf{3}\setminus\mathbf{N}}(B, u_t)]$ contains all coskewness conditions from Equation (8)-(9), and $E[f_{\mathbf{4}\setminus\mathbf{N}}(B, u_t)]$ contains all cokurtosis conditions from Equation (10)-(13), implied by independent shocks and not included in the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$.

The overidentified block-recursive SVAR GMM estimator is defined as

$$\hat{B}_{\mathbf{N}+\mathbf{D}} := \underset{B \in \mathbb{B}_{brec}}{\arg\min} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B) \end{bmatrix}' W_{\mathbf{N}+\mathbf{D}} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B) \end{bmatrix}, \tag{19}$$

with a suitable weighting matrix $W_{\mathbf{N+D}}$ and $g_{\mathbf{D}}(B) := 1/T \sum_{t=1}^{T} f_{\mathbf{D}}(B, u_t)$. Note that the overidentified block-recursive SVAR GMM estimator uses all coskewness and cokurtosis conditions implied by independent shocks. That is, the moment conditions used for estimation are the same as in the SVAR GMM estimator proposed by Keweloh (2021b). However, the latter estimator neither uses restrictions nor distinguishes between identifying and overidentifying moment conditions. In contrast to that, we allow for block-recursive restrictions. These restrictions allow to transform identifying into overidentifying moment conditions.

Consistency and asymptotic normality of the overidentified block-recursive SVAR GMM estimator in Equation (19) require that not only the identifying but also the overidentifying moment conditions are valid, which holds if the shocks are independent as assumed in Assumption 3. That is,

$$\hat{B}_{\mathbf{N+D}} \xrightarrow{p} B_0 \tag{20}$$

$$\sqrt{T}\left(\left(vec\left(\hat{B}_{\mathbf{N+D}}\right) - vec\left(B_0\right)\right) \xrightarrow{d} \mathcal{N}\left(0, V_{\mathbf{N+D}}\right), \tag{21}$$

where the formula for the asymptotic variance, $V_{\mathbf{N+D}}$, is standard and can be found in Appendix A.1. Again, under standard assumptions the weighting matrix $W^*_{\mathbf{N+D}} := S^{-1}_{\mathbf{N+D}}$ with $S_{\mathbf{N+D}} := \lim_{T\to\infty} E[g_{\mathbf{N+D}}(B_0) g_{\mathbf{N+D}}(B_0)']$, where $g_{\mathbf{N+D}}(B_0) := [g_{\mathbf{N}}(B_0)', g_{\mathbf{D}}(B_0)']'$, leads to the estimator $\hat{B}^*_{\mathbf{N+D}}$ with lowest possible asymptotic variance (see, e.g., Hall (2005)).

Adding additional valid moment conditions can never increase the asymptotic variance of the GMM estimator (see, e.g., Breusch et al. (1999)). Therefore, if the structural shocks are independent such that the overidentifying conditions hold, the asymptotic variance of $\hat{B}^*_{\mathbf{N+D}}$ is equal to or smaller than the asymptotic variance of $\hat{B}^*_{\mathbf{N}}$. If including an additional moment condition decreases the asymptotic variance of the estimator, the moment condition is called relevant, otherwise the moment condition is called redundant. A moment condition is called partially relevant for a subset of parameters if it decreases the asymptotic variance of a subset of parameters. If this is not the case, the moment condition is called partially redundant.

In the following proposition, we show that overidentifying higher-order moment conditions in $E[f_{\mathbf{D}}(B, u_t)]$ can decrease the asymptotic variance of the estimator. To this end, we consider the special case of a recursive SVAR with independent shocks. In this case, the SVAR is identified solely by second-order moment conditions and all coskewness and cokurtosis moment conditions are overidentifying. The proposition highlights that some coskewness and cokurtosis conditions are always (partially) redundant, while other conditions are relevant if certain conditions for the skewness, excess kurtosis, and elements of the inverse of B_0 are fulfilled. The proposition also implies that if at least one shock has a non-zero skewness, at least one higher-order moment condition will be relevant and consequently, the recursive SVAR GMM estimator based solely on second-order moment conditions, which is equal to frequently used estimator obtained by applying the Cholesky decomposition, is inefficient.

Proposition 3. *(Redundant and relevant moment conditions in the recursive SVAR.)*
Let $A := B_0^{-1}$ and let a_{ql} denote the element at row q and column l of A. Additionally let $i, j, k, l \in \{1, ..., n\}$ and $i \neq j \neq k \neq l$. The impact of a shock $\epsilon_{q,t}$ is equal to the unrestricted elements in the q-th row of B_0. In a recursive SVAR with independent structural shocks the following redundancy statements hold w.r.t. the identifying second-order moment conditions $E[f_{\mathbf{2}}(B, u_t)]$.

Coskewness condition:

1. $E[e(B_0)_i e(B_0)_j e(B_0)_k]$ *is redundant.*

2. $E[e(B_0)_i^2 e(B_0)_j]$ *is partially redundant for the impact of the shock $\epsilon_{q,t}$ with $q \neq j$.*

3. $E[e(B_0)_i^2 e(B_0)_j]$ *is partially redundant for the impact of the shock $\epsilon_{j,t}$ if and only if*

for $i < j$	*for $i > j$*
$\frac{2E[\epsilon_{j,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} = 0.$	$\frac{2E[\epsilon_{j,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} + E[\epsilon_{i,t}^3] a_{ij} = 0,$ $E[\epsilon_{i,t}^3] a_{i,z} = 0, \quad z = j+1, \ldots, i.$

Cokurtosis condition:

1. $E[e(B_0)_i e(B_0)_j e(B_0)_k e(B_0)_l]$ *and* $E[e(B_0)_i^2 e(B_0)_j e(B_0)_k]$ *are redundant.*

2. $E[e(B_0)_i^3 e(B_0)_j]$ *is partially redundant for the impact of the shock $\epsilon_{q,t}$ with $q \neq j$.*

3. $E[e(B_0)_i^3 e(B_0)_j]$ *is partially redundant for the impact of the shock* $\epsilon_{j,t}$ *if and only if*

for $i < j$	*for* $i > j$
$\frac{2E[\epsilon_{j,t}^3]E[\epsilon_{i,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} = 0.$	$\frac{2E[\epsilon_{j,t}^3]E[\epsilon_{i,t}^3]}{E[\epsilon_{j,t}^4]-1} a_{jj} + (E[\epsilon_{i,t}^4] - 3) a_{ij} = 0,$ $(E[\epsilon_{i,t}^4] - 3) a_{i,z} = 0, \quad z = j+1, \ldots, i.$

4. $E[e(B_0)_i^2 e(B_0)_j^2 - 1]$ *is partially redundant for the impact of the shock* $\epsilon_{q,t}$ *with* $q \neq i$ *and* $i < j$.

5. $E[e(B_0)_i^2 e(B_0)_j^2 - 1]$ *is partially redundant for the impact of the shock* $\epsilon_{i,t}$ *with* $i < j$ *if and only if*

$$E[\epsilon_{j,t}^3] E[\epsilon_{i,t}^3] a_{jz} = 0, \quad z = i, \ldots, j.$$

Proof. The proof can be found in Appendix A.4. □

In practice, the conditions in Proposition 3 cannot be verified since the matrix B_0, the skewness, and the kurtosis of the structural shocks are unknown a priori. Furthermore, Proposition 3 only covers a recursive SVAR with independent shocks, i.e., if the shocks are only mean independent or the SVAR has a different block-recursive structure, we do not have a theoretical result on which moment conditions are relevant and which are not.

4.3 Data-driven moment selection

Section 4.1 provides an identification result for block-recursive SVARs only requiring a (small) subset of cokurtosis conditions which is robust in the sense that it allows for various kinds of dependencies of the shocks. Section 4.2 stresses that there is a trade-off between robustness and efficiency of the estimator. For robustness, we leave out overidentifying conditions, which has the downside that some of these conditions may be valid and relevant, i.e., decrease the asymptotic variance of the estimator. However, an advantage is that one does not include potentially invalid overidentifying conditions, which could lead to an inconsistent overidentified block-recursive

SVAR GMM estimator in Equation (19). Additionally, valid but redundant overidentifying conditions can lead to a many moment problem and a poor finite sample performance of the overidentified block-recursive SVAR GMM estimator, compare Cheng and Liao (2015), Hall (2005), and Hall (2015). Therefore, we propose to use the pGMM estimator of Cheng and Liao (2015) to detect and include only the relevant and valid overidentifying moment conditions in a data-driven way. By including valid and relevant moment conditions in the estimation, we exploit the asymptotic efficiency gains of relevant moments. By leaving out invalid or redundant moment conditions, we can avoid inconsistent estimates and issues related to many moment conditions.

In general, the overidentifying higher-order moment conditions $E[f_{\mathbf{D}}(B,u_t)]$ can be separated into three sets: $E[f_{\mathbf{A}}(B,u_t)]$ contains valid and relevant moment conditions, $E[f_{\mathbf{R}}(B,u_t)]$ contains valid but redundant conditions, and $E[f_{\mathbf{I}}(B,u_t)]$ contains invalid moment conditions. The goal is to select the moments $E[f_{\mathbf{A}}(B,u_t)]$ and to leave out the moments $E[f_{\mathbf{R}}(B,u_t)]$ and $E[f_{\mathbf{I}}(B,u_t)]$. However, in practice the researcher does not know whether a given moment condition is invalid, redundant, or valid and relevant. Therefore, we propose to detect and select the relevant and valid overidentifying moment conditions in a data-driven way. Based on Cheng and Liao (2015), we define the block-recursive SVAR pGMM estimator

$$\{\widehat{B}^{pGMM}_{\mathbf{N+D}}, \widehat{\beta}\} := \underset{\{B,\beta\}\in\Lambda}{\arg\min} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B)-\beta \end{bmatrix}' W_{\mathbf{N+D}} \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B)-\beta \end{bmatrix} + \lambda \sum_{j\in\widetilde{D}} \omega_j \, |\beta_j|, \tag{22}$$

where $\lambda \geq 0$ is a tuning parameter specified by the researcher, $\beta \in \mathbb{R}^{k_{\mathbf{D}}}$ is the vector of slackness parameters, $\Lambda := \{\mathbb{B}_{brec}, \mathbb{R}^{1\times k_{\mathbf{D}}}\}$ is the parameter space of $\{B,\beta\}$, $\omega \in \mathbb{R}^{k_{\mathbf{D}}}$ is a vector of weights used in the penalty term, and $\widetilde{D} := \{1,\ldots,k_{\mathbf{D}}\}$ with $k_{\mathbf{D}}$ denoting the number of conditions in $E[f_{\mathbf{D}}(B,u_t)]$.

The vector of slackness parameters β allows the moment conditions $E[f_{\mathbf{D}}(B,u_t)]$ to deviate from zero without increasing the first part of the loss function and therefore, to decrease their impact on the estimation. However, each element of β gets penalized in the second part of the loss function and consequently, giving slack to overidentifying moments adds a cost, i.e., increases the loss function. The vector of weights ω and the tuning parameter λ govern the cost of giving slack to moment conditions. In particular, a smaller λ makes it cheaper to give slack to all overidentifying moments and a smaller ω_j makes it less costly to give slack to a specific overidentifying moment

j.

The pGMM estimator in Equation (22) has two special cases. First, if $\lambda = 0$, adding slack to the overidentifying moments is not penalized. Therefore, the solution of the pGMM estimator is $\widehat{B}^{pGMM}_{\mathbf{N+D}} = \hat{B}_{\mathbf{N}}$ and $\widehat{\beta} = g_{\mathbf{D}}\left(\hat{B}_{\mathbf{N}}\right)$, where $\hat{B}_{\mathbf{N}}$ is the solution of the the block-recursive SVAR GMM estimator in Equation (15) using only the identifying moments $E[f_{\mathbf{N}}(B, u_t)]$ and the weighting matrix $W_{\mathbf{N}}$, equal to the block of the weighting matrix $W_{\mathbf{N+D}}$ corresponding to the identifying conditions $E[f_{\mathbf{N}}(B, u_t)]$. Second, if $\lambda = \infty$, deviations of $\widehat{\beta}$ from zero become infinitely costly for overidentifying moments with $\omega_j > 0$. Assuming $\omega > 0$, the pGMM estimator cannot give slack to any overidentifying moment condition. Thus, $\widehat{B}^{pGMM}_{\mathbf{N+D}} = \hat{B}_{\mathbf{N+D}}$ and $\widehat{\beta} = 0$ minimize the loss function of the pGMM estimator, where $\hat{B}_{\mathbf{N+D}}$ is the solution of the the overidentified block-recursive SVAR GMM estimator in Equation (19), using the weighting matrix $W_{\mathbf{N+D}}$. Choices of λ other than $\lambda = 0$ or $\lambda = \infty$ lead to solutions which lie between these extreme cases. In practice, we recommend using cross-validation to find the optimal value of λ.

The penalty term uses weights $\omega_j \geq 0$, $\forall j \in \tilde{D}$, to shrink the elements of β differently. Let $E[f_{\mathbf{D}_j}(B, u_t)]$ for $j \in \widetilde{D}$ correspond to one specific moment of $E[f_{\mathbf{D}}(B, u_t)]$. A higher ω_j leads to more shrinkage for β_j and consequently, makes it more likely that β_j becomes zero, meaning that the corresponding moment $E[f_{\mathbf{D}_j}(B, u_t)]$ gets selected. Furthermore, $\omega_j = 0$ implies that even if the tuning parameter λ is large, there is no cost for giving slack to the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$, implying that those moments do not influence the estimated $\widehat{B}^{pGMM}_{\mathbf{N+D}}$. Since we aim to select only the relevant and valid moment conditions $E[f_{\mathbf{A}}(B, u_t)$, and not the invalid $E[f_{\mathbf{I}}(B, u_t)]$ or redundant moment conditions $E[f_{\mathbf{R}}(B, u_t)]$, we would specify $\omega_j > 0$ for all valid and relevant conditions, and $\omega_j = 0$ for all invalid or redundant conditions. To achieve this without prior knowledge on $E[f_{\mathbf{A}}(B, u_t)]$, $E[f_{\mathbf{R}}(B, u_t)]$, and $E[f_{\mathbf{I}}(B, u_t)]$, Cheng and Liao (2015) construct ω_j allowing information-based adaptive adjustment for each moment in $E[f_{\mathbf{D}}(B, u_t)]$. More precisely, they use

$$\omega_j = \frac{\mu_j^{r_1}}{|\beta_j^{*r_2}|},\ j \in \tilde{D}, \tag{23}$$

where μ_j is a measure for the empirical relevance of the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$, relative to the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$, and β_j^* is a preliminary consistent estimator of $E[f_{\mathbf{D}_j}(B_0, u_t)]$ and $r_1 \geq r_2 \geq 0$ are constants specified by the researcher. The use of $1/|\beta_j^{*r_2}|$ resembles an adaptive LASSO penalty (cf. Zou (2006)) and implies that moments with small β_j^* are subject to more shrinkage. Since β_j^* is a consistent estimator and the true value of β_j^* for a valid moment is zero, the adaptive penalty ensures that valid moments get selected. However, using only the adaptive penalty, we would unintendedly incentivize the estimator to select also redundant moments since, by definition, these are also valid. To avoid selecting redundant moments, Cheng and Liao (2015) suggest to multiply the adaptive penalty with

$$\mu_j = \rho_{\max}\left(\widehat{V}_{\mathbf{N}} - \widehat{V}_{\mathbf{N}+\mathbf{D}_j}\right),\ j \in \tilde{D}, \tag{24}$$

where $\rho_{\max}(A)$ is the maximum eigenvalue of a square matrix A and $\widehat{V}_{\mathbf{N}}$ and $\widehat{V}_{\mathbf{N}+\mathbf{D}_j}$ are consistent estimators of the efficient asymptotic variance-covariance matrices $V_{\mathbf{N}}^*$ and $V_{\mathbf{N}+\mathbf{D}_j}^*$, defined in Appendix A.1. If the maximum eigenvalue of $V_{\mathbf{N}}^* - V_{\mathbf{N}+\mathbf{D}_j}^*$ is positive, then adding moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ to the conditions $E[f_{\mathbf{N}}(B, u_t)]$ decreases the asymptotic variance of the estimator and hence, moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ is relevant. Therefore, μ_j estimates the empirical relevance of the moment $E[f_{\mathbf{D}_j}(B, u_t)]$.[7]

Cheng and Liao (2015) show that, under conditions, the pGMM estimator consistently selects the valid and relevant moments, i.e., $\lim_{T\to\infty} P(\hat{\beta}_j = 0) = 1$ if the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ is in $E[f_{\mathbf{A}}(B, u_t)]$, and does not select the invalid or redundant moments, i.e., $\lim_{T\to\infty} P(\hat{\beta}_j = 0) = 0$ if the moment condition $E[f_{\mathbf{D}_j}(B, u_t)]$ is in $E[f_{\mathbf{R}}(B, u_t)]$ or $E[f_{\mathbf{I}}(B, u_t)]$. They also derive that, under conditions, the pGMM estimator is a consistent estimator of B_0 and asymptotically normal with asymptotic variance $V_{\mathbf{N}+\mathbf{A}}$.[8] In our case, the conditions in particular require that Assumption 2 holds. However, consistency and asymptotic normality do not rely on independent shocks, i.e., Assumption 3. Even though the SVAR pGMM estimator uses the moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and $E[f_{\mathbf{D}}(B, u_t)]$ for estimation, its asymptotic variance only

[7] Cheng and Liao (2015) show that $V_{\mathbf{N}}^* - V_{\mathbf{N}+\mathbf{D}_j}^*$ is positive semidefinite for every $j \in \widetilde{D}$, implying that the maximum eigenvalue of $V_{\mathbf{N}}^* - V_{\mathbf{N}+\mathbf{D}_j}^*$ is nonnegative. Furthermore, note that both $\widehat{V}_{\mathbf{N}} \equiv \widehat{V}_{\mathbf{N}}\left(\hat{B}_{\mathbf{N}}\right)$ and $\widehat{V}_{\mathbf{N}+\mathbf{D}_j} \equiv \widehat{V}_{\mathbf{N}+\mathbf{D}_j}\left(\hat{B}_{\mathbf{N}}\right)$ are evaluated at $\hat{B}_{\mathbf{N}}$, which is obtained from Equation (15). Thereby, we do not rely on $\hat{B}_{\mathbf{N}+\mathbf{D}_j}$ to estimate $V_{\mathbf{N}+\mathbf{D}_j}^*$ since the moment associated with $\mathbf{D}_j$ may be invalid and hence, $\widehat{V}_{\mathbf{N}+\mathbf{D}_j}\left(\hat{B}_{\mathbf{N}+\mathbf{D}_j}\right)$ inconsistent for $V_{\mathbf{N}+\mathbf{D}_j}^*$.

[8] This result is not explicitly stated in Cheng and Liao (2015) but follows from their Remark 3.5 using the Cramér-Wold device, an arbitrary weighting matrix W and replacing the variance of the sample GMM estimator with the asymptotic variance. We prove the result in Appendix A.5 under Assumption 1 and 2.

depends on the moments conditions $E[f_{\mathbf{D}}(B, u_t)]$ and $E[f_{\mathbf{A}}(B, u_t)]$. That is, the SVAR pGMM estimator successfully ignores the redundant and invalid moments and decreases the asymptotic variance by incorporating the information contained in the relevant and valid moments. The weighting matrix $W^*_{\mathbf{N}+\mathbf{D}} := S^{-1}_{\mathbf{N}+\mathbf{D}}$ leads to the estimator with the lowest possible asymptotic variance (Hall, 2005), corresponding to the asymptotic variance of the oracle estimator. The oracle estimator uses only moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and $E[f_{\mathbf{A}}(B, u_t)]$ and is infeasible in practice without prior knowledge of $E[f_{\mathbf{D}}(B, u_t)]$ and $E[f_{\mathbf{A}}(B, u_t)]$. However, the SVAR pGMM estimator is as efficient as the oracle estimator asymptotically.

