Identifying Proxy VARs with Restrictions on the Forecast Error Variance

Tilmann Härtl

Department of Economics, University of Konstanz

tilmann.haertl@uni-konstanz.de

*

August 19, 2022

Abstract

The proxy VAR framework requires additional restrictions to disentangle the structural shocks when multiple shocks are identified using multiple instruments. I propose to employ restrictions on the forecast error variance (FEV). Less restrictive assumptions that bound the contributions to the FEV can replace or accompany inequality restrictions on e.g. the impulse responses. This enables or sharpens the set identification of the structural parameters. Furthermore, with the correct economic intuition the Max-Share framework can be used to point identify the structural parameters without the need for strict equality restrictions in the case when two shocks are identified with two proxy variables.

*German Science Foundation (DFG) grant number BR 2941/3-1 is gratefully acknowledged.

1 Introduction

In the last decade, proxy variables were highly prevalent in the structural vector autoregression (SVAR) literature. The proxy VAR framework was developed by Stock and Watson (2012) and Mertens and Ravn (2013). So far, it was applied to identify the effects of various structural shocks. For example, Mertens and Ravn (2013) estimate the effects of taxation shocks, Gertler and Karadi (2015) the effects of monetary policy shocks and Piffer and Podstawski (2018) the impacts of uncertainty shocks. As in the standard IV identification the proxy variables (also called instruments in this context) need to satisfy two key conditions. Reminiscent of the relevance and exogeneity assumption, the external series have to be related to the target shocks of interest while being unrelated to the remaining structural shocks that are not identified.

When multiple shocks are identified with multiple instruments, Mertens and Ravn (2013) show that additional identifying restrictions are needed in order to disentangle the structural shocks. For instance, Piffer and Podstawski (2018) have an instrument related to an uncertainty shock and one instrument for a news shock. If the two instruments are not related to structural shocks other than the uncertainty and news shock, the proxy VAR successfully rules out the other shocks as confounders. Yet, in order to disentangle the two identified shocks additional restrictions are needed. Different solution were proposed in the literature. Mertens and Ravn (2013) assume a recursive structure, meaning that one of their two identified shocks has no contemporaneous impact on a specified variable. Piffer and Podstawski (2018) distinguish between uncertainty and news shocks by enforcing that each shock is correlated more strongly to the instrument targeting it.

For the latter strategy, the inequality restriction results in the set identification of the structural parameters. Yet, the set identification strategies inherit the potential to lead to rather large and uninformative sets if the identification restrictions are not sharp enough. On the contrary hard equality restrictions which sharply identify the structural parameters, as e.g. in Mertens and Ravn (2013), are typically hard to defend. The difficulties in both cases highlight the importance of economic intuition in the identification of SVARs. I propose to use restrictions on the contribution of a specific shock to the forecast error variance (FEV) of a target variable in order to disentangle the identified shocks in the proxy VAR. Depending on the available economic intuition the FEV restrictions allows both point and set identification.

Set identification of shocks based on restrictions on the contributions to the FEV were introduced by Volpicella (2021). The main idea is to bound the contributions of the target shock to the FEV of a specified variable. These bounds induce inequality restrictions on the structural model, similar to e.g. sign restrictions, and the identified set consists of all the structural representations that satisfy them. The bounds on the FEV can either replace or accompany existing inequality restrictions. Hence, the bounds offer a possibility to identify the impulse responses, which are often the key element of interest in the SVAR environment, without the need to impose restrictions on exactly the parameters of interest. Yet, if the bounds on the FEV accompany existing inequality restrictions they help to sharpen the identification and alleviate the problem of large uninformative identified sets.

Point identification of shocks via restrictions on the FEV dates back to Faust (1998) and Uhlig (2004a) and was originally an alternative to the bias prone long-run restrictions. The shock is identified to be the shock that maximizes the contribution to the FEV of a specific variable. Francis et al. (2014) coined the term 'Max-Share' for this identification strategy and in the following I use this expression to refer to it. However, Dieppe et al. (2019) point out the as soon as more than one shock contributes to the FEV of the chosen variable, the maximization of the share identifies a combination of the contributing shocks. Due to this drawback the Max-Share approach is only in specific situations applicable. In the literature it has, for instance, been used by Barsky and Sims (2011) to identify technology news shocks and by Ben Zeev and Pappa (2017) to identify defence spending news shocks.

The merit of fusing proxy VARs and the Max-Share approach is that sharp point identification of the structural parameters is achieved without the need for strict equality restrictions. Certainly, the drawback of the pure Max-Share framework carries over to the application in the proxy VARs. Hence, the Max-Share approach only correctly disentangles the underlying shocks if one of them exclusively contributes to a variable in the system. Shocks that contribute to the FEV but are not related to the proxies are cancelled out by the proxy VAR. Yet, if the condition of exclusive contribution is not fulfilled the results will be biased. Thus, I propose to augment the Max-Share framework with an inequality restriction to disentangle the shocks while removing or reducing the bias. This strategy is limited to the case when two shocks are identified with two instruments. However, in practice this highly relevant as finding two suitable proxy variables is difficult enough. The inequality restriction is of the form that the contemporaneous impact of shock one on a specified variable is larger than the impact of shock two on the same variable. If known, one can also incorporate the margin by which the response to the one shock exceeds the response of the other shock into the inequality constraint to reduce the bias further.

The simulation study shows that the Max-Share approach successfully disentangles the shocks in the proxy VAR. In the case of exclusive contributions the basic Max-Share identification is sufficient to disentangle the structural shocks. If the basic Max-Share framework is biased the augmentation with the mentioned inequality constraint reduces or removes the bias depending on the specification of the additional constraint. The closer the inequality constraint to the actual margin of the restricted contemporaneous responses the closer the estimate to the true underlying structural parameters.

Compared to the Max-Share framework the bounds offer a more flexible approach with less restrictive assumptions and it can be easily combined with other types of restrictions. Yet, the benefits come with the cost of loosing the sharp point identification. Nevertheless, bounding the FEV is a useful identification strategy whose assumptions can be backed with economic theory and intuition. I present an empirical illustration which highlights one useful application of the bound constraints. I weaken the strict identification assumption by Mertens and Ravn (2013) and show that their results are, although qualitatively not considerably different, statistically less convincing.