5 Finite sample performance

In this section, we conduct two Monte Carlo studies. The first one illustrates that the performance of SVAR estimators can be improved substantially by exploiting the block-recursive structure. This is especially relevant for SVARs with well-justified restrictions and a large number of variables. The second Monte Carlo study focuses on how to incorporate information in overidentifying higher-order moment conditions. More concretely, we demonstrate that the pGMM estimator selects relevant and does not select redundant moment conditions in a data-driven way and thereby, improves the finite sample performance.

For both Monte Carlo experiments, we consider three different sample sizes $T = \{100, 250, 1\,000\}$ to analyze the influence of the sample size on the performance of the estimators. We independently and identically draw each structural shock ϵ_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$, from the two-component mixture

$$\epsilon_{it} \sim 0.79\, \mathcal{N}(-0.2, 0.7^2) + 0.21\, \mathcal{N}(0.75, 1.5^2),$$

where $\mathcal{N}(\mu, \sigma^2)$ indicates a normal distribution with mean μ and standard deviation σ. The shocks have skewness 0.9 and excess kurtosis 2.4.

We compare the finite sample performance of various SVAR estimators.[9] Based on the simulations presented in Keweloh (2021a), we use continuous updating estimators (CUEs) instead of GMM estimators and estimate the asymptotically efficient weighting matrix based on serially and

[9] The estimators are implemented in python and the pGMM estimator uses the solvers of Defferrard et al. (2017).

mutually independent shocks.[10] Since CUE estimators are closely related to GMM estimators, we use both terms interchangeably. More specifically, we refer to the estimators as follows:

- GMM: Continuous updating estimator based on Equation (15) using only the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$.
- oGMM: Overidentified continuous updating estimator based on Equation (19) using the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and overidentifying moment conditions $E[f_{\mathbf{D}}(B, u_t)]$.
- GMM-Oracle: Overidentified continuous updating estimator based on Equation (19) using the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ and the relevant overidentifying moment conditions $E[f_{\mathbf{A}}(B, u_t)]$.
- pGMM: Continuous updating LASSO estimator based on Equation (22).

We only indicate which block-recursive structure is imposed for estimation, when necessary (e.g., when comparing an GMM estimator without restrictions with a block-recursive GMM estimator).

5.1 Block-Recursive Structure

We simulate a SVAR with $n = 2$ and $n = 4$ variables. The mixing matrices B_0 are given by

$$B_0 = \begin{bmatrix} 10 & 5 \\ 5 & 10 \end{bmatrix} \quad \text{and} \quad B_0 = \begin{bmatrix} 10 & 5 & 0 & 0 \\ 5 & 10 & 0 & 0 \\ 5 & 5 & 10 & 5 \\ 5 & 5 & 5 & 10 \end{bmatrix}. \tag{25}$$

The Monte Carlo study analyzes the impact of imposing a block-recursive structure for GMM estimators. In the small SVAR with $n = 2$, we impose no restrictions. In the large SVAR

[10] Keweloh (2021a) demonstrates that the inability to precisely estimate S, the long-run covariance matrix of the moment conditions, and as consequence the efficient weighting matrix leads to a poor small sample performance of two-step GMM and CUE estimators. Recognizing this downside, Keweloh (2021a) proposes a novel estimator for S exploiting serially and mutually independent shocks.Keweloh (2021a) illustrates that the estimator for S substantially increases the small sample performance of the two-step GMM and CUE estimator. Additionally, Keweloh (2021a) illustrates that CUE estimators are less biased than GMM estimator in small samples.

with $n = 4$, we estimate the GMM estimator without restrictions and the block-recursive GMM estimator, using the block-recursive structure in Equation (25), i.e., we apply zero restrictions for all elements where B_0 is zero.[11]

Table 1 summarizes the results of $M = 3,500$ Monte Carlo simulations. The table shows the average of each estimated element $\bar{b}_{ij} = 1/M \sum_{m=1}^{M} \hat{b}_{ij}^m$ and the estimated mean squared error (MSE), $\hat{\sigma}_{i,j}^2 = 1/M \sum_{m=1}^{M} \left(\hat{b}_{ij}^m - b_{ij}\right)^2$, where b_{ij} denotes the element of B_0 in row i and column j and $\hat{b}_{ij}^m$ its estimated value in Monte Carlo run m. Moreover, we calculate the average over the empirical biases, $Bias := \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \left(\bar{b}_{ij} - b_{ij}\right)$, and the average over the estimated MSEs, $Var := \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \hat{\sigma}_{i,j}^2$, across estimated elements in $\hat{B}$, i.e., $w_{i,j}$ equals zero if $\hat{b}_{ij}^m$ is restricted to be zero and one over the number of estimated elements in $\hat{B}$ otherwise. Additionally, we report the number of moments used by each estimator. For each estimator, the average bias and MSE decreases with the sample size. Furthermore, the simulation highlights how the performance of the GMM estimators, which are based entirely on non-Gaussianity, decreases with an increasing model size (e.g., the average bias and MSE for each sample size is up to 2.1 and 1.9 times higher for the GMM estimator with $n = 4$ compared to the GMM estimator with $n = 2$). The Monte Carlo study illustrates how in a typical macroeconomic application, which rarely or if at all contains more than a few hundred observations, data-driven estimates based on non-Gaussianity become less reliable the more variables the SVAR contains. However, the simulation also stresses that exploiting the block-recursive structure annihilates the deterioration of the performance induced by a larger model. That is, the average bias and MSE for each sample size in Table 1 is at least 1.8 and 1.8 times higher for the GMM estimator with $n = 4$ compared to the block-recursive GMM estimator with $n = 4$. Using the block-recursive structure allows the block-recursive GMM estimator to estimate the four elements on the lower left of B_0 (each with a value of 5) only by covariance moment conditions (which explains why the average MSE of the block-recursive GMM estimator with $n = 4$ even can be lower than or comparable to the GMM estimator with $n = 2$, which relies on higher-order moment conditions).

Our results suggest that if in a given application well-justified restrictions are available, these restrictions should be used as they substantially improve the performance of the estimator.

[11] In this Monte Carlo study, we focus on GMM estimators. We include the oGMM, GMM-Oracle and pGMM estimator in the second Monte Carlo study.

Table 1: Finite sample performance of the GMM and block-recursive GMM estimator.

		n=2		n=4							
		GMM		GMM				block-recursive GMM			
$T = 100$	$\hat{B}$	9.78 (2.26)	4.90 (4.31)	9.28 (3.24)	4.63 (4.82)	0.04 (5.31)	0.07 (5.27)	9.74 (2.31)	4.91 (4.30)	.	.
		4.90 (4.24)	9.76 (2.18)	4.70 (4.87)	9.23 (3.20)	0.08 (5.32)	0.05 (5.14)	4.87 (4.43)	9.74 (2.18)	.	.
				4.68 (6.54)	4.62 (6.74)	9.27 (5.01)	4.74 (6.54)	4.86 (2.51)	4.89 (2.44)	9.63 (2.17)	4.84 (4.41)
				4.67 (6.67)	4.65 (6.53)	4.66 (6.48)	9.33 (4.93)	4.87 (2.56)	4.91 (2.45)	4.84 (4.24)	9.64 (2.34)
	#Mo	5.00		22.00				14.00			
	Bias	−0.1649		−0.3314				−0.1878			
	MSE	3.25		5.41				3.03			
		n=2		n=4							
		GMM		GMM				block-recursive GMM			
$T = 250$	$\hat{B}$	9.88 (1.10)	4.90 (2.30)	9.56 (1.64)	4.79 (2.77)	0.02 (3.19)	0.06 (3.21)	9.87 (1.07)	4.91 (2.41)	.	.
		4.98 (2.22)	9.85 (1.13)	4.77 (2.69)	9.54 (1.65)	−0.01 (3.14)	0.04 (3.26)	4.94 (2.33)	9.83 (1.15)	.	.
				4.74 (4.05)	4.83 (3.94)	9.56 (2.76)	4.83 (3.92)	4.93 (1.16)	4.91 (1.20)	9.81 (1.13)	4.92 (2.30)
				4.74 (4.10)	4.82 (3.91)	4.79 (3.86)	9.61 (2.85)	4.94 (1.14)	4.92 (1.21)	4.91 (2.32)	9.84 (1.09)
	#Mo	5.00		22.00				14.00			
	Bias	−0.0982		−0.2065				−0.1069			
	MSE	1.69		3.18				1.54			
		n=2		n=4							
		GMM		GMM				block-recursive GMM			
$T = 1000$	$\hat{B}$	9.96 (0.24)	5.00 (0.46)	9.92 (0.26)	4.99 (0.53)	0.00 (0.64)	0.02 (0.54)	9.97 (0.22)	5.02 (0.48)	.	.
		4.97 (0.48)	9.97 (0.22)	4.95 (0.51)	9.94 (0.29)	0.00 (0.61)	0.02 (0.53)	4.97 (0.46)	9.99 (0.24)	.	.
				4.95 (0.73)	4.99 (0.72)	9.92 (0.56)	4.99 (0.65)	4.98 (0.25)	5.01 (0.28)	9.96 (0.21)	4.99 (0.40)
				4.95 (0.69)	4.99 (0.66)	4.96 (0.75)	9.95 (0.43)	4.98 (0.25)	5.01 (0.27)	4.98 (0.41)	9.97 (0.20)
	#Mo	5.00		22.00				14.00			
	Bias	−0.0262		−0.0295				−0.0124			
	MSE	0.35		0.57				0.31			

The table reports the average $\bar{b}_{ij}$ and the corresponding estimated MSE (in parentheses) of each estimated element in $\hat{B}$ as well as the BIAS and MSE across estimated elements in $\hat{B}$ over $3,500$ Monte Carlo replicates. We estimate the GMM estimator without restrictions for $n = 2$ and $n = 4$, and the block-recursive GMM estimator for $n = 4$, which uses zero restrictions highlighted by the dots.

5.2 Recursive Structure

In this subsection, we simulate a recursive SVAR using $n = 4$ variables and

$$B_0 = \begin{pmatrix} 10 & 0 & 0 & 0 \\ 5 & 10 & 0 & 0 \\ 5 & 5 & 10 & 0 \\ 5 & 5 & 5 & 10 \end{pmatrix}. \tag{26}$$

For the estimation of B_0, we impose a recursive order for all considered estimators, i.e., we use zero restrictions for all elements where B_0 is zero. In this setup, the pGMM, GMM-Oracle, and the oGMM estimator are efficient estimators and have a smaller asymptotic variance than the GMM estimator, which is equivalent to the estimator obtained by applying a Cholesky decomposition. By using a recursive structure, we can apply Proposition 3 to calculate whether an overidentifying moment condition is relevant or redundant. Therefore, we can analyze whether the pGMM estimator selects relevant moment conditions and does not select redundant moment conditions. With the imposed recursive order, the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ contain 10 and the overidentifying conditions $E[f_{\mathbf{D}}(B, u_t)]$ contain 47 conditions. All moment conditions in $E[f_{\mathbf{D}}(B, u_t)]$ are valid. More precisely, 17 of overidentifying conditions are redundant and 30 overidentifying conditions are relevant.

The construction of the weights for the pGMM estimator as in Equation (23) requires an initial consistent estimate $\hat{B}$ to estimate β^* and the asymptotic variance in Equation (24). To this end, we apply the GMM estimator, which is the Cholesky estimator in this case. Moreover, we again use the assumption of independent shocks to estimate the asymptotic variance, as proposed by Keweloh (2021a). We use $r_1 = 2$ and $r_2 = 1$ in Equation (23) and additionally, we normalize the weights such that they sum to one, i.e., we use $\omega_j^* := \omega_j / \sum_{k \in \widetilde{D}} \omega_k$, allowing for straightforward comparison among the weights.

We choose the optimal λ for the pGMM estimator with 5-fold cross-validation from a sequence of 10 potential values. The maximum value of the sequence of λ's depends on the sample size, ensuring that it is large enough to select all moments j for which $\omega_j^* > 10^{-4}$.[12] We also include $\lambda = 0$ in the range of possible values to allow our estimator to simplify to the recursive SVAR. The selection of the optimal tuning parameter is based on the median of the GMM loss of each left-out fold.

Table 2 summarizes the results of $M = 3,500$ Monte Carlo simulations. We report the same summary statistics as in Table 1. In addition, we calculate the average number of moments selected by the pGMM estimator and the median of the chosen λ's for the pGMM estimator across Monte Carlo runs. In Appendix B.1, we display results including the Post-pGMM estimator which uses the moments selected by pGMM in a second stage estimation.

The GMM estimator performs well in the smallest sample size in terms of bias and MSE. However, the GMM estimator is asymptotically inefficient and has the largest MSE among all considered estimators for $T = 250$ and $T = 1000$. Due to many moments, the oGMM estimator performs worst in terms of bias and MSE among the considered estimators for $T = 100$. Yet, its performance improves with sample size and it eventually outperforms the GMM estimator in terms of MSE. The bias is highest for the oGMM and GMM-Oracle estimator across sample sizes, which might be explained by the greater number of moments used by these estimators. Note that both estimators are asymptotically efficient. Nevertheless, many moment conditions can still lead to a finite sample bias. The MSE of the GMM-Oracle estimator is already comparable to the GMM estimator in small samples. Relative to the other estimators, its MSE further decreases with the sample size and it performs best in the largest sample size. In general, the GMM-Oracle estimator is infeasible since the redundant moments are unknown a priori.[13] In contrast to that, the pGMM estimator is feasible and uses a data-driven approach to select the relevant and valid moments. The pGMM estimator performs well across all sample sizes in terms of bias and MSE. For $T = 100$, its bias and MSE is notably smaller than the one of the oGMM and the GMM-Oracle estimator and surprisingly, also smaller than the one of the GMM estimator. In the largest sample, the pGMM estimator performs similar to the oGMM and GMM-Oracle estimator in terms of

[12] We specify the maximum value of the sequence of λ's in a data-driven way using the subgradient of Equation (22) with respect to β. We give more details on how to construct the maximum value of the sequence of λ's in the cross-validation in Appendix A.6.

[13] Even if we knew the non-Gaussianity of the shocks, we would not be able to derive the oracle estimator if the block-recursive structure was not just purely recursive. In this case, we still lack the information on which moments are redundant and which are relevant.

Table 2: Finite Sample Performance of the pGMM estimator.

		GMM	oGMM	GMM-Oracle	pGMM
$T = 100$	$\hat{B}$	9.93 (1.09) . . .	9.77 (1.07) . . .	9.76 (1.07) . . .	9.96 (1.09) . . .
		4.98 (1.21) 9.86 (1.02) . .	4.90 (1.31) 9.71 (1.01) . .	4.91 (1.17) 9.70 (1.02) . .	5.00 (1.15) 9.88 (1.01) . .
		4.97 (1.49) 4.95 (1.29) 9.83 (1.12) .	4.89 (1.69) 4.88 (1.43) 9.70 (1.10) .	4.91 (1.50) 4.88 (1.26) 9.69 (1.10) .	4.98 (1.46) 4.96 (1.22) 9.85 (1.11) .
		4.96 (1.71) 4.93 (1.46) 4.91 (1.27) 9.78 (1.08)	4.90 (2.07) 4.88 (1.74) 4.88 (1.46) 9.69 (1.09)	4.92 (1.81) 4.88 (1.51) 4.88 (1.25) 9.67 (1.10)	4.99 (1.71) 4.96 (1.42) 4.95 (1.21) 9.82 (1.10)
	#Mo	10.00	57.00	40.00	24.22
	Bias	−0.0883	−0.1806	−0.1804	−0.0650
	MSE	1.27	1.40	1.28	1.25
	λ	.	.	.	71.08

		GMM	oGMM	GMM-Oracle	pGMM
$T = 250$	$\hat{B}$	9.97 (0.43) . . .	9.90 (0.40) . . .	9.90 (0.40) . . .	9.99 (0.42) . . .
		4.99 (0.51) 9.96 (0.43) . .	4.96 (0.49) 9.90 (0.40) . .	4.97 (0.44) 9.90 (0.40) . .	5.01 (0.45) 9.97 (0.41) . .
		4.98 (0.64) 5.00 (0.52) 9.93 (0.45) .	4.96 (0.65) 4.97 (0.51) 9.87 (0.42) .	4.97 (0.59) 4.97 (0.46) 9.87 (0.42) .	5.01 (0.59) 5.02 (0.46) 9.94 (0.42) .
		4.98 (0.72) 4.99 (0.61) 4.98 (0.51) 9.91 (0.45)	4.97 (0.73) 4.96 (0.61) 4.97 (0.49) 9.86 (0.42)	4.98 (0.65) 4.97 (0.54) 4.96 (0.44) 9.85 (0.42)	5.02 (0.66) 5.01 (0.55) 5.00 (0.44) 9.92 (0.43)
	#Mo	10.00	57.00	40.00	27.20
	Bias	−0.0311	−0.0676	−0.0656	−0.0114
	MSE	0.53	0.51	0.48	0.48
	λ	.	.	.	118.92

		GMM	oGMM	GMM-Oracle	pGMM
$T = 1000$	$\hat{B}$	10.00 (0.11) . . .	9.98 (0.10) . . .	9.98 (0.10) . . .	10.00 (0.11) . . .
		5.00 (0.13) 9.99 (0.11) . .	4.99 (0.12) 9.97 (0.10) . .	4.99 (0.11) 9.97 (0.10) . .	5.00 (0.11) 9.99 (0.10) . .
		4.99 (0.15) 4.99 (0.13) 9.99 (0.11) .	4.99 (0.13) 4.99 (0.11) 9.98 (0.10) .	4.99 (0.13) 4.99 (0.10) 9.98 (0.10) .	5.00 (0.13) 5.00 (0.11) 10.00 (0.10) .
		4.99 (0.19) 4.99 (0.15) 4.99 (0.13) 9.98 (0.11)	4.99 (0.16) 4.99 (0.14) 4.99 (0.11) 9.97 (0.10)	4.99 (0.15) 4.99 (0.13) 4.99 (0.11) 9.97 (0.10)	5.00 (0.16) 5.00 (0.13) 5.00 (0.11) 9.98 (0.10)
	#Mo	10.00	57.00	40.00	29.59
	Bias	−0.0076	−0.0158	−0.0158	−0.0021
	MSE	0.13	0.12	0.11	0.12
	λ	.	.	.	75.34

The table reports the average $\bar{b}_{ij}$ and the corresponding estimated MSE (in parentheses) of each estimated element in $\hat{B}$ as well as the BIAS and MSE across estimated elements in $\hat{B}$ over 3,500 Monte Carlo replicates for the GMM estimator, the oGMM estimator, the GMM-Oracle estimator, and the pGMM estimator. All estimator use zero restrictions which are highlighted by the dots.

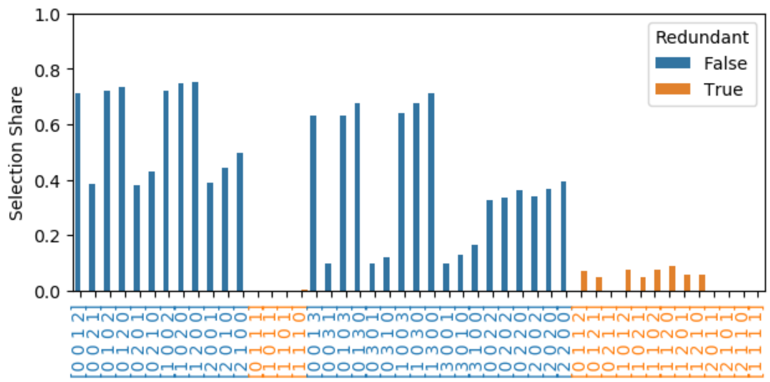
MSE and best in terms of bias.[14] The simulation shows that the pGMM estimator can, without prior specification, distinguish informative from non-informative overidentifying moments, which solves the many moments problem of the oGMM estimator and allows to exploit information in overidentifying higher-order moments already in small samples.