Section two commences with the introduction of the baseline SVAR framework and the introduction of the proxy VAR. Section three describes the usage of restrictions on the FEV for the identification of the structural VAR. It starts with the more general set identification approach via the bounds on the FEV and ends with the more strict point identification via the Max-Share framework. Section four and five present the results of the simulation study and of the empirical illustration, respectively.

2 Econometric Framework

2.1 The Structural VAR

The starting point is the k dimensional stationary structural VAR(p) model:

$$y_t = \sum_{m=0}^{p} A_m y_{t-l} + B w_t, \quad t = 1, ...T, \tag{1}$$

where the $k \times 1$ vector w_t depicts the economically meaningful structural shocks (e.g. Kilian & Lütkepohl, 2017). The $k \times k$ impact matrix B maps the reduced form innovations into the structural shock, $u_t = B w_t$. The elements of the $k \times 1$ white noise vector u_t are the reduced form innovations. To shorten the notation one can rewrite the SVAR in (1) as:

$$y_t = A x_t + B w_t, \quad t = 1, ...T, \tag{2}$$

with $x_t = (y'_{t-1}, ..., y'_{t-p})'$ and $A = (A_1, ..., A_p)$. A constant is omitted for the brevity of the notation but it can be included in a straightforward way.

Common to most identification strategies is the normalization $\mathbb{E}(w_t w_t') \equiv \Sigma_w = I_K$, which yields the set of covariance restriction $\mathbb{E}(u_t u_t') \equiv \Sigma = BB'$. Without further assumptions these restrictions do not suffice to pin down the structural parameters as the resulting system of equation has many possible solutions. The Cholesky decomposition of Σ, denoted by Σ_c, satisfies these covariance restrictions. Yet, they will also hold for every rotation with an $k \times k$ orthonormal matrix Q, $\Sigma = \Sigma_c \Sigma_c' = \Sigma_c Q Q' \Sigma_c'$. Giacomini, Kitagawa, and Read (2021) refer to this representation of the SVAR as the 'orthogonal reduced form'.

The moving average representation of the SVAR is then given by:

$$y_t = \sum_{m=0}^{\infty} C_m \Sigma_c Q w_{t-m}, \quad t = 1, ...T, \tag{3}$$

where the $k \times k$ matrices C_m contain the moving average coefficients which give the response of the system to the reduced form innovations m periods ago. The impulse response of variable i to shock j at horizon h is given by:

$$\eta_{i,j,h} = e_i' C_m \Sigma_c q_j, \tag{4}$$

where e_i and e_j are the ith and jth column of I_k, respectively, and q_j is the jth column of Q.

Apart from the impulse response functions the FEV decomposition is are an element of interest in the SVARs. In this paper the FEV decomposition is particularly of importance as identifying restrictions are placed on it. To formalize the FEV decomposition let the h-step-ahead forecast of y_t be:

$$y_{t+h|t} = \sum_{m=0}^{\infty} C_{h-m} u_{t-m}. \tag{5}$$

The h-step ahead forecast error is then given by:

$$y_{t+h} - y_{t+h|t} = \sum_{m=0}^{h-1} C_m \Sigma_c Q w_{t+h-m}, \tag{6}$$

and the h-step ahead forecast error covariance matrix is represented by:

$$\Omega(h) = \sum_{m=0}^{h-1} C_m \Sigma_c Q Q' \Sigma_c' C_m' = \sum_{m=0}^{h-1} C_m \Sigma C_m'. \tag{7}$$

The contribution from shock j to the total forecast variance of variable i at horizon

h is then:

$$\Omega_{i,j}(h) = \frac{e_i'(\sum_{m=0}^{h-1} C_m \Sigma_c q_j q_j' \Sigma_c' C_m')e_i}{e_i'(\sum_{m=0}^{h-1} C_m \Sigma C_m')e_i}, \tag{8}$$

where e_i is the ith column of the identity matrix I_k and q_j is the jth column of Q.

In order to identify the structural parameters of interest restrictions have to be placed on the model. As mentioned before two different identification schemes exist. Set identification amounts to finding all the rotation matrices Q that satisfy the identification restrictions. In turn, they define the identified set for e.g. the impulse response functions or the FEV decomposition. Common set identification restrictions are, for instance, inequality restriction on the structural impulse responses. The stricter the identifying restrictions, the smaller the identified set. In point identification schemes the restrictions are such that only one admissible rotation matrix Q exists. The typical point identification restrictions are equality restrictions on the elements of the impact matrix B. For example, $k(k-1)/2$ independent equality restrictions on B are sufficient to point identify the structural parameters.

The proxy VAR framework that is introduced in the next subsection allows both for point and set identification. If one shock is identified using one proxy variable the parameters are point identified up to scale. With multiple shocks and multiple proxies additional restrictions are needed in order to disentangle the shocks. In the latter case the just mentioned sign or equality restrictions are one possibility and depending on the type of imposed restrictions the structural parameters are either point or set identified.

2.2 Proxy VAR

For the the proxy VAR framework I loosely follow the framework by Giacomini et al. (2021) as the inference for the set identification part will be based on their work. As in the standard instrumental variable (IV) framework, the proxy variables - also called instruments interchangeably - have to satisfy two key assumptions. Without loss of generality, let z_t be a $l \times 1$ vector of instruments that are related to the first l structural shocks in w_t. The two following two conditions have to be satisfied:

$$\mathbb{E}(z_t w_{(1:l),t}) = \Psi \quad \text{and} \quad \mathbb{E}(z_t w_{(k-l+1:k),t}) = 0, \tag{9}$$

where Ψ is an $l \times l$ matrix of full rank. These two conditions resemble the relevance and exogeneity conditions of the standard IV approach. The instruments have to be related to the target shocks and unrelated to the remaining structural shocks.