Table 2 indicates that the average number of selected moments increases only slightly as T

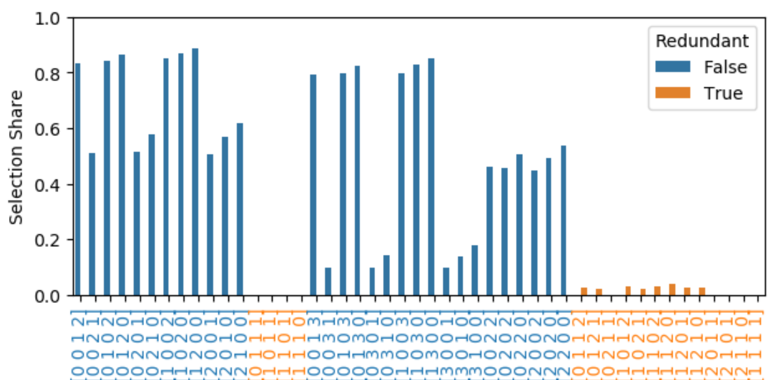
[14] The Post-pGMM estimator reported in Appendix B.1 performs similar to the pGMM estimator.

increases. Even for $T = 1000$, the pGMM estimator only selects 20 out of 30 valid and relevant overidentifying moments in addition to the 10 identifying moments. That said, the remaining 10 moments would only decrease the MSE from 0.12 to 0.11, indicating that the moments not being selected would not lower the MSE much. Figure 2 illustrates that pGMM estimator only selects relevant moments and manages to leave out redundant moments, especially as T increases. Moreover, the share of selections of each moment across all Monte Carlo runs rises with the sample size for the majority of relevant moments. In Figure B.2, we plot the average weight of each moment across Monte Carlo runs. By comparing Figure 2 and Figure B.2, we argue that there is a clear correlation between the average weight and the number of selections of each moment. More precisely, all redundant moments have an average weight which is very close to zero and hence, they are not selected by the pGMM estimator.

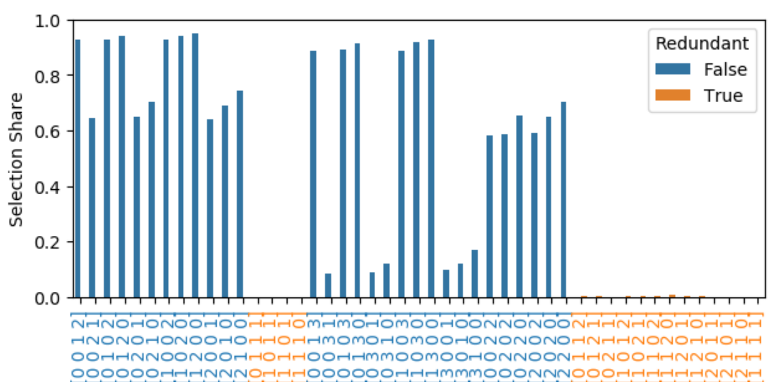
Figure 2: Share of Selections of Moments across Monte Carlo Runs



(a) $T = 100$



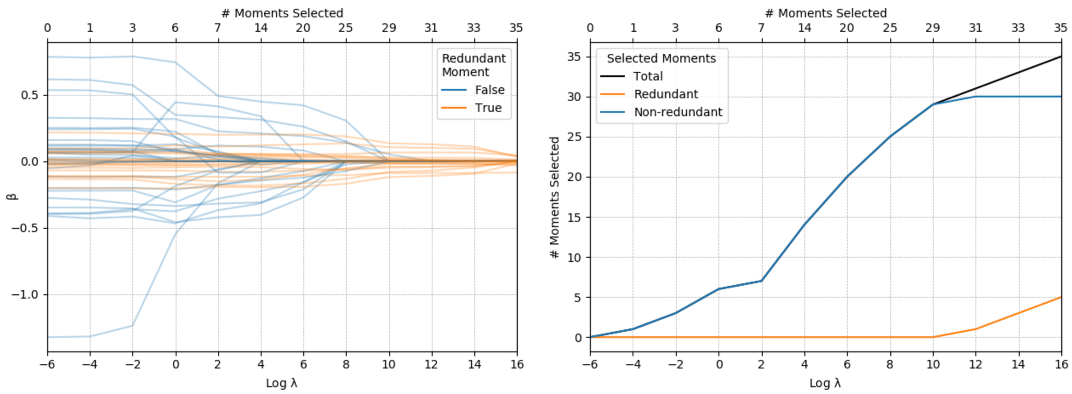
(b) $T = 250$



(c) $T = 1000$

Note: The figure shows how often each moment gets selected across $M = 3,500$ Monte Carlos simulations. Redundant moment (orange) and relevant moments (blue) are displayed on the x-axis. Each x-axis label abbreviates a moment condition, e.g., $[0, 1, 2, 1]$ corresponds to $E[e(B)^0_{1,t}\, e(B)^1_{2,t}\, e(B)^2_{3,t}\, e(B)^1_{4,t}]$.

Figure 3: Illustration of Influence of λ on β.



(a) Trace Plot (b) Selected Moments

Note: Panel (a) of the figure shows the values of β in dependence on $log(\lambda)$ for one Monte Carlo run for $T = 100$ and the corresponding number of selected moments in $\widetilde{D}$. Panel (b) of the figure splits the number of selected moments into the number of selected redundant and the number of selected relevant moments for each $log(\lambda)$.

Figure 3 highlights the influence of λ on β and hence, on the number of selected moment conditions for one Monte Carlo run.[15] For instance, for $\log(\lambda) = -6$ no overidentifying moment conditions are selected and the solution of the pGMM estimator corresponds to the one of the GMM estimator. Further, the number of selected moments increases as λ increases, i.e., the penalty shrinks the elements of β to zero. Furthermore, the relevant moments get selected first when λ increases and we do not select any redundant moment until λ becomes very large.

6 Application of the block-recursive SVAR: Disentangling speculative demand and supply shocks in the oil market

In this section, we propose a SVAR model for the oil market to analyze the impact of flow and speculative supply and of flow and speculative demand shocks on the real oil price. A flow supply shock for oil represents an exogenous deviation in the present amount of oil coming out of the ground and a flow demand shock for oil an exogenous deviation in the present amount of oil being consumed. A speculative oil supply shock represents a shift in the expected future oil supply and a speculative oil demand shock a shift in the expected future oil demand.

[15] For the purpose of illustration, we use a wider range of of λ values for this plot.

We consider a SVAR with monthly data from January 1974 to December 2019 of the form

$$\begin{bmatrix} O_t \\ Y_t \\ OP_t \\ SR_t \end{bmatrix} = \alpha + \sum_{i=1}^{12} A_i \begin{bmatrix} O_{t-i} \\ Y_{t-i} \\ OP_{t-i} \\ SR_{t-i} \end{bmatrix} + \begin{bmatrix} u_t^O \\ u_t^Y \\ u_t^{OP} \\ u_t^{SR} \end{bmatrix}. \tag{27}$$

The variable O_t is the log difference of global oil production, Y_t is the log difference of industrial production, measuring economic activity, OP_t is the growth rate of real oil price, and SR_t are real monthly stock returns.[16] We decompose the reduced form shocks u_t into four structural shocks with

$$\begin{bmatrix} u_t^O \\ u_t^Y \\ u_t^{OP} \\ u_t^{SR} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{41} \end{bmatrix} \begin{bmatrix} \varepsilon_t^s \\ \varepsilon_t^d \\ \varepsilon_t^{s-exp} \\ \varepsilon_t^{d-exp} \end{bmatrix}, \tag{28}$$

where ε_t^s is a flow supply shock for oil, ε_t^d is a flow demand shock for oil, ε_t^{s-exp} is a speculative oil supply shock, and ε_t^{d-exp} is a speculative oil demand shock. The block-recursive restrictions in Equation (28) imply that oil production and economic activity behave sluggishly and can contemporaneously only respond to flow supply and demand shocks, whereas oil prices and stock returns can immediately incorporate all available information and contemporaneously respond to flow and speculative supply and demand shocks.

The simultaneous relationship is estimated using the block-recursive SVAR pGMM estimator.[17] In line with the Monte Carlo simulations, we apply continuous updating for the weighting

[16] Global oil production is given by the global crude oil including lease condensate production obtained from the U.S. EIA. We obtain industrial production by the monthly industrial production index in the OECD and six major other countries from Baumeister and Hamilton (2019). The real oil price is equal to the refiner's acquisition cost of imported crude oil from the U.S. EIA deflated by the U.S. CPI. Real stock prices correspond to the aggregate U.S. stock index constructed by the OECD deflated by the U.S. CPI.

[17] In Appendix B.2, we conduct various robustness checks. In particular, we estimate the block-recursive SVAR using the GMM estimator from Equation (15) and the overidentified GMM estimator from Equation (19). Estimates using the white fast SVAR GMM estimator proposed by Keweloh (2021b) and the PML estimator proposed by Gouriéroux et al. (2017) are qualitatively similar and available on request. Additionally, we report results for different specifications of the variables in the block-recursive SVAR, including specifications where oil production and real economic activity are included in deviations from a trend, and where the oil price is measured in log levels.

matrix and use the assumption of serially and mutually independent shocks to estimate the asymptotically efficient weighting matrix as proposed by Keweloh (2021a). With the imposed block-recursive structure, we can divide the moment conditions into 14 identifying conditions $E[f_{\mathbf{N}}(B, u_t)]$ and 43 overidentifying conditions $E[f_{\mathbf{D}}(B, u_t)]$. We use the same specifications to construct the weights as in the Monte Carlo simulation, i.e., we use $r_1 = 2$ and $r_2 = 1$ in Equation (23). For the cross-validation, we consider a range of 28 values for λ, including $\lambda = 0$. The maximum value of λ is chosen such that all conditions $E[f_{\mathbf{D}}(B, u_t)]$ for which $\omega_j / \sum_{k \in \widetilde{D}} \omega_k > 10^{-7}$ get selected. With the chosen $\lambda = 34679$, which is the 27th value of the considered sequence, 12 coskewness and 12 cokurtosis conditions are selected.[18]

For each estimated structural shock, Table 3 shows the estimated skewness, kurtosis and p-value of the Jarque-Bera test. To ensure identification, at most one structural shock in each block may be Gaussian. In our block-recursive structure, each block contains only two shocks and, therefore, it is sufficient for identification to show that at least one structural shock in each block is non-Gaussian. Furthermore, the block-recursive structure implies that each of the two unmixed innovations in the first block is equal to a linear combination of the two structural shocks in the first block, i.e., if both structural shocks are Gaussian, the two unmixed innovations have to be Gaussian as well. However, the skewness, kurtosis, and Jarque-Bera test for normality clearly suggest that the unmixed innovations in the first block are non-Gaussian and, hence, that at least one structural shock in the first block is non-Gaussian. Consequently, the first block is identified. Moreover, the unmixed innovations in the second block are equal to a linear combination of the structural shocks in the second block (the argument follows from Equation (A.5) in the proof of Proposition 2). Again, the skewness, kurtosis and Jarque-Bera test for normality clearly suggest that the unmixed innovations in the second block are non-Gaussian, implying that at least one structural shock in the second block is non-Gaussian. Thus, the second block is also identified. Consequently, the block-recursive SVAR is identified.

In Figure 4, we show impulse response functions (IRFs).

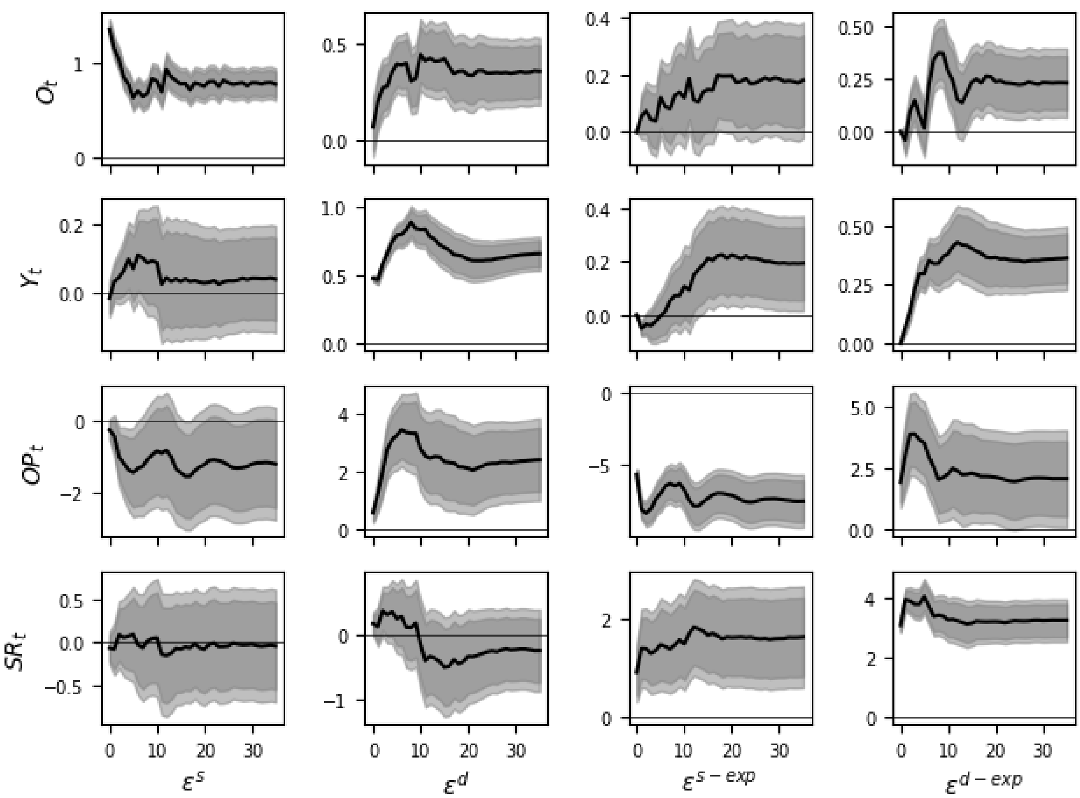
[18] Additionally, we compute the block-recursive SVAR pGMM estimator using the plugin rule $\lambda = k_{\mathbf{D}}^{r_2/4} T^{(-0.5 - r_2/4)}$, where $k_{\mathbf{D}}$ denotes the number of overidentifying moment conditions, see Cheng and Liao (2015). The estimator selects 8 coskewness and 6 cokurtosis conditions.

Table 3: Non-Gaussianity of the estimated structural shocks

	ε_t^s	ε_t^d	ε_t^{s-exp}	ε_t^{d-ext}
Skewness	−0.97	−0.21	0.46	−0.82
Kurtosis	9.92	4.58	6.79	6.88
JB-Test	0.00	0.00	0.00	0.00

Note: Skewness, kurtosis and the p-value of the Jarque-Bera test for normality.

Figure 4: Impulse Responses of the block-recursive SVAR pGMM estimator.



Note: Impulse responses to the estimated structural shocks for the block-recursive SVAR pGMM estimator. Confidence bands are symmetric 68% and 80% bands based on standard errors and 1000 replications. The rows show the cumulative responses. The x-axis displays monthly lags.

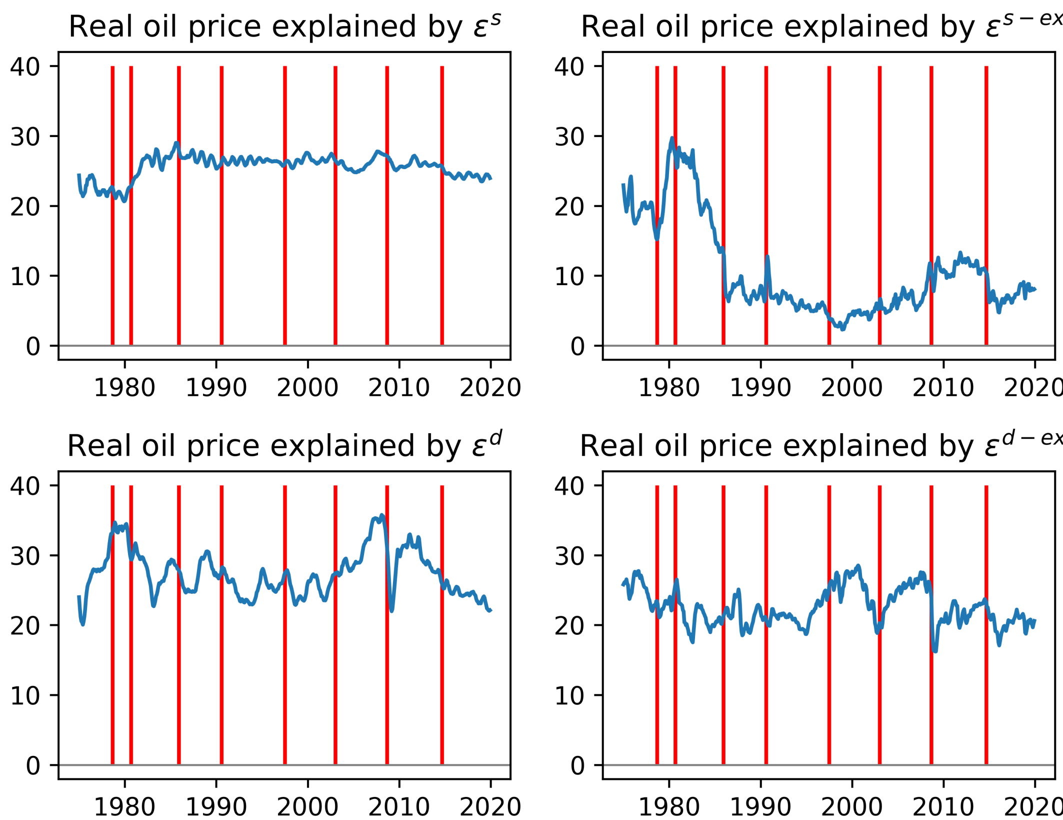
With the block-recursive structure, labeling of the shocks in the plot of the IRFs is straightforward. In the first block, there is only one shock which leads to a significant immediate increase of economic activity and, thus, an immediate increase in demand for oil. We label this shock as the flow demand shock and the remaining shock in the first block as the flow supply shock.

In the second block, one shock leads to an immediate increase of the real oil price and to a long-run increase of economic activity. We label this shock as the speculative oil demand shock. The remaining shock in the second block leads to an immediate decrease of the oil price and to an increase of economic activity and oil production in the long-run, which corresponds to the speculative oil supply shock.

Our results show that flow supply shocks immediately increase oil production and decrease the real oil price and flow demand shocks increase real economic activity and the real oil price. Moreover, oil production responds to the demand shock with a lagged increase. Interestingly, it seems that real stock returns do not respond significantly to flow demand and supply shocks. With respect to the speculative shocks, we find that a supply expectation shock leads to an increase of oil production and of real economic activity after one year. Furthermore, it immediately and permanently decreases the real oil price and increases real stock returns. A speculative demand shock increases oil production and real economic activity. Additionally, the speculative demand shocks leads to an immediate increase of the real oil price and of real stock returns.

Figure 5 shows the contribution of the estimated structural shocks to the evolution of the real oil price. Figure B.3 in Appendix B.2 shows the historical evolution of the real oil price.

Figure 5: Real oil price evolution explained by the estimated structural shocks.