Assume that the proxies follow:

$$\Gamma_0 z_t = \Lambda w_t + \sum_{l=0}^{p_m} \Gamma_m z_{t-l} + \nu_t, \quad t = 1, ..., T. \tag{10}$$

The process in (10) indicates that the proxies are related to the structural shocks. Giacomini et al. (2021) assume that $(w_t', \nu_t)'|\mathcal{F}_{t-1} \sim N(0_{(k+l)\times l}, I_{k+l})$, where $\mathcal{F}_{t-1}$ is the information set at time $t-1$.

The assumptions in (9) together with process (10) yield:

$$\mathbb{E}(z_t w_t') = \Gamma_0^{-1}\Lambda = [\Psi, 0_{l\times(k-l)}]. \tag{11}$$

Plugging model (2) into the process in (10) and left-multiplying by Γ_0^{-1} yields:

$$z_t = Dy_t + Gx_t + \sum_{m=0}^{p_z} H_m z_{t-m} + v_t, \quad t = 1, ..., T, \tag{12}$$

where $D = \Gamma_0^{-1}\Lambda B^{-1}, G = -\Gamma_0^{-1}\Lambda A$ and $H_m = \Gamma_0^{-1}\Gamma_l$ for each $m = 1, ..., p_z$.

Giacomini et al. (2021) show that (11) can also be represented by:

$$\mathbb{E}(z_t w_t') = D\Sigma_c Q = [\Psi, 0_{l\times(k-l)}], \tag{13}$$

implying that the relevance assumption rank$(\Psi) = l$ is fulfilled if and only if rank$(D) = l$. The exogeneity and relevance assumption regarding the proxies restricts the rotation matrices Q such that they follow the structure in (13). In this fashion the proxy VAR shrinks the identified set.

In the following I deviate from the notation of Giacomini et al. (2021) and use the proxy VAR framework by Piffer and Podstawski (2018). This allows me to handle both the bounds on the FEVD and the combination with Max-Share approach in the same proxy VAR framework. The next section describes how the robust bayesian inference algorithm of Giacomini et al. (2021) is adapted to the representation of the proxy SVAR below.

Following Piffer and Podstawski (2018), I decompose the reduced form errors into two components:

$$u_t = B_z w_{(1:l),t} + B_{-z} w_{(k-l+1:k),t}, \quad t = 1, ..., T, \tag{14}$$

where B_z is the $k\times l$ block of the impact matrix B that contains the first l colums and B_{-z} is the according remaining part of B. B_z contains the structural parameters of the shocks related to the proxies whose identification is the goal of the proxy VAR. I refer to this matrix as the 'proxy impact matrix'. Considering (14) together with the assumption regarding the proxies in (9) yields $\mathbb{E}(u_t z_t') = B_z \Psi' = Z$. Equation (13) lets me rewrite this expected value:

$$\mathbb{E}(u_t z_t') = B\mathbb{E}(w_t z_t') = B\mathbb{E}(z_t w_t')' = \Sigma D', \tag{15}$$

with $B = \Sigma_c Q$ and $\mathbb{E}(z_t w_t') = D\Sigma_c Q$. Hence, $\Sigma D' = Z = B_z \Psi'$. Partitioning the matrix $\Sigma D'$ and B_z yields:

$$\Sigma D' = Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \ \ and \ \ B_z = \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} \tag{16}$$

and thus

$$B_{21} = Z_2 Z_1^{-1} B_{11} = Z_l B_{11}, \tag{17}$$

where Z_1 is the upper $l \times l$ block of the matrix Z and B_{11} is the upper $l \times l$ block of B_z. Hence,

$$B_z = \begin{pmatrix} B_{11} \\ Z_l B_{11} \end{pmatrix}, \tag{18}$$

and if the upper $l \times l$ block B_{11} is identified the remaining block of the proxy impact matrix is identified as well. In order to identify the upper block of B_z decompose the matrices of the standard covariance restrictions $\mathbb{E}(u_t u_t') \equiv \Sigma = BB'$ such that:

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{21} \\ B_{12} & B_{22} \end{pmatrix}, \tag{19}$$

where Σ_{11} is the upper left $l \times l$ block of Σ. B_{11} is again the upper $l \times l$ block of B^z and therefore the upper left block of B. The remaining blocks of the two matrices have the according dimensions. It can be shown that $B_{11}B_{11}' = \Sigma_{11} - B_{12}B_{12}'$ (see Piffer & Podstawski, 2018) with:

$$B_{12}B_{12}' = (\Sigma_{21} - Z\Sigma_{11})'\Pi^{-1}(\Sigma_{21} - Z\Sigma_{11}), \tag{20}$$

$$\Pi = \Sigma_{22} + Z'\Sigma_{11}Z' - \Sigma_{21}Z' - Z\Sigma_{21}'. \tag{21}$$

Similar to the covariance restrictions $\mathbb{E}(u_t u_t') \equiv \Sigma = BB'$, the equation $B_{11} = \Sigma_{11} - B_{12}B_{12}'$ does not pin down the parameters of B_{11} uniquely. Let B_{11}^c be the Cholesky decomposition of $\Sigma_{11} - B_{12}B_{12}'$, then every rotation of B_{11}^c with an $l \times l$ orthonormal matrix Q will also satisfy $B_{11} = B_{11}^c B_{11}^{c'} = B_{11}^c QQ' B_{11}^{c'} = \Sigma_{11} - B_{12}B_{12}'$.

The exogeneity and relevance restriction regarding the proxies is satisfied for B_z by construction. Hence, the identification boils down to finding the set of $l \times l$ orthonormal matrices Q that satisfy the additional identifying restrictions, e.g. inequality restriction on structural parameters. In the next section I describe how inequality restrictions on the FEV like in Volpicella (2021) fit into this framework.

It is also possible to point identify the structural parameters, what again comes down to finding the one rotation matrix for which the restrictions are satisfied. If the resulting recursive structure of B_{11}^c for the contemporaneous impacts of the identified shocks is economically justifiable, the Cholesky decomposition immediately point identifies the structural shocks. This is, for instance, the identification assumption used in Mertens and Ravn (2013) and the corresponding rotation matrix is just $Q = I_l$.