Note: In each of the panels, we simulate the real oil price (blue line) by setting all but one of the shocks to zero (and for ease of interpretation, we also set $\alpha = 0$ in Equation (27)). The red vertical bars indicate the following events: Iranian Revolution (1978 : 9), Iran Iraq War (1980 : 9), collapse of OPEC (1985 : 12), Persian Gulf War (1990 : 8), Asian Financial Crisis of (1997 : 7), Iraq War (2003 : 1), the collapse of Lehman Brothers (2008 : 9), and the oil price decline in mid 2014.

Figure 5 suggests that the increase of the real oil price from 1978 to 1981 is mainly driven by flow supply and speculative supply shocks. Moreover, we find that the decline of the real oil price from 1981 to 1985 is largely explained by speculative supply shocks. Additionally, the decrease in real oil prices after the collapse of OPEC in 1985 and the peak of real oil prices during the Persian Gulf War in 1990 can to a large extent be explained explained by speculative supply shocks. The run-up in the real oil prices from 2003 to 2008 is driven by flow demand, speculative demand, and speculative supply shocks. Flow demand and speculative demand shocks explain the plunge of the real oil price during the financial crisis in 2008. Additionally, most of the recovery of the real oil price after the financial crisis is explained by demand shocks. The collapse of the real oil price since mid 2014 is related to flow demand, speculative demand, and speculative supply shocks.

The IRFs in Figure 4 show no evidence against a recursive structure of the shocks in the first block. That said, our results clearly suggest that the second block does not have a recursive structure since the two structural shocks in the second block have an immediate impact on both reduced form shocks in the second block. As a robustness-check and to illustrate the impact of misspecification in the second block, we estimate a recursive specification as proposed in Kilian and Park (2009). That is, we restrict b_{12} and b_{34} in Equation (28) to zero. In this case, the interpretation of the shocks changes and we refer to the third and fourth shock as speculative oil price shock and residual stock market shock, respectively.

Figure B.5 in Appendix B.2 displays the IRFs of the recursive SVAR. The response of the real oil price to flow supply and demand shocks in the recursive model is similar to the the one in the block-recursive model. The speculative oil price shock leads to an decrease of the real oil price. However, none of the remaining variables shows any significant response to the speculative oil price shock, except for economic activity which shows a small negative reaction in the first seven month. In the recursive SVAR for the oil market, we cannot distinguish between speculative supply and speculative demand shocks. Rather, the speculative oil price shock contains a mixture of the speculative supply and speculative demand shock. However, the impact of the speculative oil price shocks on oil production and the economy should depend on the source of the speculative oil price shock and, thus, it is not surprising that we are unable to find a clear response of oil production, economic activity, and the stock market to the speculative oil price shock in the recursive specification.

As a further robustness-check, we estimate the SVAR without any restrictions on the interaction, i.e., we estimate the model without the zero restrictions given in Equation (28). In this case, the labeling of the shocks is the same as in Equation (28). However, the difference is that oil production and economic activity can now contemporaneously respond to speculative supply and demand shocks. Figure B.6 and Figure B.7 in Appendix B.2 show the corresponding IRFs. Overall, the unrestricted responses in Figure B.6 are comparable to the block-recursive responses in Figure 4. However, the confidence bands are broader and there is no significant response of the real oil price to flow supply and (almost) no significant response to flow demand shocks.

7 Conclusion

For a non-Gaussian block-recursive SVAR, we derive a small set of identifying moment conditions based on mean independent shocks. Additionally, we derive overidentifying moment conditions from independent shocks and show that these conditions can decrease the asymptotic variance of the block-recursive SVAR estimator. In particular, we prove that the frequently applied Cholesky estimator may be inefficient. Since some of the overidentifying moment conditions may be redundant, i.e., may not decrease the asymptotic variance, or be invalid, i.e., may lead to inconsistent estimates, we employ the block-recursive SVAR pGMM estimator to select only the relevant and valid overidentifying moment conditions.

We demonstrate in a Monte Carlo experiment that imposing a block-recursive structure substantially increases the finite sample performance compared to unrestricted estimators. Furthermore, a second Monte Carlo experiment highlights that, for a given block-recursive structure, the block-recursive SVAR pGMM estimator selects only relevant moment conditions and thereby, increases finite sample precision compared to the block-recursive SVAR GMM estimator and overidentified block-recursive SVAR GMM estimator.

Our application analyzes the impact of flow and speculative supply and flow and speculative demand shocks in the oil market. We argue that there are some but not enough well-justified restrictions available to identify the SVAR based on second moments. Traditional approaches would either rely on additional less credible restrictions or refrain from using any restrictions and solely rely on non-Gaussianity. The proposed block-recursive estimator allows to utilize only the well-justified restrictions and, therefore, offers a compromise between both approaches. The application illustrates that by combining data-driven identification with traditional zero restrictions we are able to gain deeper insights into the transmission of demand and supply shocks in the oil market.

References

Andrews, D. W. (1999). Consistent moment selection procedures for generalized method of moments estimation. *Econometrica 67*(3), 543–563.

Baumeister, C. and J. D. Hamilton (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.

Baumeister, C. and L. Kilian (2016). Understanding the decline in the price of oil since june 2014. *Journal of the Association of Environmental and resource economists 3*(1), 131–158.

Blanchard, O. J. (1989). A traditional interpretation of macroeconomic fluctuations. *The American Economic Review*, 1146–1164.

Blanchard, O. J. and D. Quah (1993). The dynamic effects of aggregate demand and supply disturbances: Reply. *The American Economic Review 83*(3), 653–658.

Breusch, T., H. Qian, P. Schmidt, and D. Wyhowski (1999). Redundancy of moment conditions. *Journal of econometrics 91*(1), 89–111.

Byrne, J. P., M. Lorusso, and B. Xu (2019). Oil prices, fundamentals and expectations. *Energy Economics 79*, 59–75.

Cheng, X. and Z. Liao (2015). Select the valid and relevant moments: An information-based LASSO for GMM with many moments. *Journal of Econometrics 186*(2), 443–464.

Defferrard, M., R. Pena, and N. Perraudin (2017, Dec). PyUNLocBoX: Optimization by Proximal Splitting.

Gouriéroux, C., A. Monfort, and J.-P. Renne (2017). Statistical Inference for Independent Component Analysis: Application to Structural VAR Models. *Journal of Econometrics 196*(1), 111–126.

Guay, A. (2021). Identification of structural vector autoregressions through higher unconditional moments. *Journal of Econometrics 225*(1), 27–46.

Hall, A. R. (2005). *Generalized Method of Moments.* Oxford university press.

Hall, A. R. (2015). Econometricians have their moments: Gmm at 32. *Economic Record 91*, 1–24.

Hall, A. R., A. Inoue, K. Jana, and C. Shin (2007). Information in generalized method of moments estimation and entropy-based moment selection. *Journal of Econometrics 138*(2), 488–512.

Juvenal, L. and I. Petrella (2015). Speculation in the oil market. *Journal of applied econometrics 30*(4), 621–649.

Keweloh, S. A. (2021a). A feasible approach to incorporate information in higher moments in structural vectorautoregressions. *Discussion papers SFB 823*.

Keweloh, S. A. (2021b). A Generalized Method of Moments Estimator for Structural Vector Autoregressions Based on Higher Moments. *Journal of Business & Economic Statistics 39*(3), 772–782.

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), 1053–69.

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

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

Kilian, L. and C. Park (2009). The impact of oil price shocks on the us stock market. *International Economic Review 50*(4), 1267–1287.

Lanne, M., K. Liu, and J. Luoto (2021). Identifying structural vector autoregression via large economic shocks. *Available at SSRN 3910532*.

Lanne, M. and J. Luoto (2021). GMM estimation of non-Gaussian structural vector autoregression. *Journal of Business & Economic Statistics 39*(1), 69–81.

Lanne, M., H. Lütkepohl, and K. Maciejowska (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control 34*(2), 121–131.

Lanne, M., M. Meitz, and P. Saikkonen (2017). Identification and Estimation of Non-Gaussian Structural Vector Autoregressions. *Journal of Econometrics 196*(2), 288–304.

Lewis, D. J. (2021, 05). Identifying Shocks via Time-Varying Volatility. *The Review of Economic Studies*.

Lütkepohl, H. and A. Netšunajev (2017). Structural vector autoregressions with smooth transition in variances. *Journal of Economic Dynamics and Control 84*, 43–57.

Mertens, K. and M. O. Ravn (2013). The dynamic effects of personal and corporate income tax changes in the united states. *American economic review 103*(4), 1212–47.

Moussa, Z. and A. Thomas (2021). Identifying oil supply news shocks and their effects on the global oil market.

Olea, J. L. M., M. Plagborg-Møller, and E. Qian (2022). Svar identification from higher moments: Has the simultaneous causality problem been solved?

Olea, J. L. M., M. Plagbrg-Møller, and E. Qian (2021). Svar identification from higher moments: Has the simultaneous causality problem been solved?

Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics 85*(4), 777–792.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 1–48.

Stock, J. H. and M. W. Watson (2012). Disentangling the channels of the 2007-2009 recession. Technical report, National Bureau of Economic Research.

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

Wainwright, M. J. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (lasso). *IEEE transactions on information theory 55*(5), 2183–2202.

Zha, T. (1999). Block recursion and structural vector autoregressions. *Journal of Econometrics 90*(2), 291–316.

Zou, H. (2006). The adaptive LASSO and its oracle properties. *Journal of the American Statistical Association 101*(476), 1418–1429.

A Supplementary Notation and Proofs

We include the formulas in Appendix A.1 and A.2 for completeness, even though they are standard textbook results (cf. Hall (2005)).

A.1 Asymptotic variance of the block-recursive SVAR GMM estimator

The asymptotic variance of the overidentified block-recursive SVAR GMM estimator defined in Equation (19) is given by

$$V_{\mathbf{N}} := M_{\mathbf{N}} S_{\mathbf{N}} M'_{\mathbf{N}} \tag{A.1}$$

where

$$M_{\mathbf{N}} := \left(G'_{\mathbf{N}} W_{\mathbf{N}} G_{\mathbf{N}}\right)^{-1} G'_{\mathbf{N}} W_{\mathbf{N}}, \qquad S_{\mathbf{N}} := \lim_{T \to \infty} E\left[T g_{\mathbf{N}}(B_0) g_{\mathbf{N}}(B_0)\right],$$
$$G_{\mathbf{N}} := E\left[\frac{\partial f_{\mathbf{N}}(B_0, u_t)}{\partial vec(B)'}\right].$$

Consequently, using the weighting matrix $W^*_{\mathbf{N}} := S^{-1}_{\mathbf{N}}$ leads to the estimator $\hat{B}^*$ with the asymptotic variance

$$V^*_{\mathbf{N}} := (G'_{\mathbf{N}} S^{-1}_{\mathbf{N}} G_{\mathbf{N}})^{-1}, \tag{A.2}$$

which is the lowest possible asymptotic variance (see Hall (2005)).

A.2 Asymptotic variance of the (overidentified) block-recursive SVAR GMM estimator

The asymptotic variance of the overidentified block-recursive SVAR GMM estimator defined in Equation (19) is given by

$$V_{\mathbf{N+D}} := M_{\mathbf{N+D}} S_{\mathbf{N+D}} M'_{\mathbf{N+D}}, \tag{A.3}$$

where

$$M_{\mathbf{N+D}} := \left(G'_{\mathbf{N+D}} W_{\mathbf{N+D}} G\right)^{-1} G'_{\mathbf{N+D}} W_{\mathbf{N+D}}, \qquad S_{\mathbf{N+D}} := \lim_{T\to\infty} E\left[g_{\mathbf{N+D}}(B_0) g_{\mathbf{N+D}}(B_0)'\right],$$

$$G_{\mathbf{N+D}} := \begin{bmatrix} G_{\mathbf{N}} \\ G_{\mathbf{D}} \end{bmatrix}, \qquad g_{\mathbf{N+D}}(B_0) := \begin{bmatrix} g_{\mathbf{N}}(B_0) \\ g_{\mathbf{D}}(B_0) \end{bmatrix},$$

$$G_{\mathbf{D}} := E\left[\frac{\partial f_{\mathbf{D}}(B_0, u_t)}{\partial vec(B)'}\right].$$

Using the weighting matrix $W^*_{\mathbf{N+D}} := S^{-1}_{\mathbf{N+D}}$ leads to the estimator $\hat{B}^*_{\mathbf{N+D}}$ with the asymptotic variance

$$V^*_{\mathbf{N+D}} := (G'_{\mathbf{N+D}} S^{-1}_{\mathbf{N+D}} G_{\mathbf{N+D}})^{-1}, \tag{A.4}$$

which is the lowest possible asymptotic variance (see Hall (2005)). To construct $V_{\mathbf{N+D}_j}$ and $V^*_{\mathbf{N+D}_j}$, $j \in \widetilde{D}$, we replace the moment conditions $f_{\mathbf{D}_j}(B, u_t)$ by moment condition $f_{\mathbf{D}_j}(B, u_t)$, $j \in \widetilde{D}$, in Equation (A.3) and (A.4).

A.3 Identification in the block-recursive SVAR

Proof of Proposition 1.

For ease of notation, we omit the time index t and w.l.o.g., consider an example with two blocks[19]

$$\begin{bmatrix} u_{p_1} \\ u_{p_2} \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 \\ B_{21,0} & B_{22,0} \end{bmatrix} \begin{bmatrix} \varepsilon_{p_1} \\ \varepsilon_{p_2} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix},$$

where u_{p_1} and u_{p_2} contain the reduced form shocks of the first and second block, ε_{p_1} and ε_{p_2} contain the structural shocks of the first and second block, and $B_{11,0}$, $B_{21,0}$, $B_{22,0}$, B_{11}, B_{21}, and B_{22} are the corresponding blocks of the matrices B_0 and B.

First, let $E[f_{\mathbf{2_{p_1}}}(B,u)] = 0$ contain all (co-)variance conditions of shocks in the first block. The block-recursive structure implies that $u_{p_1} = B_{11,0}\varepsilon_{p_1}$. If at most one structural shock in the first block has zero excess kurtosis, it follows from Lanne and Luoto (2021) that the conditions containing only shocks in the first block

$$E \begin{bmatrix} f_{\mathbf{2_{p_1}}}(B,u) \\ f_{\mathbf{4_{p_1}}}(B,u) \end{bmatrix} = 0$$

locally identify $B_{11} = B_{11,0}$, the impact of the shocks in the first block on the variables in the first block.

Second, let $E\left[f_{\mathbf{2_{p_1 p_2}}}(B,u)\right] = 0$ contain all covariance conditions belonging to shocks in both blocks. At the local solution $B_{11} = B_{11,0}$, the covariance conditions containing shocks of both blocks only hold if $B_{21} = B_{21,0}$. To see this, rewrite the covariance conditions as

[19] If the SVAR contains more than two blocks, the procedure outlined in the proof can be repeated multiple times to identify arbitrary many blocks. For example, a SVAR with three blocks

$$\begin{bmatrix} u_{p_1} \\ u_{p_2} \\ u_{p_3} \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 & 0 \\ B_{21,0} & B_{22,0} & 0 \\ B_{32,0} & B_{32,0} & B_{33,0} \end{bmatrix} \begin{bmatrix} \epsilon_{p_1} \\ \epsilon_{p_2} \\ \epsilon_{p_3} \end{bmatrix} \text{ can be written as } \begin{bmatrix} u_{p_1} \\ \tilde{u}_{p_2} \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 \\ \tilde{B}_{21,0} & \tilde{B}_{22,0} \end{bmatrix} \begin{bmatrix} \epsilon_{p_1} \\ \tilde{\epsilon}_{p_2} \end{bmatrix},$$

with $\tilde{u}_{p_2} = [u'_{p_2}, u'_{p_3}]'$, $\tilde{B}_{22,0} = \begin{bmatrix} B_{22,0} & 0 \\ B_{32,0} & B_{33,0} \end{bmatrix}$, $\tilde{B}_{21,0} = \begin{bmatrix} B_{21,0} \\ B_{31,0} \end{bmatrix}$,and $\tilde{\epsilon}_{p_2} = [\epsilon'_{p_2}, \epsilon'_{p_3}]'$. Our proof then shows how to identify $B_{11,0}$, $\tilde{B}_{21,0} = \begin{bmatrix} B_{21,0} \\ B_{31,0} \end{bmatrix}$, and ϵ_{p_1}. Defining

$$\begin{bmatrix} z_{p_2} \\ z_{p_3} \end{bmatrix} := \begin{bmatrix} u_{p_2} \\ u_{p_3} \end{bmatrix} - \begin{bmatrix} B_{21,0} \\ B_{31,0} \end{bmatrix} \epsilon_{p_1} \text{ then yields } \begin{bmatrix} z_{p_2} \\ z_{p_3} \end{bmatrix} = \begin{bmatrix} B_{22,0} & 0 \\ B_{32,0} & B_{33,0} \end{bmatrix} \begin{bmatrix} \epsilon_{p_2} \\ \epsilon_{p_3} \end{bmatrix},$$

which is another block-recursive SVAR with two blocks.

$E\left[e_{p_2}(B)e_{p_1}(B)'\right] = 0$. With the partitioned inverse of B and the block-recursive structure, it holds that $e_{p_2}(B) = -B_{22}^{-1}B_{21}B_{11}^{-1}B_{11,0}\varepsilon_{p_1} + B_{22}^{-1}\left(B_{21,0}\varepsilon_{p_1} + B_{22,0}\varepsilon_{p_2}\right)$. Therefore, with $B_{11} = B_{11,0}$ it holds that

$$E\left[e_{p_2}(B)e_{p_1}(B)'\right] = -B_{22}^{-1}B_{21}E\left[\varepsilon_{p_1}\varepsilon_{p_1}'\right] + B_{22}^{-1}B_{21,0}E\left[\varepsilon_{p_1}\varepsilon_{p_1}'\right] + B_{22,0}E\left[\varepsilon_{p_2}\varepsilon_{p_1}'\right].$$

With $E\left[\varepsilon_{p_1}\varepsilon_{p_1}'\right] = I$ and $E[\varepsilon_{p_2}\varepsilon_{p_1}'] = 0$, the condition $E\left[e_{p_2}(B)e_{p_1}(B)'\right] = 0$ implies $0 = -B_{22}^{-1}\left(B_{21} - B_{21,0}\right)$ at $B_{11} = B_{11,0}$. Therefore, at the local solution $B_{11} = B_{11,0}$ the covariance conditions $E\left[f_{\mathbf{2_{p_1p_2}}}(B,u)\right]$, globally identify $B_{21} = B_{21,0}$ the impact of shocks in the first block on variables in the second block.

Finally, let $E[f_{\mathbf{2_{p_2}}}(B,u)] = 0$ contain all (co-)variance conditions of shocks in the second block. At the solution $B_{11} = B_{11,0}$ and $B_{21} = B_{21,0}$ the unmixed innovations of the second block $e_{p_2}(B)$ are mixtures of the structural shocks in the second block and are not influenced by shocks from the first block. This follows from the partitioned inverse of B and the block-recursive structure such that $e_{p_2}(B) = B_{22}^{-1}B_{22,0}\varepsilon_{p_2}$. If at most one structural shock in the second block has zero excess kurtosis, it then again follows from Lanne and Luoto (2021) that at the solution $B_{11} = B_{11,0}$ and $B_{21} = B_{21,0}$ the remaining conditions containing only shocks in the second block

$$E\begin{bmatrix} f_{\mathbf{2_{p_2}}}(B,u) \\ f_{\mathbf{4_{p_2}}}(B,u) \end{bmatrix} = 0$$

locally identify $B_{22} = B_{22,0}$, meaning the impact of shocks in the second block on variables in the second block.

□

Proof of Proposition 2.