A special case arises when $l = 2$, meaning that two shocks are identified with two instrument. In the proxy VAR the rotation matrix will be of dimension 2×2. In this case, knowing one column of the rotation matrix also gives the second column of the rotation matrix Q up to a sign normalization. If the first column of a 2×2 orthogonal matrix is known the second column is pinned down up to sign through following equations:

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}, \qquad 1 = q_{21}^2 + q_{22}^2 \quad \text{and} \quad 1 = q_{11}^2 + q_{12}^2.$$

Hence, restrictions on one of the two shocks are sufficient to identify both shocks of interest. The second shock is pinned down due to the properties of orthogonal matrices. I make use of this special case in combination of the proxy VAR with the Max-Share framework where I point identify two shocks with restrictions on only one of the two shocks.

To avoid confusion, in the following sections every rotation matrix is of dimension $l \times l$. For the parts describing identification of the proxy VAR with the Max-Share approach $l = 2$.

3 Proxy VAR with Restrictions on the FEV

3.1 Proxy VARs with Bounds on the FEV

The bounds on the contributions to the FEV where introduced by Volpicella (2021) and this section applies them to the proxy SVAR framework. In doing so, I loosely follow the notation of Volpicella (2021). Such bounds on the FEV are inequality restrictions in the spirit of the well known sign restrictions on impulse response parameter, an thus the structural parameters are set identified. Naturally, the challenges the set identification literature deals with also apply to this identification scheme.

3.1.1 Bounding the contribution to the FEV

In the proxy VAR framework the contribution of shock j to the FEV of variable i at horizon h is:

$$\Omega_{ij}(h) = \frac{e_i'(\sum_{m=0}^{h-1} C_m B_z q_j q_j' B_z' C_m')e_i}{e_i'(\sum_{m=0}^{h-1} C_m \Sigma C_m')e_i}. \tag{22}$$

Uhlig (2004b) shows that equation (22) can also be written as:

$$\Omega_{i,j}(h) = q_j' R_{i,h} q_j, \tag{23}$$

where

$$R_{i,h} = \frac{\sum_{m=0}^{h-1} c_{i,m}' c_{i,m}}{e_i'(\sum_{m=0}^{h-1} C_m \Sigma C_m')e_i}, \tag{24}$$

with $c_{i,m} = e_i C_m B_z$ is the ith row vector of $C_m B_z$. $R_{i,h}$ is a positive semidefinite and symmetric $l \times l$ real matrix.

Given equation (23) the bounds on the contribution to the FEV of variable i by shock j at horizon h can be represented by:

$$\underline{\tau}_{i,j,h} \leq q_j' R_{i,h} q_j \leq \overline{\tau}_{i,j,h},$$

where $\underline{\tau}_{i,j,h}$ and $\overline{\tau}_{i,j,h}$ depict the lower and upper bound, respectively, and $0 \leq \underline{\tau}_{i,j,h} \leq \overline{\tau}_{i,j,h} \leq 1$. Let $\mathcal{I}_j$ be a set of indices that depict whether the FEV of variable i is bounded and $\mathcal{H}_{ij}$ collects the horizons $h = 0, 1, ...$ for which these bounds are imposed. The whole set of bound constraints is then characterized by

$$\underline{\tau}_{i,j,h} \leq q_j' R_{i,h} q_j \leq \overline{\tau}_{i,j,h}, \text{ for } i \in \mathcal{I}_j \text{ and } h \in \mathcal{H}_{ij}.$$

These bounds on the contributions to the FEV can also be applied together with already existing set identifying inequality restrictions, like e.g. sign restrictions. Furthermore, restrictions on the correlations of the proxies with the identified shocks are possible. These type of restrictions constrain the elements of Ψ. They can be checked employing the routine used by Piffer and Podstawski (2018). The identified set is then characterized by all the rotation matrices Q for which these FEV bounds and other potential restrictions are satisfied. As pointed out by Volpicella (2021) such bounds on the FEV contributions can be derived either through economic theory or simply by strong beliefs due to economic intuition.

3.1.2 Nonemptiness of the Set

The bound restrictions, together with potential additional restrictions, are subject to set-identification specific considerations. On the one hand, if the bounds are not restrictive enough one gets potentially large identified sets which yield a fuzzy identification of the underlying structural effects. If, on the other hand, the bounds are to restrictive the identified set might be empty because no structural representation of the model satisfies them.

Unfortunately, no formal guidance helps to assess the restrictions in this regard, what in turn highlights the importance of the economic theory or intuition behind them. Yet, if the identified set is empty, this might be a sign that the imposed restrictions are not reasonable.

Furthermore, it is important to know whether the set is empty for the estimation procedure. Volpicella (2021) provides sufficient conditions for the nonemptiness of the identified set when only one shock is restricted. These sufficient conditions also apply in the same fashion to the proxy VAR framework. Recall that the contribution of the target shock j to the FEV of variable i is:

$$\Omega_{ij}(h) = q_j' R_{i,h} q_j, \tag{25}$$

Let λ_m^{ih} the real eigenvalues of $R_{i,j}$ with $i \in \mathcal{I}_j$, $h \in \mathcal{H}_{ij}$ and $m = 1, ..., l$. Uhlig (2004b) shows that finding the maximum (minimum) of (25) with respect to q_j amounts to finding the largest (smallest) eigenvalue λ_m^{ih} of $R_{i,j}$ and the maximum (minimum) is achieved by using the corresponding eigenvector q_m as a rotation vector q_j. Hence, the eigenvalues λ_m^{ih} correspond to the contributions to the FEV of variable i at horizon h.

Proposition 3.1 follows from Proposition 3.1 of Volpicella (2021) and gives sufficient conditions for the nonemptiness of the identified set when a single target shock j is restricted.

Proposition 3.1. (Nonemptiness) If the following conditions hold:

(a) $\exists i \in \mathcal{I}_j, \exists h \in \mathcal{H}_{ij} \mid \underline{\tau}_{i,j,h} \leq \lambda_m^{ih} \leq \overline{\tau}_{i,j,h},\ Rq_m = \lambda_m^{ih} q_m$ for some $m = 1, ..., l$,

(b) given q_m from (a), $\underline{\tau}_{i,j,h} \leq q_m' R_{i,h} q_m \leq \overline{\tau}_{i,j,h} \forall i \in \mathcal{I}_j$ and $\forall h \in \mathcal{H}_{ij}$, and all other additional restrictions are satisfied,

then the identified set is non-empty and bounded.