To simplify the notation let

$$\tilde{u}_1 := [u_1, ..., u_{p_i-1}]', \qquad \tilde{e}_1(B) := [e_1(B), ..., e_{p_i-1}(B)]', \qquad \tilde{\varepsilon}_1 := [\varepsilon_1, ..., \varepsilon_{p_i-1}]',$$

$$\tilde{u}_2 := [u_{p_i}, ..., u_{p_{i+1}-1}]', \qquad \tilde{e}_2(B) := [e_{p_i}(B), ..., e_{p_{i+1}-1}(B)]', \qquad \tilde{\varepsilon}_2 := [\varepsilon_{p_i}, ..., \varepsilon_{p_{i+1}-1}]',$$

$$\tilde{u}_3 := [u_{p_{i+1}}, ..., u_n]', \qquad \tilde{e}_3(B) := [e_{p_{i+1}}(B), ..., e_n(B)]', \qquad \tilde{\varepsilon}_3 := [\varepsilon_{p_{i+1}}, ..., \varepsilon_n]',$$

such that $\tilde{u}_1$, $\tilde{e}_1(B)$, and $\tilde{\varepsilon}_1$ contain all reduce form shocks, unmixed innovations, and structural shocks in blocks preceding the ith block of $\mathbb{B}_{brec}$, $\tilde{u}_2$, $\tilde{e}_2(B)$, and $\tilde{\varepsilon}_2$ contain the innovations and shocks in the i-th block of $\mathbb{B}_{brec}$, and $\tilde{u}_3$, $\tilde{e}_3(B)$, and $\tilde{\varepsilon}_3$ contain the innovations and shocks following block i of $\mathbb{B}_{brec}$. Moreover, we denote parts of the B_0 matrix as follows

$$\begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \\ \tilde{u}_3 \end{bmatrix} = \begin{bmatrix} B_{11,0} & 0 & 0 \\ B_{21,0} & B_{22,0} & 0 \\ B_{31,0} & B_{32,0} & B_{33,0} \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \tilde{\varepsilon}_3 \end{bmatrix},$$

and B_{11}, B_{21}, B_{31}, B_{22}, B_{32}, and B_{33} denote the respective parts of a given B matrix.

With the block-recursive structure and the partitioned inverse, it holds that

$$\begin{aligned} \tilde{e}_1(B) &= B_{11}^{-1} B_{11,0} \tilde{\varepsilon}_1, \\ \tilde{e}_2(B) &= -B_{22}^{-1} B_{21} B_{11}^{-1} B_{11,0} \tilde{\varepsilon}_1 + B_{22}^{-1} \left(B_{21,0} \tilde{\varepsilon}_1 + B_{22,0} \tilde{\varepsilon}_2 \right). \end{aligned}$$

For any matrix B satisfying $E\left[f_{\mathbf{2}}(B, u_t)\right] = 0$ and, therefore, $0 = E\left[\tilde{e}_2(B)\tilde{e}_1(B)'\right]$ it holds that $0 = -B_{22}^{-1}\left(B_{21,0} - B_{21}B_{11}^{-1}B_{11,0}\right)B_{11,0}'(B_{11}^{-1})'$ and, thus, $B_{21} = B_{21,0}B_{11,0}^{-1}B_{11}$. Any B Matrix satisfying the condition $0 = E\left[\tilde{e}_2(B)\tilde{e}_1(B)'\right]$ thus yields innovations of the second block equal to

$$\tilde{e}_2(B) = B_{22}^{-1} B_{22,0} \tilde{\varepsilon}_2, \tag{A.5}$$

meaning the innovations of the second block are equal to a linear combination of the structural shocks in the second block. Applying the identification result from Lanne and Luoto (2021) yields that the conditions $E\left[f_{\mathbf{4}_{\tilde{\mathbf{p}}_\mathbf{i}}}(B, u_t)\right] = 0$ locally identify $B_{22,0}$.

Analogously, with the block-recursive structure and the partitioned inverse it holds that

$$\begin{aligned} \tilde{e}_3(B) = &- B_{33}^{-1} \begin{bmatrix} B_{31} & B_{32} \end{bmatrix} \begin{bmatrix} B_{11} & 0 \\ B_{21} & B_{22} \end{bmatrix}^{-1} \begin{bmatrix} B_{11,0} & 0 \\ B_{21,0} & B_{22,0} \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \end{bmatrix} \\ &+ B_{33}^{-1} \left(\begin{bmatrix} B_{31,0} & B_{32,0} \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \end{bmatrix} + B_{33,0} \tilde{\varepsilon}_3 \right). \end{aligned}$$

With $B_{21} = B_{21,0}B_{11,0}^{-1}B_{11}$ it follows that

$$
\begin{aligned}
\tilde{e}_3(B) = & - B_{33}^{-1}\begin{bmatrix} B_{31} & B_{32}\end{bmatrix}\begin{bmatrix} B_{11}^{-1} & 0 \\ -B_{22}^{-1}B_{21,0}B_{11,0}^{-1}B_{11}B_{11}^{-1} & B_{22}^{-1}\end{bmatrix}\begin{bmatrix} B_{11,0} & 0 \\ B_{21,0} & B_{22,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2\end{bmatrix} \\
& + B_{33}^{-1}\left(\begin{bmatrix} B_{31,0} & B_{32,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2\end{bmatrix} + B_{33,0}\tilde{\varepsilon}_3\right) \\
= & - B_{33}^{-1}\begin{bmatrix} B_{31} & B_{32}\end{bmatrix}\begin{bmatrix} B_{11}^{-1} & 0 \\ -B_{22}^{-1}B_{21,0}B_{11,0}^{-1} & B_{22}^{-1}\end{bmatrix}\begin{bmatrix} B_{11,0} & 0 \\ B_{21,0} & B_{22,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2\end{bmatrix} \\
& + B_{33}^{-1}\left(B_{31,0}\tilde{\varepsilon}_1 + B_{32,0}\tilde{\varepsilon}_2 + B_{33,0}\tilde{\varepsilon}_3\right) \\
= & - B_{33}^{-1}\begin{bmatrix} B_{31}B_{11}^{-1} - B_{32}B_{22}^{-1}B_{21,0}B_{11,0}^{-1} & B_{32}B_{22}^{-1}\end{bmatrix}\begin{bmatrix} B_{11,0} & 0 \\ B_{21,0} & B_{22,0}\end{bmatrix}\begin{bmatrix}\tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2\end{bmatrix} \\
& + B_{33}^{-1}\left(B_{31,0}\tilde{\varepsilon}_1 + B_{32,0}\tilde{\varepsilon}_2 + B_{33,0}\tilde{\varepsilon}_3\right).
\end{aligned}
$$

Hence, at $B_{22} = B_{22,0}$ the condition $E\left[f_{\mathbf{2}}(B, u_t)\right] = 0$ implies $0 = E[\tilde{e}_3(B)\tilde{e}_2(B)']$ and therefore,

$$
0 = B_{33}^{-1}(-B_{32}B_{22}^{-1}B_{22,0} + B_{32,0})
$$

which implies $B_{32} = B_{32,0}$. □

A.4 Redundant and relevant moment conditions in the recursive SVAR

The proof of Proposition 3 requires to verify the redundancy conditions from Breusch et al. (1999). However, verifying these conditions is a lengthy task. We derive analytical expressions for the conditions in Online Appendix C and summarize them in Lemma 14 in Online Appendix C. The following proof of Proposition 3 uses Lemma 8 and 14 in Online Appendix C.

Proof of Proposition 3.

In the recursive SVAR, the identifying moment conditions $E[f_{\mathbf{N}}(B, u_t)]$ only contain second-order moment conditions and therefore, are referred to as $E[f_{\mathbf{2}}(B, u_t)]$ in this proof.

Breusch et al. (1999) show that overidentifying moment conditions $E[f_{\mathbf{D}}(B, u_t)]$ are redundant w.r.t. the identifying moment conditions $E[f_{\mathbf{2}}(B,u_t)]$ if and only if

$$G_{\mathbf{D}} = S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}},$$

where

$$G_{\mathbf{D}} := E\left[\frac{\partial f_{\mathbf{D}}(B_0,u_t)}{\partial vec(B)'}\right], \qquad G_{\mathbf{2}} := E\left[\frac{\partial f_{\mathbf{2}}(B_0,u_t)}{\partial vec(B)'}\right],$$
$$S_{\mathbf{2}} := \lim_{T\to\infty} E\left[g_{\mathbf{2}}(B_0)g_{\mathbf{2}}(B_0)'\right], \qquad S_{\mathbf{D2}} := \lim_{T\to\infty} E\left[g_{\mathbf{D}}(B_0)g_{\mathbf{2}}(B_0)'\right].$$

Moreover, Breusch et al. (1999) show that overidentifying moment conditions $E[f_{\mathbf{D}}(B,u_t)]$ are partially redundant w.r.t. $E[f_{\mathbf{2}}(B,u_t)]$ for a subset of coefficients $b \subset vec(B)$ w.r.t. the moment conditions $E[f_{\mathbf{2}}(B,u_t)]$ if and only if

$$G_{\mathbf{D}}^{b} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b} = \left(G_{\mathbf{D}}^{\neg b} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{\neg b}\right)\left(\left(G_{\mathbf{2}}^{\neg b}\right)' S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{\neg b}\right)\left(\left(G_{\mathbf{2}}^{\neg b}\right)' S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b}\right), \tag{A.6}$$

where

$$G_{\mathbf{2}}^{b} := E\left[\frac{\partial f_{\mathbf{2}}(u_t,B_0)}{\partial b'}\right], \qquad G_{\mathbf{D}}^{b} := E\left[\frac{\partial f_{\mathbf{D}}(u_t,B_0)}{\partial b'}\right],$$
$$G_{\mathbf{2}}^{\neg b} := E\left[\frac{\partial f_{\mathbf{2}}(u_t,B_0)}{\partial (\neg b)'}\right], \qquad G_{\mathbf{D}}^{\neg b} := E\left[\frac{\partial f_{\mathbf{D}}(u_t,B_0)}{\partial (\neg b)'}\right],$$

and where $\neg b$ denotes all unrestricted elements of B not contained in b. With Lemma 8 it holds that ${G_{\mathbf{2}}^{b_i}}' S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_j} = 0$ for $i,j \in \{1,\ldots,n\}$ with $i \neq j$. Therefore, for any vector $b_i = [b_{ii},...,b_{ni}]$ representing the impact of the ith structural shock $\epsilon_{i,t}$ it holds that ${G_{\mathbf{2}}^{b_i}}' S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{\neg b_i}$ is zero. Therefore, for any vector $b_i = [b_{ii},...,b_{ni}]$ the right hand side of Equation (A.6) is zero and hence the partial redundancy condition simplifies to

$$G_{\mathbf{D}}^{b_i} - S_{\mathbf{D2}} S_{\mathbf{2}}^{-1} G_{\mathbf{2}}^{b_i} = 0.$$

The statements then follow from Lemma 14. $\square$

A.5 Asymptotic variance of the block-recursive SVAR pGMM estimator

We show how to derive the asymptotic variance of the pGMM estimator, $V_{\mathbf{N}+\mathbf{A}}$, based on Remark 3.5 of Cheng and Liao (2015). We first show Lemma 1 and then apply the result in Remark 3.5 of Cheng and Liao (2015). Recall that $E[f_{\mathbf{I}}(B, u_t)]$ and $E[f_{\mathbf{R}}(B, u_t)]$ denote the sets of invalid and redundant moment conditions, respectively. Denote $E[f_{\mathbf{U}}(B, u_t)]$ as moment conditions either in $E[f_{\mathbf{I}}(B, u_t)]$ or $E[f_{\mathbf{R}}(B, u_t)]$ and the number of moment conditions $E[f_{\mathbf{U}}(B, u_t)]$ by $k_{\mathbf{U}}$. Similarly, we denote $k_{\mathbf{A}}$ as the number of moment conditions in $E[f_{\mathbf{A}}(B, u_t)]$. Further, define the number of unrestricted elements in $vec(B)$ as d_B. In the proof of Lemma 1, we use the indices $1 \equiv \mathbf{N}+\mathbf{A}$, $2 \equiv (\mathbf{N}+\mathbf{A}, \mathbf{U})$, $3 \equiv (\mathbf{U}, \mathbf{N}+\mathbf{A})$, and $4 \equiv \mathbf{U}$ to keep notation uncluttered. Let $\iota^* = (\iota', \mathbf{0}'_{k_{\mathbf{U}}})'$ where $\iota = (1, \ldots, 1)'$ is a $d_B \times 1$ vector, i.e., $\iota^{*\prime} A \iota^*$ gives the leading $d_B \times d_B$-upper west block of an arbitrary $(d_B + k_{\mathbf{U}}) \times (d_B + k_{\mathbf{U}})$ matrix A.

Lemma 1.

$$\iota^{*\prime} \left(\Gamma' W \Gamma\right)^{-1} \left(\Gamma' W S_{\mathbf{N}+\mathbf{D}} W \Gamma\right) \left(\Gamma' W \Gamma\right)^{-1} \iota^* = V_{\mathbf{N}+\mathbf{A}},$$

where

$$\Gamma := \begin{bmatrix} G_{\mathbf{N}+\mathbf{A}} & \mathbf{0}_{(k_{\mathbf{N}}+k_{\mathbf{A}})\times k_{\mathbf{U}}} \\ G_{\mathbf{U}} & -I_{k_{\mathbf{U}}} \end{bmatrix}, \qquad V_{\mathbf{N}+\mathbf{A}} := M_{\mathbf{N}+\mathbf{A}} S_{\mathbf{N}+\mathbf{A}} M'_{\mathbf{N}+\mathbf{A}}$$

with

$$M_{\mathbf{N}+\mathbf{A}} := \left(G'_{\mathbf{N}+\mathbf{A}} W^{pi}_{\mathbf{N}+\mathbf{A}} G_{\mathbf{N}+\mathbf{A}}\right)^{-1} G'_{\mathbf{N}+\mathbf{A}} W^{pi}_{\mathbf{N}+\mathbf{A}}, \qquad S_{\mathbf{N}+\mathbf{A}} := \lim_{T\to\infty} E\left[g_{\mathbf{N}+\mathbf{A}}(B_0) g_{\mathbf{N}+\mathbf{A}}(B_0)'\right],$$

$$G_{\mathbf{N}+\mathbf{A}} := \begin{bmatrix} G_{\mathbf{N}} \\ G_{\mathbf{A}} \end{bmatrix}, \qquad W^{pi}_{\mathbf{N}+\mathbf{A}} := \left(W_{\mathbf{N}+\mathbf{A}} - W_{\mathbf{N}+\mathbf{A},\mathbf{I}\cup\mathbf{R}} W^{-1}_{\mathbf{I}\cup\mathbf{R}} W_{\mathbf{I}\cup\mathbf{R},\mathbf{N}+\mathbf{A}}\right),$$

$$G_{\mathbf{A}} := E\left[\frac{\partial f_{\mathbf{A}}(B_0, u_t)}{\partial vec(B)'}\right], \qquad W_{\mathbf{N}+\mathbf{D}} := \begin{bmatrix} W_{\mathbf{N}+\mathbf{A}} & W_{\mathbf{N}+\mathbf{A},\mathbf{I}\cup\mathbf{R}}, \\ W_{\mathbf{I}\cup\mathbf{R},\mathbf{N}+\mathbf{A}} & W_{\mathbf{I}\cup\mathbf{R}} \end{bmatrix},$$

$$W_{\mathbf{N}+\mathbf{A}} \in \mathbb{R}^{(k_{\mathbf{N}}+k_{\mathbf{A}})\times(k_{\mathbf{N}}+k_{\mathbf{A}})}, \qquad W_{\mathbf{N}+\mathbf{A},\mathbf{I}\cup\mathbf{R}} \in \mathbb{R}^{(k_{\mathbf{N}}+k_{\mathbf{A}})\times(k_{\mathbf{D}}-k_{\mathbf{A}})},$$

$$W_{\mathbf{I}\cup\mathbf{R},\mathbf{N}+\mathbf{A}} = W'_{\mathbf{N}+\mathbf{A},\mathbf{I}\cup\mathbf{R}}, \qquad W_{\mathbf{I}\cup\mathbf{R}} \in \mathbb{R}^{(k_{\mathbf{D}}-k_{\mathbf{A}})\times(k_{\mathbf{D}}-k_{\mathbf{A}})}.$$

Proof. Recall that $G_{\mathbf{N+A}}$ and $G_{\mathbf{U}}$ have dimension $(k_{\mathbf{N}} + k_{\mathbf{A}}) \times d_B$ and $k_{\mathbf{U}} \times d_B$, respectively. We define

$$L := \begin{bmatrix} L_1 & L_2 \\ L_3 & L_4 \end{bmatrix} := \left(\Gamma' W \Gamma\right)^{-1}.$$

Additionally, let

$$N := \begin{bmatrix} N_1 & N_2 \\ N_3 & N_4 \end{bmatrix} := \left(\Gamma' W S_{\mathbf{N+D}} W \Gamma\right),$$

and denote the inverse of W by

$$W^{ipi} := \begin{bmatrix} W_1^{ipi} & W_2^{ipi} \\ W_3^{ipi} & W_4^{ipi} \end{bmatrix} := W^{-1} = \begin{bmatrix} W_1 & W_2 \\ W_3 & W_4 \end{bmatrix}^{-1}.$$

Let $W_1^{pi} := \left(W_1 - W_2 W_4^{-1} W_3\right)$. Then, by the partitioned inverse, $W_1^{ipi} := \left(W_1^{pi}\right)^{-1}$. By similar arguments as leading to (2.18) in the Online Appendix of Cheng and Liao (2015), we get that

$$L_1 = \left(G_1' \left(W_1 - W_2 W_4^{-1} W_3\right) G_1\right)^{-1} = \left(G_1' W_1^{pi} G_1\right)^{-1}$$

and, by using the partitioned inverse formula again, and similar arguments as leading to (2.10), (2.11) and (2.18) in the Online Appendix of Cheng and Liao (2015), that

$$\begin{aligned} L_3 &= -W_4^{-1} \left(-G_1' W_2 - G_4' W_4\right)' \left(G_1' W_1^{pi} G_1\right)^{-1} \\ &= \left(W_4^{-1} W_3 G_1 + G_4\right) L_1 \\ &= X L_1, \end{aligned} \tag{A.7}$$

where we used that $W_4' = W_4$, $W_3 = W_2'$ and $X := \left(W_4^{-1} W_3 G_1 + G_4\right)$. Further, let

$$H := \begin{bmatrix} H_1 & H_2 \\ H_3 & H_4 \end{bmatrix} := W S_{\mathbf{N+D}} W,$$

where

$$\begin{aligned}
H_1 &:= W_1S_1W_1 + W_2S_3W_1 + W_1S_2W_3 + W_2S_4W_3 \\
H_2 &:= W_1S_1W_2 + W_2S_3W_2 + W_1S_2W_4 + W_2S_4W_4 \\
H_3 &:= W_3S_1W_1 + W_4S_3W_1 + W_3S_2W_3 + W_4S_4W_3 \\
H_4 &:= W_3S_1W_2 + W_4S_3W_2 + W_3S_2W_4 + W_4S_4W_4.
\end{aligned}$$

Note that $H_3 = H_2'$ since $W_3 = W_2'$, $W_1 = W_1'$, $W_4 = W_4'$, $S_3 = S_2'$, $S_1 = S_1'$ and $S_4 = S_4'$. Hence, similar to (2.11) in the Online Appendix of Cheng and Liao (2015),

$$\begin{aligned}
N_1 &= G_1'H_1G_1 + G_4'H_3G_1 + G_1'H_2G_4 + G_4'H_4G_4 \\
&= G_1'H_1G_1 + G_4'H_2'G_1 + G_1'H_2G_4 + G_4'H_4G_4 \\
N_2 &= -G_1'H_2 - G_4'H_4 \\
N_3 &= N_2' \\
N_4 &= H_4.
\end{aligned}$$

Then,

$$\begin{aligned}
\iota^{*\prime}\left(\Gamma'W\Gamma\right)^{-1}\left(\Gamma'WS_{\mathbf{N+D}}W\Gamma\right)\left(\Gamma'W\Gamma\right)^{-1}\iota^* &= \iota^{*\prime}LNL\iota^* \\
&= L_1N_1L_1 + L_2N_3L_1 + L_1N_2L_3 + L_2N_4L_3 \\
&= L_1N_1L_1 + L_3'N_3L_1 + L_1N_2L_3 + L_3'N_4L_3 \\
&\overset{\text{(A.7)}}{=} L_1N_1L_1 + L_1'X'N_2'L_1 + L_1N_2XL_1 + L_1'X'N_4XL_1 \\
&= L_1\left(N_1 + X'N_2' + N_2X + X'N_4X\right)L_1, \qquad \text{(A.8)}
\end{aligned}$$

where we used that $L_1' = L_1$, $L_3' = L_2$, and $N_3' = N_2$.