If the contribution to the FEV of a single variable i is bounded the condition reduces to a simple check whether one of the eigenvalues λ_m^{ih} lies within the bounds. When additional restrictions, like sign restrictions, are imposed one also has to check whether they are satisfied for $q_j = q_m$.

Apart from Volpicella (2021) Proposition 3.1 not only applies when only one shock is constrained. In the special case of $l = 2$, Proposition 3.1 also helps to detect nonemptiness also when both shocks are subject to identification restrictions. Knowing q_m amounts to knowing the whole 2×2 rotation matrix Q_m due to the properties of orthonormal matrices. Hence, in step (b) of Proposition 3.1 also the restrictions on the second shock can be checked.

Another difference to Volpicella (2021) is the number of eigenvalues that are available for assessment of the nonemptiness. In the proxy VAR only l eigenvalues are at hand compared to the k eigenvalues in Volpicella (2021). In practice the sufficient conditions will be met more frequently compared to the case with only l eigenvalues. On top of that, the largest and smallest eigenvalue in the proxy VAR case represents the maximum and minimum contribution to the FEV at the specific horizon. Hence, if $l = 2$ only the maximum and minimum contribution can be used to check the nonemptiness. If the bounds do not encompass the extreme values the sufficient conditions are not satisfied. Yet, in practice it might be interesting to set the bounds on the FEVD such that they are close to the maximum or minimum and the empirical illustration highlights this case.

If the sufficient conditions of Proposition 3.1 are not fulfilled a different approach helps to approximate the nonemptiness of the identified set. As often done in the literature, one can draw a specified number of matrices Q from the orthonormal space. If none of this draws satisfies the restrictions one can conclude that the set is empty.

Lastly, one can see that the set is empty if the upper bound $\overline{\tau}_{i,j,h}$ is smaller than the minimum eigenvalue for one of the imposed bounds, or if the lower bound $\underline{\tau}_{i,j,h}$ is larger than the maximum eigenvalue. In these cases the just mentioned alternative is obsolete.

3.1.3 Estimation and Inference

This subsection introduces the robust bayesian inference framework by Giacomini et al. (2021). The benefit of this approach is that it avoids specifying a prior over the rotation matrices Q that is not updated by the data. Baumeister and Hamilton (2015) show that when the prior for the rotation matrices Q is a uniform distribution over that space of orthonormal matrices, the common approach in the literature, the structural parameters are influenced by the prior distribution even asymptotically.

The remedy is to use a distribution-free approach. In the robust bayesian inference the endpoints of the identified set are calculated numerically or analytically if possible. The endpoints, or boundaries, of the identified set are the maximum and minimum values of the structural parameters of interest given all admissible rotation matrices Q. Giacomini et al. (2021) show that this procedure yields prior robust inference over the structural parameters.

As mentioned previously, I adapt the algorithm used in Giacomini et al. (2021) to the specification of the proxy VAR framework in the previous section. I incorporate the exogeneity and relevance restriction regarding the proxies via the proxy impact matrix B_z and only rotate the upper block B_{11} with an orthonormal matrix Q. Apart from the benefit that I can fit bounding the FEVD and the Max-Share approach in the same proxy VAR framework, two additional advantages arise. One merit is that I do not need to draw the rotation matrices Q subject to the exogeneity restriction as depicted in (13). For large iteration counts of the inference algorithm this potentially saves some computation time. Further, I avoid a specific ordering of the variables in the VAR. The ordering convention defined in Giacomini et al. (2021) might be difficult to incorporate in practice when it is not obvious which structural shock is linked to which variable in the VAR system.

To describe the bayesian algorithm let $\phi \in \Phi$ collect all the reduced form parameters in (2) and (12). For the following algorithm it is not important which prior for ϕ is used as long as one is capable to draw from the posterior distribution of the reduced form parameters. For the results derived in the next sections I follow Giacomini et al. (2021) and use an (improper) Jeffrey's prior.

Suppose that the impulse responses $\eta_{i,j,h} = e_i' C_m B_z q_j$ are the structural parameters of interest. The upper and lower boundary of the identified set with respect to the imposed restrictions are depicted by $u_{i,h}(\phi)$ and $l_{i,h}(\phi)$. Algorithm 1 describes how to conduct robust bayesian inference for the identified set of $\eta_{i,j,h}$.

Algorithm 1.

Step 1: Obtain draws ϕ from its posterior distribution and compute B_{11}^c.

Step 2: Check whether the identified set is empty. If the set is empty go back to Step 1. If the set is non-empty proceed with Step 3.

Step 3: Compute the boundaries of the identified set:

$$\begin{aligned}
l_{i,h}(\phi) = & \min_{Q} e_i' C_m B_z q_j \\
s.t \quad & \underline{\tau}_{i,j,h} \leq q_j' R_{i,h}(\phi) q_j \leq \overline{\tau}_{i,j,h}, \quad \forall i \in \mathcal{I}_j \text{ and } \forall h \in \mathcal{H}_{ij}, \\
& QQ' = I_l, \\
& \text{potential sign restrictions and/or restrictions on } \Psi.
\end{aligned}$$

The upper boundary $u_{i,h}(\phi)$ is obtained analogously.

Step 4: Repeat Steps 2 and 3 N times.

Step 5: Approximate the set of posterior means and the robust credible region as described in Giacomini et al. (2021).

Especially, Step 2 differs from the algorithm of Giacomini et al. (2021) as Proposition 3.1 helps to gauge whether the identified set is empty. If the sufficient conditions of Proposition 3.1 are not fulfilled a specified number of rotation matrices Q are drawn to approximate the set as being empty if none of the draws satisfies the identification restrictions. In this case, Step 2 differs from Giacomini et al. (2021) as the $l \times l$ rotation matrices Q do not need to be drawn considering the exogeneity conditions for the proxies. These conditions are already incorporated in the construction of B_z. Step 3 differs in the specification of the proxy VAR, and thus the dimension of the rotation matrix Q. Second, the added constraint in the maximization problem that represents the restrictions on the FEV are a distinction to the algorithm in Giacomini et al. (2021).