Next, define $Y := N_1 + X'N_2' + N_2X + X'N_4X$. Then, multiplying out gives

$$\begin{aligned}
Y =& G_1'H_1G_1 + G_4'H_3G_1 + G_1'H_2G_4 + G_4'H_4G_4 + \left(G_1'W_2W_4^{-1} + G_4'\right)\left(-H_2'G_1 - H_4'G_4\right) \\
&+ \left(-G_1'H_2 - G_4'H_4\right)\left(W_4^{-1}W_2'G_1 + G_4\right) + \left(G_1'W_2W_4^{-1} + G_4'\right)H_4\left(W_4^{-1}W_2'G_1 + G_4\right) \\
=& G_1'W_2W_4^{-1}H_4W_4^{-1}W_2'G_1 + G_1'H_1G_1 - G_1'W_2W_4^{-1}H_2'G_1 - G_1'H_2W_4^{-1}W_2'G_1 \\
=& G_1'\left(W_2W_4^{-1}H_4W_4^{-1}W_2' + H_1 - W_2W_4^{-1}H_2' - H_2W_4^{-1}W_2'\right)G_1 \\
=& G_1'\left(W_2W_4^{-1}W_3S_1W_2W_4^{-1}W_3 + W_1S_1W_1 - W_2W_4^{-1}W_3S_1W_1 - W_1S_1W_2W_4^{-1}W_3\right)G_1 \\
=& G_1'\left(W_1 - W_2W_4^{-1}W_3\right)S_1\left(W_1 - W_2W_4^{-1}W_3\right)G_1 \\
=& G_1'W_1^{pi}S_1W_1^{pi}G_1 \qquad \text{(A.9)}
\end{aligned}$$

Plugging (A.9) into (A.8), we obtain

$$\begin{aligned}
&\iota^{*\prime}\left(\Gamma'W\Gamma\right)^{-1}\left(\Gamma'WS_{\mathbf{N+D}}W\Gamma\right)\left(\Gamma'W\Gamma\right)^{-1}\iota^{*} \\
&= L_1\left(G_1'W_1^{pi}S_1W_1^{pi}\right)G_1L_1 \\
&= \left(G_1'W_1^{pi}G_1\right)^{-1}\left(G_1'W_1^{pi}S_1W_1^{pi}G_1\right)\left(G_1'W_1^{pi}G_1\right)^{-1} \\
&= \left(G_{\mathbf{N+A}}'W_{\mathbf{N+A}}^{pi}G_{\mathbf{N+A}}\right)^{-1}\left(G_{\mathbf{N+A}}'W_{\mathbf{N+A}}^{pi}S_{\mathbf{N+A}}W_{\mathbf{N+A}}^{pi}G_{\mathbf{N+A}}\right)\left(G_{\mathbf{N+A}}'W_{\mathbf{N+A}}^{pi}G_{\mathbf{N+A}}\right)^{-1}
\end{aligned}$$

which was to show. □

Note that in the following proposition, we treat the number of valid and relevant moment conditions, $k_{\mathbf{A}}$, and the number of invalid moment conditions, $k_{\mathbf{I}}$, as fixed constants to keep our asymptotic results for the pGMM estimator in line with the asymptotic results for the block-recursive SVAR GMM estimator in Equation (19). Cheng and Liao (2015) allow both $k_{\mathbf{A}}$ and $k_{\mathbf{I}}$ to increase with the sample size. However, their results also hold when the number of moment conditions is fixed.

Proposition 4. *Assume that the Assumptions in Theorem 3.3 of Cheng and Liao (2015) hold. Further, assume that* $E\left[\frac{\partial f_{\mathbf{A}}(B_0,u_t)}{\partial vec(B)'}\right] = \frac{\partial E[f_{\mathbf{A}}(B_0,u_t)]}{\partial vec(B)'}$ *and Assumption 1 and 2 hold. Then,*

$$\sqrt{T}\left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0, V_{\mathbf{N+A}})$$

Proof. Define $\Sigma_{CL} := \left(\Gamma' W \Gamma\right)^{-1} \left(\Gamma' W S_{\mathbf{N+D}} W \Gamma\right) \left(\Gamma' W \Gamma\right)^{-1}$ and $\gamma = \left(\nu', \mathbf{0}'_{k_{\mathbf{U}}}\right)'$ where $\nu \in \mathbb{R}^{d_B}$ is an arbitrary vector. Then, by Remark 3.5 of Cheng and Liao (2015),

$$\left|\left|\Sigma_{CL}^{1/2}\, \gamma\right|\right|^{-1} \sqrt{T} \nu' \left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0,1),$$

where $||a|| := \sqrt{a'a}$ is the ℓ_2-norm of an arbitrary vector a.

Note that Lemma 1 immediately implies $\left|\left|\Sigma_{CL}^{1/2} \gamma\right|\right| = \sqrt{\gamma' \Sigma_{CL} \gamma} = \sqrt{\nu' V_{\mathbf{N+A}}(W) \nu}$. Hence,

$$\left|\left|V_{\mathbf{N+A}}(W)^{1/2} \nu\right|\right|^{-1} \sqrt{T} \nu' \left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0,1),$$

where $V_{\mathbf{N+A}}(W)$ is the asymptotic variance of $vec(\hat{B}_{\mathbf{N+D}})$ since it holds that

$$\nu^{*\prime} V_{\mathbf{N+A}}(W) \nu^* = \left|\left|V_{\mathbf{N+A}}(W)^{1/2} \nu\right|\right|^{-2} \nu' V_{\mathbf{N+A}}(W) \nu = 1$$

where $\nu^* := \left|\left|V_{\mathbf{N+A}}(W)^{1/2} \nu\right|\right|^{-1} \nu$.

Consequently, using the Cramér-Wold device, we get

$$\sqrt{T} \left(vec(\hat{B}_{\mathbf{N+D}}) - vec(B_0)\right) \xrightarrow{d} \mathcal{N}(0, V_{\mathbf{N+A}}).$$

□

A.6 Choice of maximum λ in the cross-validation

In the following, we illustrate how to choose the maximum value of λ in the cross-validation. Define the loss function of the pGMM estimator as

$$L^*(B,\beta) := L(B,\beta) + \lambda \sum_{i \in \widetilde{D}} \omega_i\, |\beta_i|, \tag{A.10}$$

where $L(B,\beta) := \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B,\beta) \end{bmatrix}' W \begin{bmatrix} g_{\mathbf{N}}(B) \\ g_{\mathbf{D}}(B,\beta) \end{bmatrix}$.

Further, let $z \in \partial ||\beta||_1$, where $z \in \mathbb{R}^{k_{\mathbf{D}}}$, denote the subgradient for the ℓ_1-norm evaluated at β, i.e.,

$$z_i = sign(\beta_i),\ \text{if}\, \beta_i \neq 0,$$
$$z_i \in [-1,1], \quad \text{if}\, \beta_i = 0, \tag{A.11}$$

for $i = 1, \ldots, k_{\mathbf{D}}$ (Wainwright, 2009). Then, the first order condition of the pGMM estimator with respect to β_i, $i = 1, \ldots, k_{\mathbf{D}}$, evaluated at β and B is

$$\frac{\partial L^*(B,\beta)}{\partial \beta_i} = \frac{\partial L(B,\beta)}{\partial \beta_i} + \lambda\omega_i\, z_i = 0 \tag{A.12}$$

Note that $\omega_i \geq 0$. However, if $\omega_i = 0$, β_i is not penalized and therefore, we only consider $i \in \widetilde{P} := \{j \in \tilde{D} |\ \omega_j > 0\}$ for which, by definition, $\omega_i > 0$ when choosing the maximum value of λ in the cross-validation. By (A.11) and (A.12), $\beta = \mathbf{0} = (0, \ldots, 0)'$ and $B = B_0$ minimize the loss function in (A.10) only if

$$\frac{1}{\omega_i}\ \frac{\partial L(B_0, \mathbf{0})}{\partial \beta_i} \in \lambda[-1,1],$$

for $i \in \widetilde{P}$. Thus,

$$\max_{i \in \widetilde{P}} \left| \frac{1}{\omega_i}\ \frac{\partial L(B_0, \mathbf{0})}{\partial \beta_i} \right| \leq \lambda.$$

This motivates us to use

$$\lambda_{\max} = \max_{i \in \widetilde{P}} \left| \frac{1}{\omega_i}\ \frac{\partial L(B_0, \mathbf{0})}{\partial \beta_i} \right|.$$

as the largest value in the cross-validation. Note that any $\lambda > \lambda_{\max}$ would not have an effect on β as $\lambda_{\max}$ already shrinks all elements of β to zero. In practice, we replace B_0 and ω_i by consistent estimators to obtain $\lambda_{\max}$. Furthermore, we consider a weight ω_j to be positive and hence, $j \in \widetilde{P}$, if $\omega_j / \sum_{k \in \widetilde{D}} \omega_k > 10^{-4}$.

B Supplementary Figures and Tables

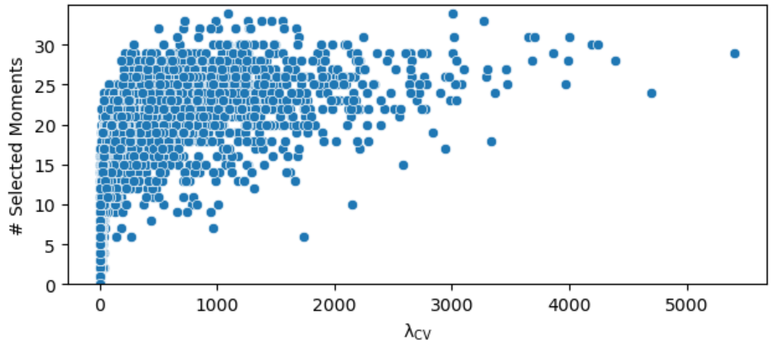
B.1 Finite sample performance

Table B.1: Finite sample performance including Post-LASSO.

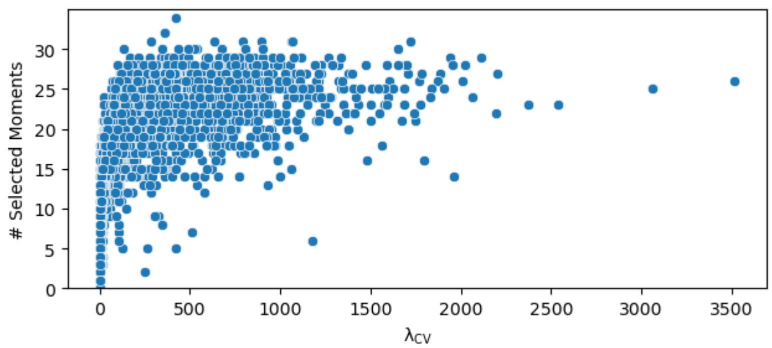
		GMM	oGMM	GMM-Oracle	pGMM	Post-pGMM
$T = 100$	$\hat{B}$	$\begin{bmatrix} \underset{(1.09)}{9.93} & . & . & . \\ \underset{(1.21)}{4.98} & \underset{(1.02)}{9.86} & . & . \\ \underset{(1.49)}{4.97} & \underset{(1.29)}{4.95} & \underset{(1.12)}{9.83} & . \\ \underset{(1.71)}{4.96} & \underset{(1.46)}{4.93} & \underset{(1.27)}{4.91} & \underset{(1.08)}{9.78} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.07)}{9.77} & . & . & . \\ \underset{(1.31)}{4.90} & \underset{(1.01)}{9.71} & . & . \\ \underset{(1.69)}{4.89} & \underset{(1.43)}{4.88} & \underset{(1.10)}{9.70} & . \\ \underset{(2.07)}{4.90} & \underset{(1.74)}{4.88} & \underset{(1.46)}{4.88} & \underset{(1.09)}{9.69} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.07)}{9.76} & . & . & . \\ \underset{(1.17)}{4.91} & \underset{(1.02)}{9.70} & . & . \\ \underset{(1.50)}{4.91} & \underset{(1.26)}{4.88} & \underset{(1.10)}{9.69} & . \\ \underset{(1.81)}{4.92} & \underset{(1.51)}{4.88} & \underset{(1.25)}{4.88} & \underset{(1.10)}{9.67} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.09)}{9.96} & . & . & . \\ \underset{(1.15)}{5.00} & \underset{(1.01)}{9.88} & . & . \\ \underset{(1.46)}{4.98} & \underset{(1.22)}{4.96} & \underset{(1.11)}{9.85} & . \\ \underset{(1.71)}{4.99} & \underset{(1.42)}{4.96} & \underset{(1.21)}{4.95} & \underset{(1.10)}{9.82} \end{bmatrix}$	$\begin{bmatrix} \underset{(1.06)}{9.84} & . & . & . \\ \underset{(1.09)}{4.96} & \underset{(1.01)}{9.79} & . & . \\ \underset{(1.39)}{4.94} & \underset{(1.19)}{4.93} & \underset{(1.11)}{9.77} & . \\ \underset{(1.61)}{4.94} & \underset{(1.38)}{4.93} & \underset{(1.18)}{4.91} & \underset{(1.11)}{9.73} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	24.22	24.22
	Bias	−0.0883	−0.1806	−0.1804	−0.0650	−0.1256
	MSE	1.27	1.40	1.28	1.25	1.21
	λ	.	.	.	71.08	.
		GMM	oGMM	GMM-Oracle	pGMM	Post-pGMM
$T = 250$	$\hat{B}$	$\begin{bmatrix} \underset{(0.43)}{9.97} & . & . & . \\ \underset{(0.51)}{4.99} & \underset{(0.43)}{9.96} & . & . \\ \underset{(0.64)}{4.98} & \underset{(0.52)}{5.00} & \underset{(0.45)}{9.93} & . \\ \underset{(0.72)}{4.98} & \underset{(0.61)}{4.99} & \underset{(0.51)}{4.98} & \underset{(0.45)}{9.91} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.40)}{9.90} & . & . & . \\ \underset{(0.49)}{4.96} & \underset{(0.40)}{9.90} & . & . \\ \underset{(0.65)}{4.96} & \underset{(0.51)}{4.97} & \underset{(0.42)}{9.87} & . \\ \underset{(0.73)}{4.97} & \underset{(0.61)}{4.96} & \underset{(0.49)}{4.97} & \underset{(0.42)}{9.86} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.40)}{9.90} & . & . & . \\ \underset{(0.44)}{4.97} & \underset{(0.40)}{9.90} & . & . \\ \underset{(0.59)}{4.97} & \underset{(0.46)}{4.97} & \underset{(0.42)}{9.87} & . \\ \underset{(0.65)}{4.98} & \underset{(0.54)}{4.97} & \underset{(0.44)}{4.96} & \underset{(0.42)}{9.85} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.42)}{9.99} & . & . & . \\ \underset{(0.45)}{5.01} & \underset{(0.41)}{9.97} & . & . \\ \underset{(0.59)}{5.01} & \underset{(0.46)}{5.02} & \underset{(0.42)}{9.94} & . \\ \underset{(0.66)}{5.02} & \underset{(0.55)}{5.01} & \underset{(0.44)}{5.00} & \underset{(0.43)}{9.92} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.41)}{9.93} & . & . & . \\ \underset{(0.44)}{4.98} & \underset{(0.41)}{9.92} & . & . \\ \underset{(0.57)}{4.98} & \underset{(0.45)}{4.99} & \underset{(0.43)}{9.89} & . \\ \underset{(0.64)}{4.99} & \underset{(0.54)}{4.98} & \underset{(0.44)}{4.97} & \underset{(0.44)}{9.87} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	27.20	27.20
	Bias	−0.0311	−0.0676	−0.0656	−0.0114	−0.0480
	MSE	0.53	0.51	0.48	0.48	0.48
	λ	.	.	.	118.92	.
		GMM	oGMM	GMM-Oracle	pGMM	Post-pGMM
$T = 1000$	$\hat{B}$	$\begin{bmatrix} \underset{(0.11)}{10.00} & . & . & . \\ \underset{(0.13)}{5.00} & \underset{(0.11)}{9.99} & . & . \\ \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{9.99} & . \\ \underset{(0.19)}{4.99} & \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.10)}{9.98} & . & . & . \\ \underset{(0.12)}{4.99} & \underset{(0.10)}{9.97} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.16)}{4.99} & \underset{(0.14)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.10)}{9.98} & . & . & . \\ \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.10)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.15)}{4.99} & \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.97} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{10.00} & . & . & . \\ \underset{(0.11)}{5.00} & \underset{(0.10)}{9.99} & . & . \\ \underset{(0.13)}{5.00} & \underset{(0.11)}{5.00} & \underset{(0.10)}{10.00} & . \\ \underset{(0.16)}{5.00} & \underset{(0.13)}{5.00} & \underset{(0.11)}{5.00} & \underset{(0.10)}{9.98} \end{bmatrix}$	$\begin{bmatrix} \underset{(0.11)}{9.99} & . & . & . \\ \underset{(0.11)}{5.00} & \underset{(0.11)}{9.98} & . & . \\ \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.10)}{9.98} & . \\ \underset{(0.16)}{5.00} & \underset{(0.13)}{4.99} & \underset{(0.11)}{4.99} & \underset{(0.11)}{9.97} \end{bmatrix}$
	#Mo	10.00	57.00	40.00	29.59	29.59
	Bias	−0.0076	−0.0158	−0.0158	−0.0021	−0.0122
	MSE	0.13	0.12	0.11	0.12	0.12
	λ	.	.	.	75.34	.

The table reports the average $\bar{b}_{ij}$ and the corresponding estimated MSE (in parentheses) of each estimated element in $\hat{B}$ as well as the BIAS and MSE across estimated elements in $\hat{B}$ over $3,500$ Monte Carlo replicates for the GMM estimator, the oGMM estimator, the GMM-Oracle estimator, the pGMM estimator, and the Post-pGMM estimator. The Post-pGMM estimator uses only the overidentifying moment conditions selected by the pGMM estimator for the estimation of the block-recursive SVAR. All estimator use zero restrictions which are highlighted by the dots.

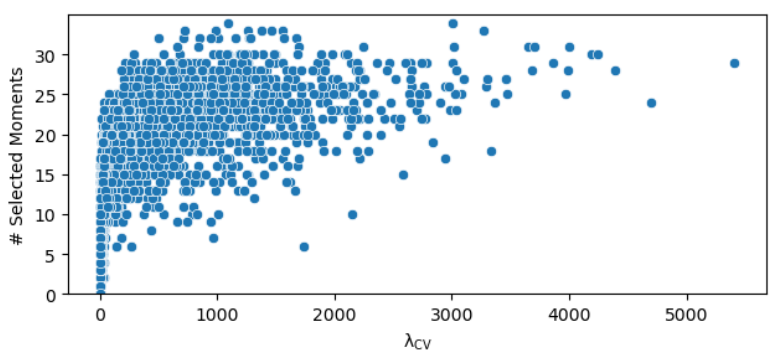
Figure B.1: Relationship of chosen λ_{CV} and Number of Selected Moments across Monte Carlo runs.



(a) $T = 100$



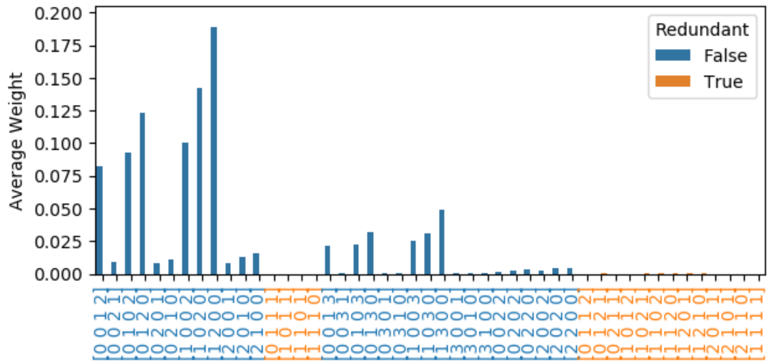
(b) $T = 250$



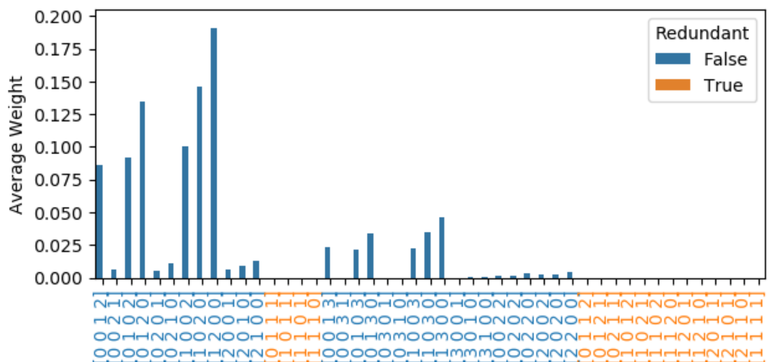
(c) $T = 1000$

Note: The figure shows the chosen λ_{CV} in the cross-validation and the corresponding number of selected moments for each of the $M = 3,500$ Monte Carlo simulations.

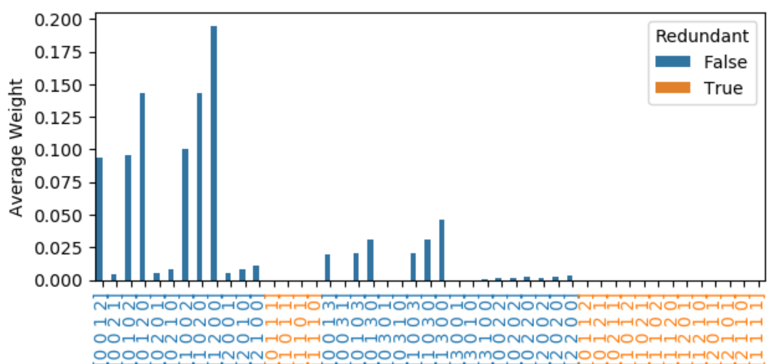
Figure B.2: Average Weight of Moments across Monte Carlo runs.



(a) $T = 100$



(b) $T = 250$



(c) $T = 1000$

Note: The figure shows the average weight of each moment across $M = 3,500$ Monte Carlo simulations. Redundant moment (orange) and relevant moments (blue) are displayed on the x-axis. Each x-axis label abbreviates a moment condition, e.g., $[0, 1, 2, 1]$ corresponds to $E[e(B)^0_{1,t}\, e(B)^1_{2,t}\, e(B)^2_{3,t}\, e(B)^1_{4,t}]$.

B.2 Empirical illustration

This section contains supplementary material and robustness checks for the application presented in Section 6.

Table B.2 shows descriptive statistics of the variables used in the SVAR. Table B.3 shows the

Table B.2: Descriptive statistics.

	Mean	Median	Std. deviation	Variance	Skewness	Kurtosis
O_t	0.078	0.19	1.5	2.26	-1.66	10.8
Y_t	0.20	0.29	0.60	0.37	-1.2	5.21
OP_t	0.32	0.03	7.31	53.4	0.06	4.46
SR_t	0.34	0.62	3.61	13.03	-0.82	3.67

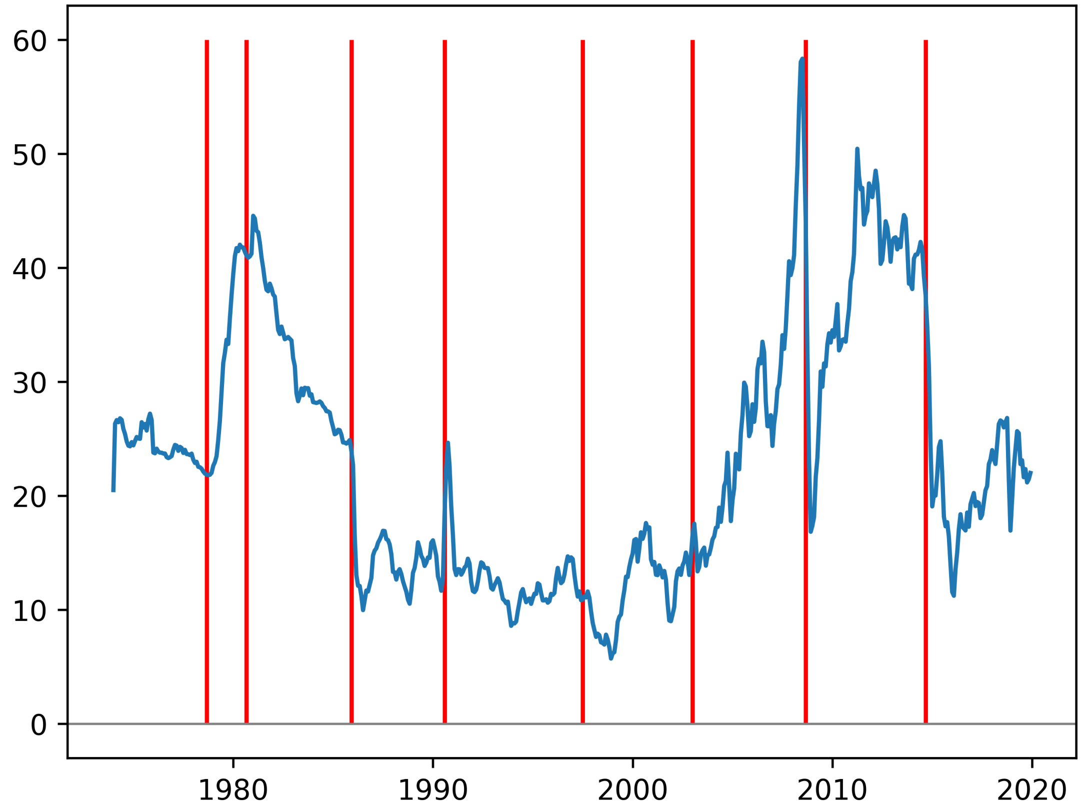
correlation between the estimated structural shocks from the block-recursive SVAR pGMM estimator and the reduced form shocks. Figure B.3 shows the historical evolution of the real oil

Table B.3: Correlation of reduced form and estimated structural shocks.

	u^O	u^Y	u^{OP}	u^{SR}
ε^s	1	-0.03	-0.13	-0.05
ε^d	0.06	1	0.12	0.06
ε^{s-exp}	-0.08	0.02	0.94	-0.27
ε^{d-exp}	-0.05	0.02	0.33	0.96

price.

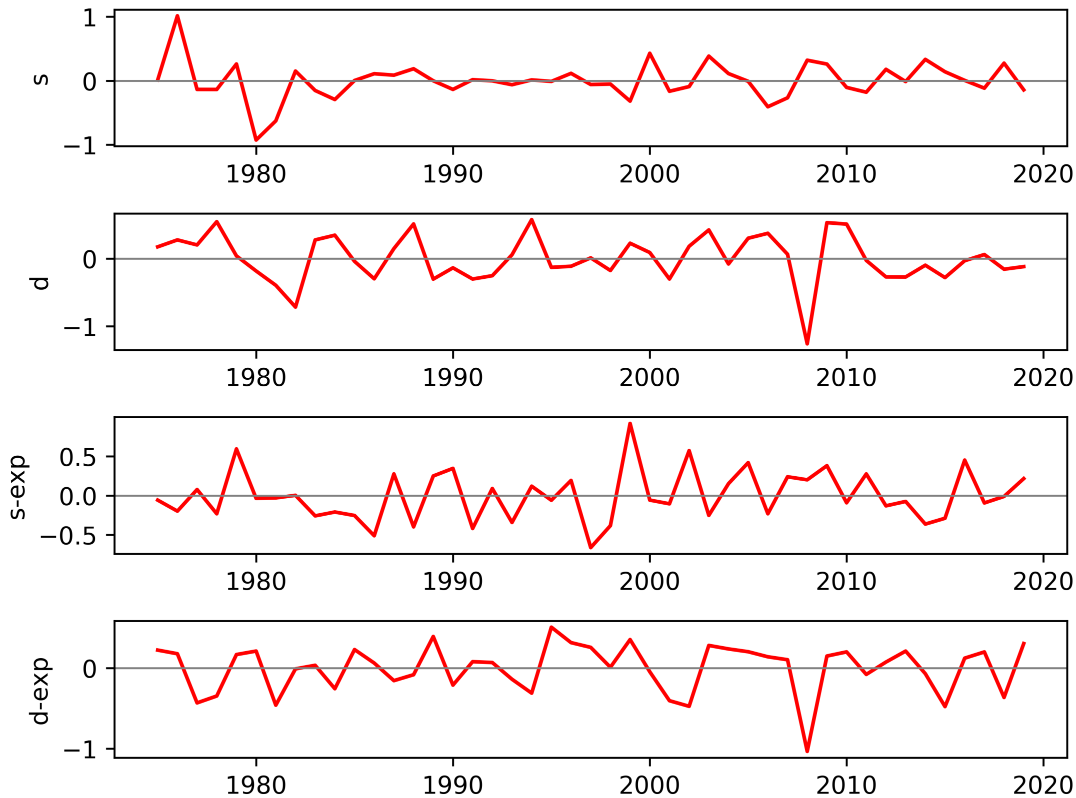
Figure B.3: Real oil price.



Note: The vertical bars indicate the following events: Iranian Revolution 1978 : 9, Iran Iraq War 1980 : 9, collapse of OPEC 1985 : 12, Persian Gulf War 1990 : 8, Asian Financial Crisis of 1997 :7, Iraq War 2003 : 1, the collapse of Lehman Brothers (2008 : 9), and the oil price decline in mid 2014.

Figure B.4 shows the estimated structural shocks across years.

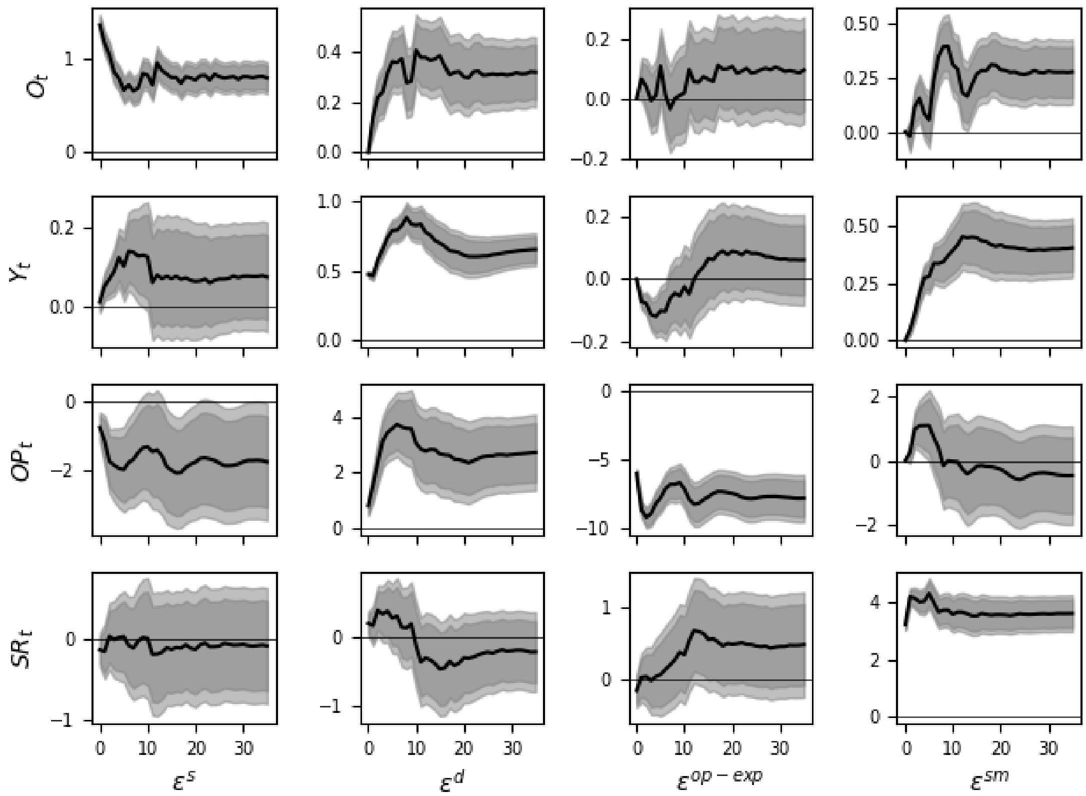
Figure B.4: Estimated structural shocks, averaged to annual frequency.



Note: The figure shows the average across years for each estimated structural shocks of the block-recursive SVAR pGMM estimator.

Figure B.5 shows the IRF for the recursive oil market SVAR from Section 6 estimated with the GMM estimator from Equation (15). In the recursive SVAR, the GMM estimator is just identified and equal to the estimator obtained by applying the Cholesky decomposition to the variance-covariance matrix of the reduced form shocks.

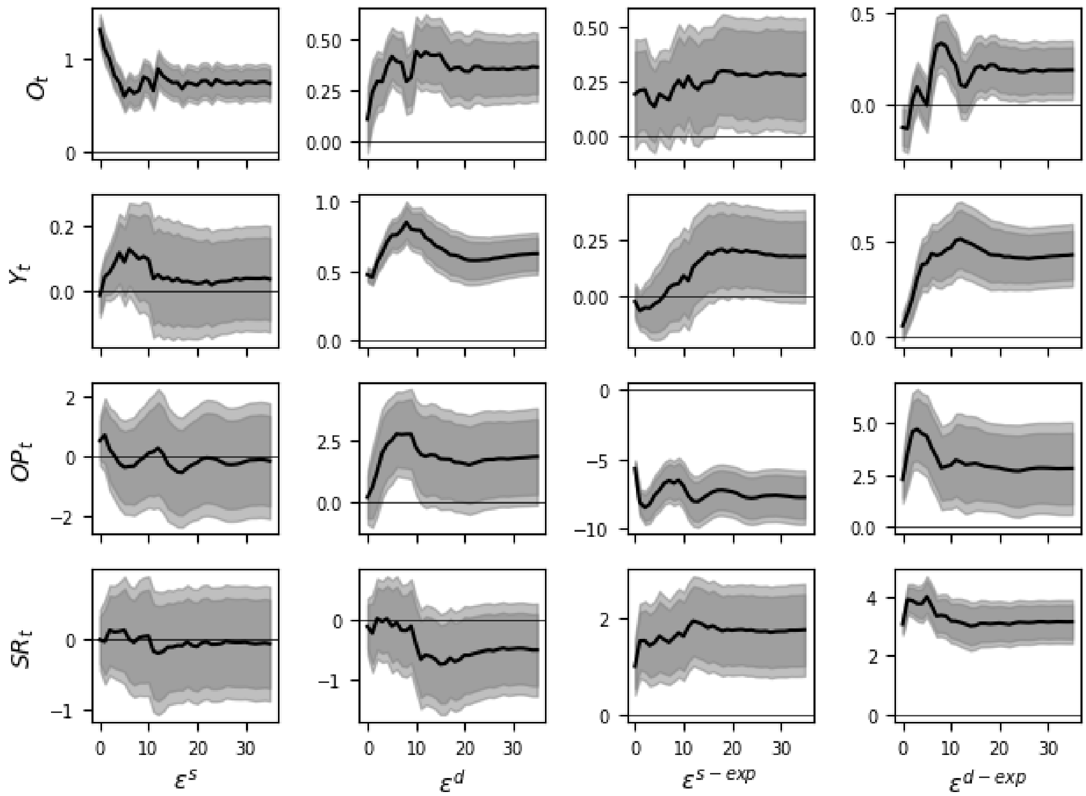
Figure B.5: Impulse Responses of the recursive SVAR GMM estimator.



Note: Impulse responses to the recursive oil market SVAR from Section 6 estimated with the GMM estimator from Equation (15), equal to the estimator obtained by applying the Cholesky decomposition to the variance-covariance matrix of the reduced form shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses. The shock ε^{op-exp} denotes a speculative oil price shock and the shock ε^{sm} represents a residual stock market shock.

Figure B.6 shows the IRF for the unrestricted oil market SVAR from Section 6 estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

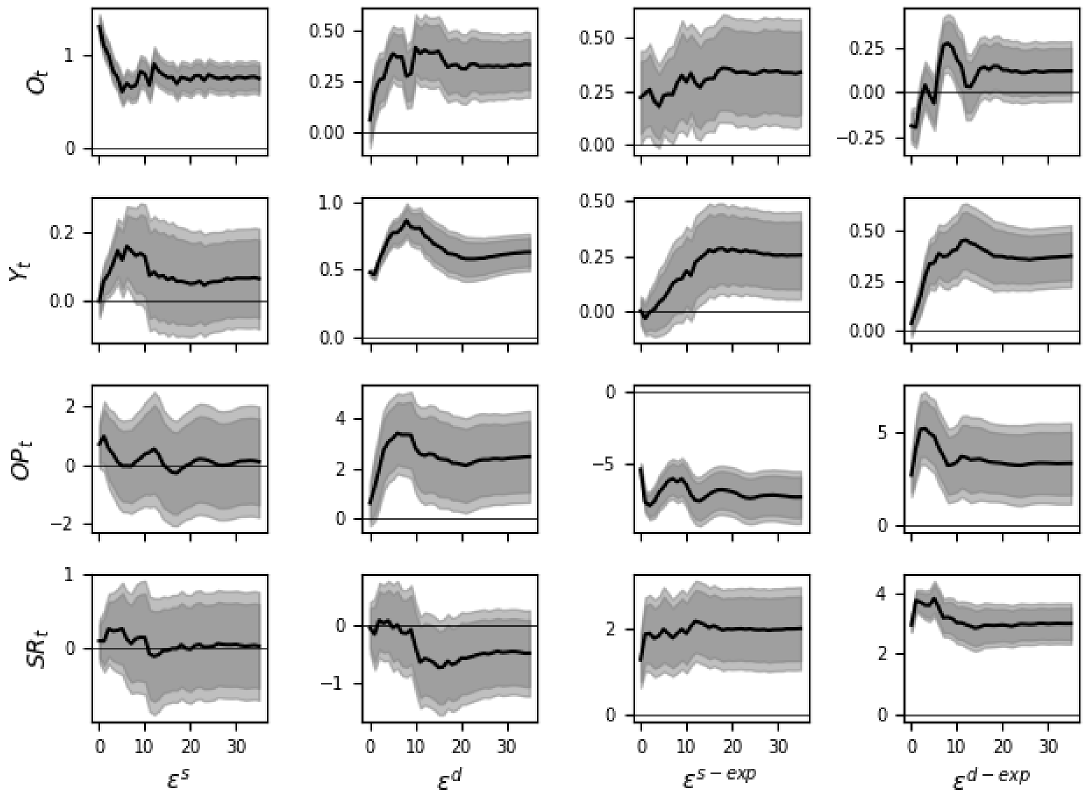
Figure B.6: Impulse Responses of the unrestricted SVAR GMM estimator.