Step 3 poses a nonconvex optimization problem. Hence, the typical approaches to handle with gradient based optimization techniques are necessary. The simple remedy is to use different initial values and to compute the maximum or minimum over the set of solutions which are derived with the different initial values.

Giacomini et al. (2021) also provide an algorithm to approximate the boundaries of the identified set in order to check the convergence of the numerical optimization or simply as an alternative.

Algorithm 2. Replace Step 3 of Algortihm 1 with:

Step 3: Draw Q until N draws hat satisfy the identification restrictions are reached. For each $Q_n,\ 1, ..., N$ compute $\eta_{n,i,j,h} = e_i' C_m B_z q_{n,j}$ and approximate $u_{i,h}(\phi)$ and $l_{i,h}(\phi)$ by the maximum and minimum of $\eta_{n,i,j,h}$ over all N draws.

Montiel Olea and Nesbit (2021) show that the random sampling approximation of Algorithm 2 can be represented as a supervised learning problem. They provide the number of admissible draws of Q that are needed to learn the set with a certain precision. Generally, the approximated set will be smaller than the true set, yet with a sufficient amount of draws the approximation error will be small. The theoretical results of Montiel Olea and Nesbit (2021) can be used to judge the precision of the approximation at a certain amount of draws N.

Giacomini et al. (2021) argue that this approximation might be favourable under certain circumstances. Firstly, if the VAR system is large and one is interested in the impulse responses for many variables at many horizons. Drawing many rotation matrices Q is computationally less costly than optimizing for every variable at every horizon. This is especially true with the representation of the proxy VAR used in this paper as it is quite easy to draw simple $l \times l$ rotation matrices considering that l is small in most empirical applications. Second, if not only the impulse responses but also e.g. the FEV decomposition is of interest the approximation has an advantage. Each draw of Q can be used to compute the impulse responses and FEV decomposition of each variable at every horizon, while the optimization has to be carried out for each parameter, variable and horizon individually.

3.2 Proxy VAR and Max-Share

This sections describes how the shocks in the proxy VAR can be disentangled using the Max-Share framework that was introduced by Faust (1998) and Uhlig (2004b). The key assumption behind the Max-Share approach is, that the shock of interest j is the one that maximizes the contribution to the FEV of a target variable i over the considered horizon $\underline{h}$ up to $\overline{h}$. Recall, that the contributions to the FEV in the proxy VAR are given by:

$$\Omega^z_{i,j}(h) = \frac{e_i'(\sum_{m=0}^{h-1} C_m B_z q_j q_j' B_z' C_m')e_i}{e_i'(\sum_{m=0}^{h-1} C_m \Sigma C_m')e_i}, \tag{26}$$

where q_j is the jth column of the $l \times l$ orthonormal matrix Q. Then identification of the target shock amounts to finding the rotation vector q_j for which this maximum is achieved:

$$q_j^* = \text{argmax} \sum_{h=\underline{h}}^{\overline{h}} \Omega^z_{i,j}(h) \quad \text{s.t.} \quad q_j' q_j = 1. \tag{27}$$

The closed form solution of the baseline Max-Share maximization problem via the eigenvalues presented by Uhlig (2004b) also applies to the proxy VAR case in (27).

In the following I only consider the case when two shocks are identified with two instruments. In practice the case of to instruments is highly relevant, as finding multiple convincing proxy variables is difficult and finding two of them is already a challenging task. If two shocks are identified using two instruments the rotations matrix Q has dimension 2×2. Then identifying the first column of a 2×2 orthogonal matrix also identifies the second column up to a sign normalization due to the properties of orthonormal matrices:

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}, \qquad 1 = q_{21}^2 + q_{22}^2 \quad \text{and} \quad 1 = q_{11}^2 + q_{12}^2. \tag{28}$$

The conditions of the 2×2 orthogonal matrices imply that each element of Q is $-1 \leq q_{ij} \geq 1$. Further they imply that $q_{21} = \sqrt{1 - q_{11}^2}$, $q_{22} = \pm q_{11}$ and $q_{12} = \pm q_{21}$. Hence, everything can be written in terms of q_{11} and the identification of one shock via the Max-Share framework also gives the structural parameters of the second shock up to a sign normalization. Suppose, the first shock is identified with the

Max-Share strategy, then the impulse responses of the second shock are pinned down up to sign. With more than two instruments the Max-Share approach does not immediately grants identification of the remaining shocks that are related to the instruments. Though, one rarely has more than two convincing instruments at hand.

3.2.1 Ruling out Confounders via the Proxy VAR

Rather recently Dieppe et al. (2019) highlighted that the key identification assumption of the Max-Share framework is violated if another shock also contributes to the FEV of the target variable. To give an example, assume that the technology shock is accountable for most of the FEV of a total factor productivity (TFP) measurement. If the technology shock is identified as the shock that maximizes the contribution to the TFP measurement, the results will be biased if also other shocks contribute to the FEV of the TFP measurement. Dieppe et al. (2019) present strategies how to circumvent this problem of confounding shocks in the baseline Max-Share identification without proxies. In this section I focus on the remedies to the bias concerns that the proxy VAR framework offers.

First and foremost, the proxy VAR helps with the confounding shocks as it rules out confounding shocks that are not related to the proxies. Proposition 3.2 formalizes this property of the Max-Share approach in the proxy VAR. The proof is relegated to the appendix.

Proposition 3.2. (Exclusive Contribution) Suppose two valid proxy variables z_1 and z_2 are available which are related to shocks w_1 and w_2. If $\Omega_{i,2}(h) = 0$ holds for $\underline{h} \leq h \leq \overline{h}$. Then

$$q_1^* = \text{argmax} \sum_{h=\underline{h}}^{\overline{h}} \Omega_{i,1}^z(h) \quad \text{s.t.} \quad q_1' q_1 = 1,$$

point identifies the shocks w_1 and w_1 up to a sign normalization.