Note: Impulse responses to the estimated structural shocks for the unrestricted oil market SVAR from Section 6 estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure B.7 shows the IRF for the unrestricted oil market SVAR from Section 6 estimated with the overidentified GMM estimator from Equation (19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

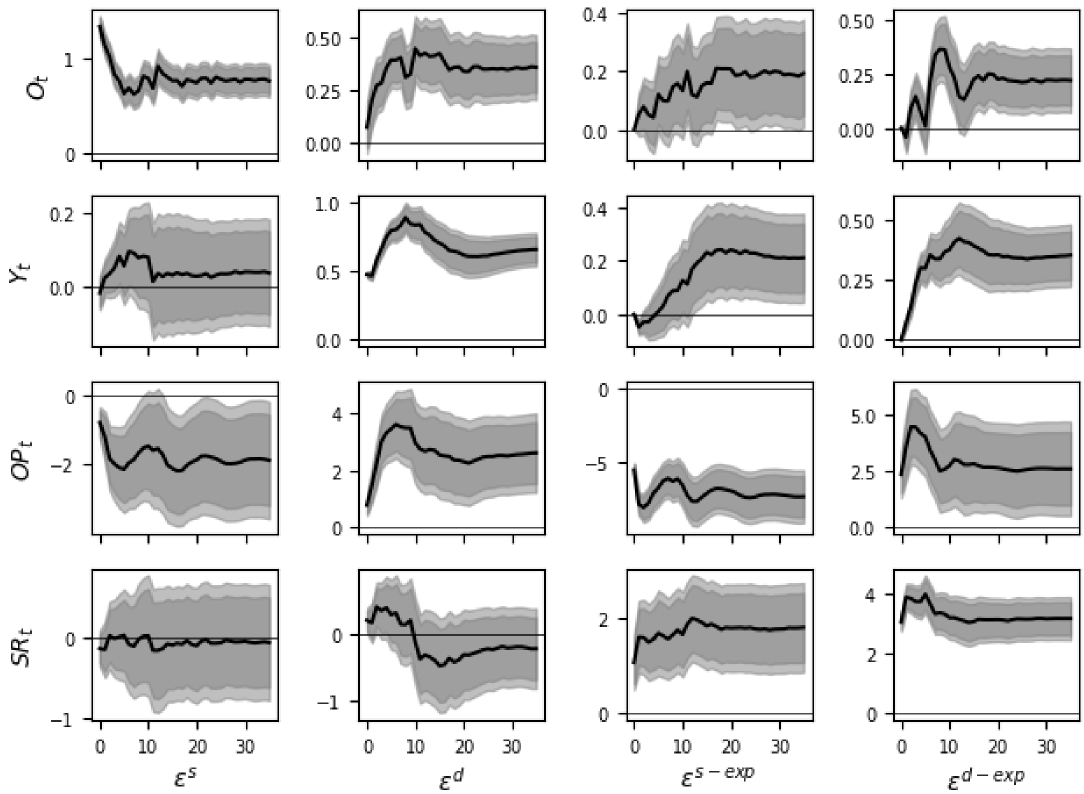
Figure B.7: Impulse Responses of the unrestricted SVAR oGMM estimator.



Note: Impulse responses to the estimated structural shocks for the unrestricted oil market SVAR from Section 6 estimated with the overidentified GMM estimator from Equation (19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure B.8 shows the IRF for the block-recursive oil market SVAR from Section 6 estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

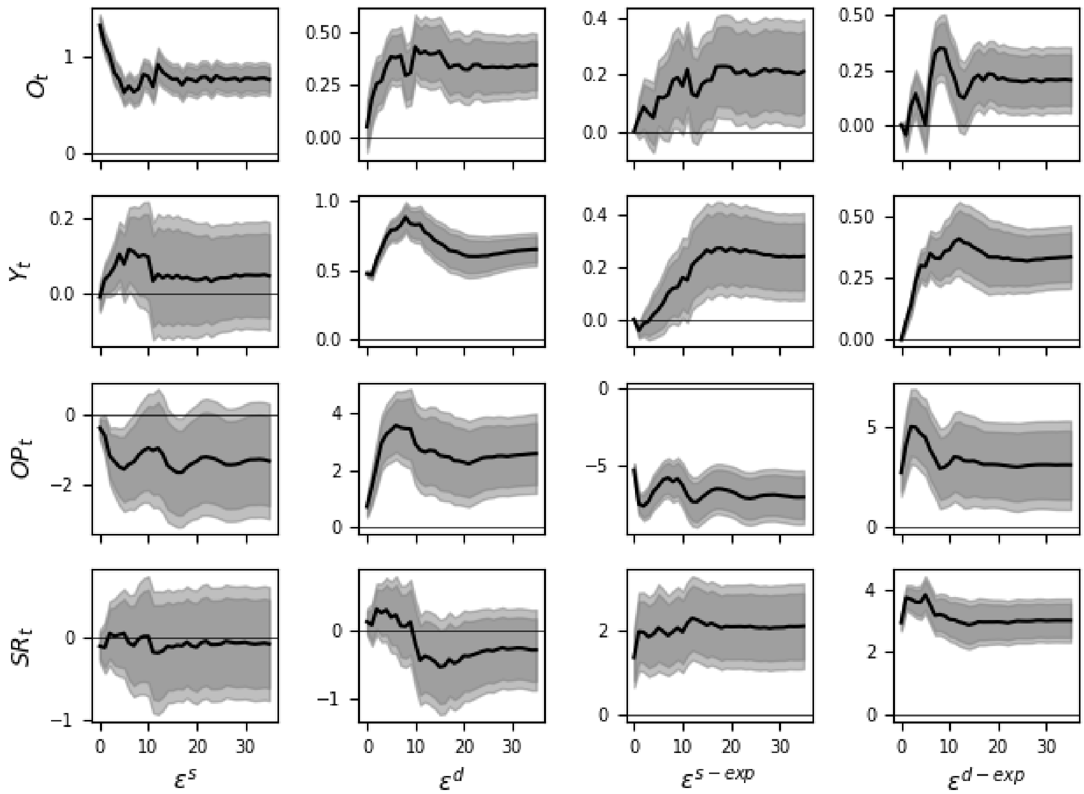
Figure B.8: Impulse Responses of the block-recursive SVAR GMM estimator.



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 6 estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure B.9 shows the IRF for the block-recursive oil market SVAR from Section 6 estimated with the overidentified GMM estimator from Equation (19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

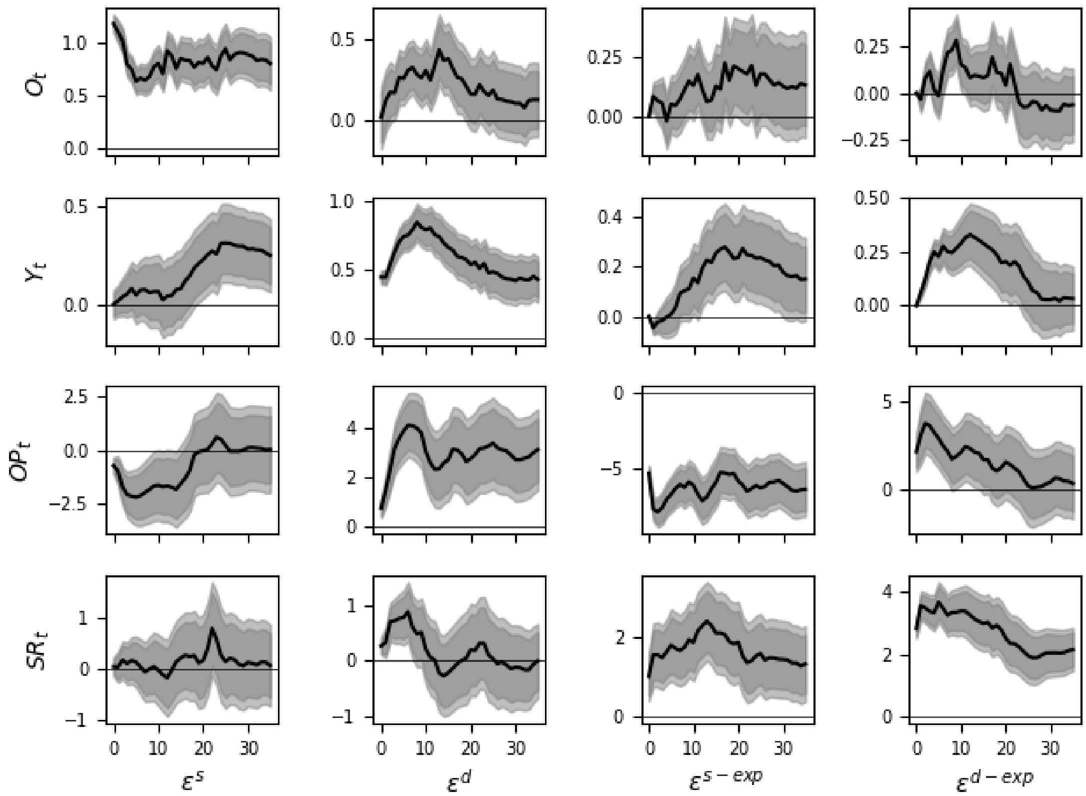
Figure B.9: Impulse Responses of the block-recursive SVAR oGMM estimator.



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 6 estimated with the overidentified GMM estimator from Equation (19) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure B.10 shows the IRF for the block-recursive oil market SVAR from Section 6 using 24 lags estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

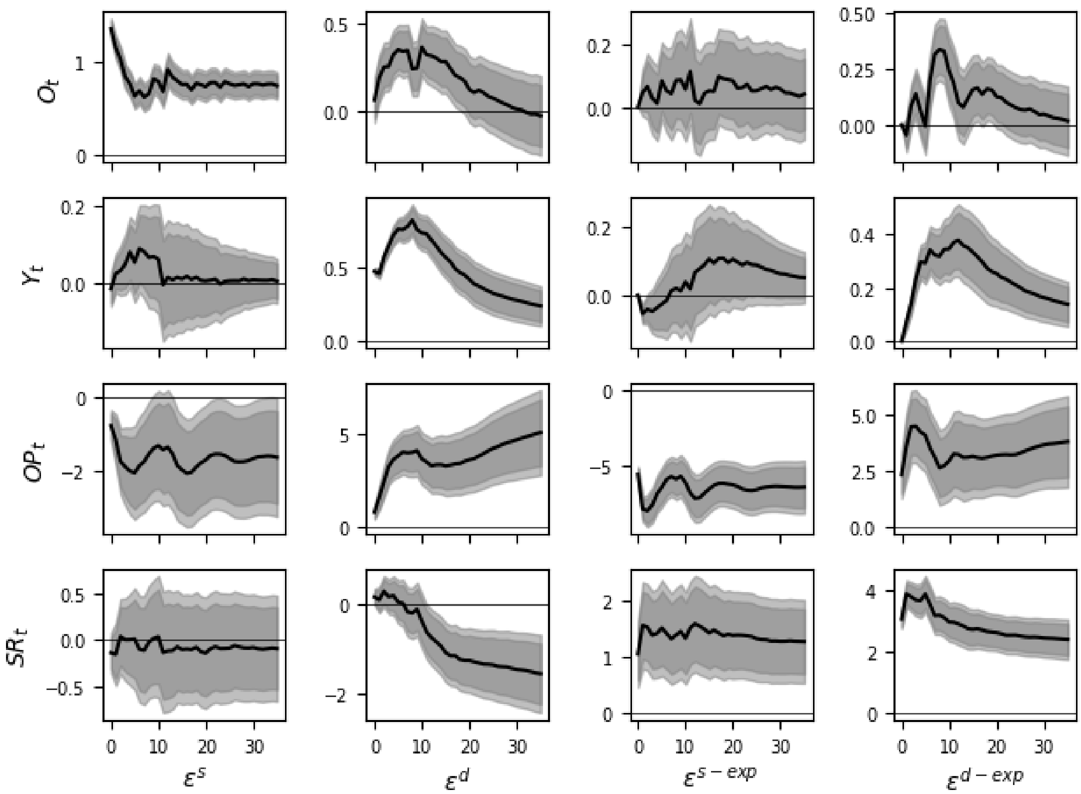
Figure B.10: Impulse Responses of the block-recursive SVAR GMM estimator using 24 instead of 12 lags.



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 6 24 estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows show the cumulative responses.

Figure B.11 shows the IRF for the block-recursive oil market SVAR from Section 6 using the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

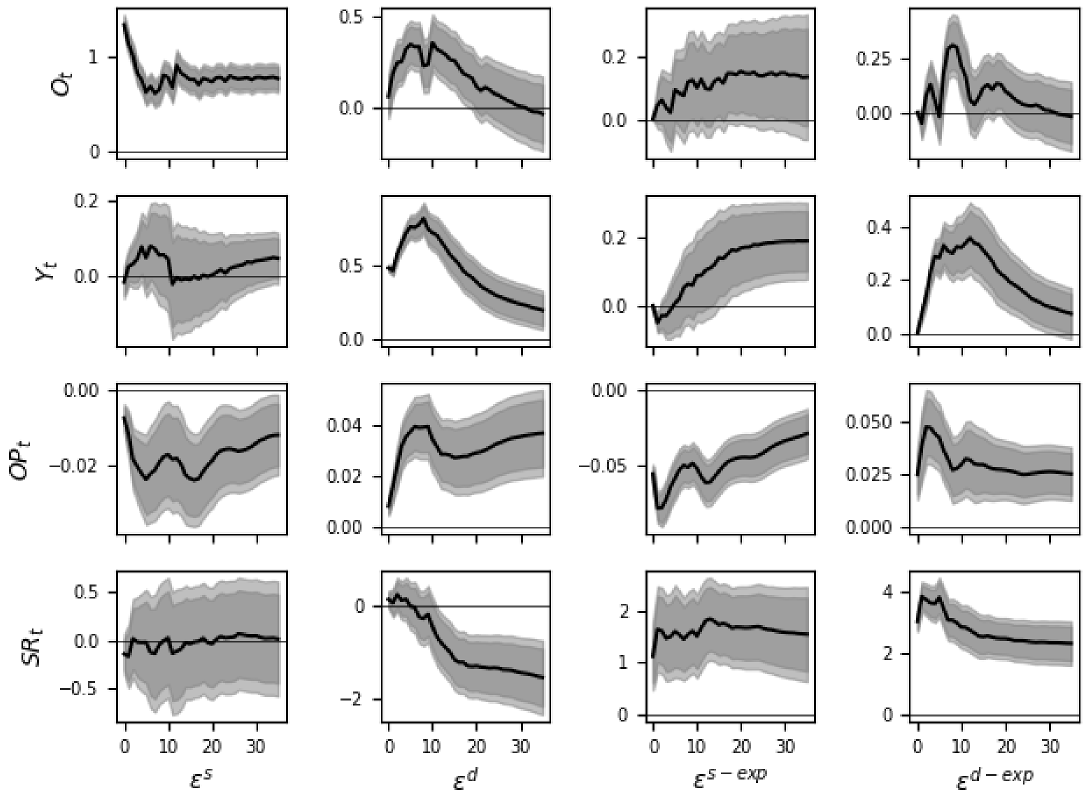
Figure B.11: Impulse Responses of the block-recursive SVAR estimator using different specification for industrial production.



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 6 using the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows O_t, OP_t, and SR_t show the cumulative responses.

Figure B.12 shows the IRF for the block-recursive oil market SVAR from Section 6 using log of real oil price instead of real oil price growth and the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

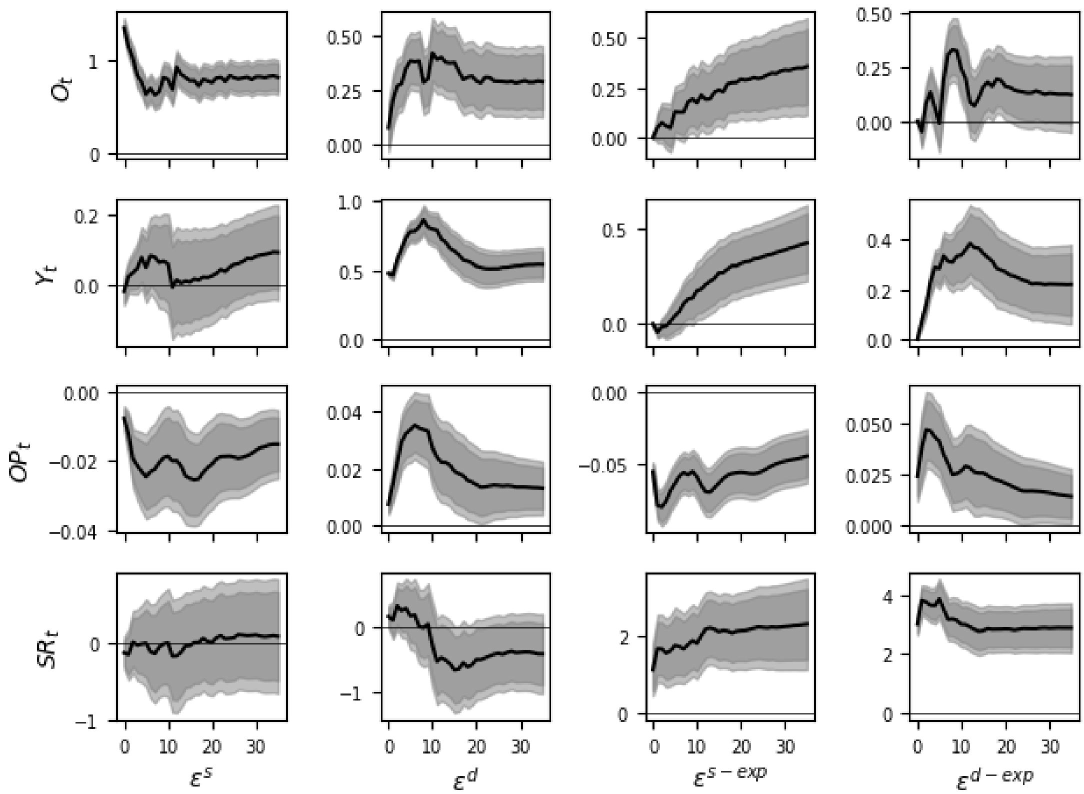
Figure B.12: Impulse Responses of the block-recursive SVAR GMM estimator using different specification for industrial production and real oil price.



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 6 using log of real oil price instead of real oil price growth and the percentage deviation of industrial production from a linear trend instead of the log difference of industrial production. The SVAR is estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows O_t and SR_t show the cumulative responses.

Figure B.13 shows the IRF for the block-recursive oil market SVAR from Section 6 using log of real oil price instead of real oil price growth. The SVAR is estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks.

Figure B.13: Impulse Responses of the block-recursive SVAR GMM estimator using different specification for real oil price.



Note: Impulse responses to the estimated structural shocks for the block-recursive oil market SVAR from Section 6 using log of real oil price instead of real oil price growth. The SVAR is estimated with the GMM estimator from Equation (15) where the weighting matrix is continuously updated and estimated based on the assumption of serially and mutually independent shocks. Confidence bands are symmetric 68% and 80% bands based on standard errors and 500 replications. The rows O_t, Y_t, and SR_t show the cumulative responses.