Proposition 3.2 establishes that if out of the shocks that are related to the instruments one exclusively contributes to the FEV of the target variable, both shocks are correctly identified by the Max-Share framework. For example, assume two proxies are used to identify two shocks. One of the shocks contributes the most to the FEV of a specified variable and out of the two shocks that are related to the proxies it contributes exclusively to the FEV of this variable. Suppose another shock that is not related to the proxies also contributes to the FEV of this target variable. Yet, as it is not related to the proxies it is ruled out as a confounding shock. As out of the two shocks that are related to the proxies only one exclusively contributes to the FEV of the target variable, the Max-Share framework correctly disentangles and identifies both shocks of interest.

The case of exclusive contribution is a rather strict assumption which also offers the possibility to identify the shocks through imposing a recursive structure on the proxy impact matrix. Yet, the next subsection describes how the Max-Share approach can be employed without the assumption of exclusive contribution.

3.2.2 Tackling the Bias

If the assumption of exclusive contribution for one of the two identified shocks is not reasonable, additional inequality restrictions $IQ_{h,i,j}(q_{11})$ which augment the Max-Share framework can help to nevertheless correctly identify the true structural parameters. Proposition 3.3 states a sufficient condition under which the structural parameters are identified. The rotation matrix that corresponds to the true structural parameters is denoted by Q^* with its columns q_j^*. In the following I restrict the first element of Q, q_{11} to be positive what simply helps with writing down the conditions for successful identification. In practice this means the shock is normalized to be an expansionary or contractionary - shock depending on the application - during the identification. Yet, after successful identification one can transform the shock into its contractionary/expansionary counterpart as usual by multiplying its respective column of the impact matrix or rotation matrix Q^* with -1.

Proposition 3.3. Let $q_{11} \in [0,1]$. Suppose two valid proxy variables z_1 and z_2 are available which are related to shocks w_1 and w_2. Suppose out of w_1 and w_2, w_1 predominantly, but not exclusively, contributes to the FEV of variable i for $\underline{h} \leq h \leq \overline{h}$. The augmented Max-Share approach

$$
\begin{aligned}
q_1^* = \operatorname{argmax} \sum_{h=\underline{h}}^{\overline{h}} \Omega_{i,1}^z(h) \qquad (29)\\
&\text{s.t.}\\
&q_1' q_1 = 1,\\
&q_1' q_2 = 0,\\
&q_{11} \geq 0,\\
&IQ_{h,i,1}(q_{11}) \lesseqqgtr \epsilon,
\end{aligned}
$$

identifies the true structural parameters if

(a) $IQ_{h,i,1}(q_{11}) \lesseqqgtr \epsilon$ implies a single binding linear restriction on q_{11} and

(b) the restrictions is set such that it implies $q_{11} \lesseqgtr q_{11}^*$.

Whether condition (a) of Proposition 3.3 is satisfied depends on the type of inequality restriction that is placed while condition (b) requires the economic intuition behind the chosen restriction to be correct.

This paper considers two types of inequality restrictions $IQ_{h,i,1}(q_1)$. Firstly, a standard sign restriction on the impulse response of the first shock on a specified variable for a single horizon h, i.e. $\eta_{i,1,h} \lesseqgtr \epsilon$. The second constraint restricts the difference of the of impulse responses of the two identified shocks on a single variable i for a single horizon h, i.e. $D_{i,h} = \eta_{i,1,h} - \eta_{i,2,h} \lesseqgtr \epsilon$, what translates to on shock having a larger impact than the other shock on the variable and horizon. The latter type of restriction adds another possibility to incorporate economic intuition into the identification and has the advantage that condition (a) of Proposition 3.3 can always be fulfilled. This sections proceeds to give sufficient conditions under which condition (a) is satisfied for the two above mentioned types of restrictions. Afterwards the implications of condition (b) are discussed.

Recall, that the structural impulse response of variable i to the first shock at horizon h is given by $\eta_{i,1,h} = e_i' C_h B_z q_1$, where e_i is the first column of the identity matrix I_2. Let

$$C_h B_z = \begin{pmatrix} CB_{1,h,1} & CB_{1,h,2} \\ CB_{2,h,1} & CB_{2,h,2} \\ \vdots & \vdots \\ CB_{k,h,1} & CB_{k,h,2} \end{pmatrix}.$$

Proposition 3.4. Let $q_{11} \in [0,1]$. Suppose a single linear restriction is placed on $\eta_{i,1,h}$. If the signs of $CB_{i,h,1}q_{11}$ and $CB_{i,h,2}q_{21}$ are the opposite the linear restriction of the form $\eta_{i,1,h} \lesseqgtr \epsilon$ implies a single linear restriction on q_{11} for $q_{11} \in [0,1]$.

Whether the condition of Proposition 3.4 is satisfied or not depends on the application and the variable whose response is restricted. In practice this condition is easy to check. Let Q^{max} be the rotation matrix obtained by the baseline Max-Share framework under the condition that $q_{11}^{max} \geq 0$. Then CB_{i1} and CB_{i2} are available reduced form parameters and q_{11} is normalized to be positive. The according sign of q_{21} is obtained via the sign of q_{21}^{max}. Yet, this is solely a sufficient condition and even if the it is not satisfied one can check whether the chosen restriction on $\eta_{i,1,h}$ implies a linear restriction on q_{11} by looking at the impulse response function $\eta_{i,1,h}(q_{11})$ over the domain $q_{11} \in [0,1]$.

For the difference restriction condition (a) of Proposition 3.3 can always be fulfilled with the correct normalization of the second column of Q. Recall that

$q_{22} = \pm q_{11}$ and $q_{12} = \pm q_{21}$. Hence, the second column is only pinned down up to sign due to the first column of Q. As for the first shock the second column of Q is normalized such that it either represents an expansionary or contractionary shock in the identification. Converting it back into a the opposing shock after successful identification is described above based on the first shock.

Proposition 3.5. Let $q_{11} \in [0, 1]$. Suppose the restriction is of the form $D_{i,h} \lesseqgtr \epsilon$. Let $\eta_{i,2,h}$ correspond to $q_{12} \geq 0$ and $-\eta_{i,2,h}$ to $q_{12} \leq 0$. Either the restriction $D_{i,h} = \eta_{i,1,h} - \eta_{i,2,h} \lesseqgtr \epsilon$ or $D_{i,h} = \eta_{i,1,h} - (-\eta_{i,2,h}) \lesseqgtr \epsilon$ implies a single linear restriction on q_{11} for $q_{11} \in [0, 1]$.

Algorithm 3 provides a guideline how to determine the needed normalization for the second column of the orthogonal matrix Q such that the difference restriction definitely implies a linear restriction on q_{11}.

Algorithm 3.

1. Carry out the baseline Max-Share identification under the condition $q_{11} \geq 0$ and check whether q_{21}^{max} is positive or negative.

2. Compute $DC_{i,h} = CB_{i,h,1} - CB_{i,h,2}$ and $SC_{i,h} = CB_{i,h,1} + CB_{i,h,2}$, where i and h correspond to the variable and horizon that is restricted. Record the signs of the two quantities.

3.1 If q_{21}^{max} is positive:

   (i) If the signs of $DC_{i,h}$ and $SC_{i,h}$ are the same normalize q_2 such that $q_{22} = q_{11}$ and $q_{12} = -q_{21}$.

   (ii) If the signs are the opposite normalize q_2 such that $q_{22} = -q_{11}$ and $q_{12} = q_{21}$.

3.2 If q_{21}^{max} is negative:

   (i) If the signs of $DC_{i,h}$ and $SC_{i,h}$ are the same normalize q_2 such that $q_{22} = q_{11}$ and $q_{12} = -q_{21}$.

   (ii) If the signs are the opposite normalize q_2 such that $q_{22} = -q_{11}$ and $q_{12} = q_{21}$.

In practice the normalization of the second column of Q alters how ϵ has to be specified as either $\eta_{i,1,h} - \eta_{i,2,h}$ or $\eta_{i,1,h} - (-\eta_{i,2,h})$ is restricted. One possible problem with this normalization of the second column is that the difference $D_{i,h}$ potentially turns into a sum if $\eta_{i,1,h}$ and $\eta_{i,2,h}/-\eta_{i,2,h}$ have the opposite sign. If that is the case it is difficult to specify a meaningful restriction as it is difficult to gauge the sum of two responses at a given horizon. Choosing the opposing normalization of the second column Q solves this issue but then condition (a) of Proposition 3.3 is not guaranteed. Hence, for the simple inequality restriction condition (a) of Proposition 3.3 is potentially not guaranteed and for the difference restriction it is potentially difficult to come up with a restriction that satisfies condition (b) of Proposition 3.3. Yet, condition (a) is just a sufficient condition and even if it is not guaranteed, the augmented Max-Share approach can work. See Appendix A for more details.

Condition (b) of Proposition 3.3 requires the restriction to be set such that $q_{11} \lesseqqgtr q_{11}^*$. For the simple sign restriction on the impulse responses this means that one response of the identified shock at one horizon has to be known. For the difference restriction one needs to gauge the difference in the responses of the shocks correctly. In practice such knowledge has to be drawn from economic theory, previous results or pure economic intuition.

Yet, if the true ϵ^* is unknown an inequality restriction can nevertheless help to reduce the bias. Let $\delta_{i,j,h} = \eta_{i,j,h}(q_{11}^*) - \eta_{i,j,h}(q_{11})$ be the bias of the structural impulse response of variable i at horizon h to shock j and $\epsilon^{max} = IQ_{i,j,h}(q_{11}^{max})$.

Proposition 3.6. (Reducing the Bias) Let $q_{11} \in [0, 1]$. Suppose $IQ_{h,i,1}(q_{11})$ is continuous in q_{11} and $IQ_{h,i,1}(q_{11}) \gtreqqless \epsilon$ implies a binding linear restriction on q_{11}. Let q_{11}^{res} be the implied restriction such that $q_{11} \gtreqqless q_{11}^{res}$. If $\epsilon^{max} < \epsilon < \epsilon^*$ ($\epsilon^{max} > \epsilon > \epsilon^*$) and $\eta_{i,j,h}(q_{11})$ is a strictly monotonic function at $q_{11} \in [q_{11}^{res}, q_{11}^*]$ ($q_{11} \in [q_{11}^*, q_{11}^{res}]$) the absolute bias $|\delta_{i,j,h}|$ decreases as $\epsilon^* - \epsilon$ ($\epsilon - \epsilon^*$) decreases.

If the condition of Proposition 3.4 holds or the second column of Q is normalized accordingly to Algorithm 3 then the underlying functions are strictly monotonic at $q_{11} \in [0, 1]$. Hence, $\epsilon^{max} < \epsilon^{res} < \epsilon^*$ implies $q_{11}^{max} < q_{11}^{res} < q_{11}^*$ or $q_{11}^{max} > q_{11}^{res} > q_{11}^*$. Further it implies that $|q_{11}^* - q_{11}^{res}|$ decreases as $\epsilon^* - \epsilon^{res}$ decreases. The same logic holds when $\epsilon^{max} > \epsilon^{res} > \epsilon^*$.

Thus, checking whether the bias is guaranteed to decrease in this case boils down to checking whether $\eta_{i,j,h}(q_{11})$ is strictly monotonous over the relevant domain. If the condition of Proposition 3.4 is fulfilled for the variable and horizon, $\eta_{i,1,h}(q_{11})$ is strictly monotonous for $q_{11} \in [0, 1]$, and thus also over the relevant domain. Following the idea behind Proposition 3.4, for the second shock on has to check whether $CB_{i,h,1}q_{12}$ and $CB_{i,h,2}q_{22}$ have opposite signs in order to determine whether $\eta_{i,2,h}(q_{11})$ is strictly monotonous or not.

Approaching ϵ^* from ϵ^{max} therefore reduces the bias. Yet, if one overshoots and i.e. $\epsilon^{max} < \epsilon^* < \epsilon^{res}$ is likely to increase again depending on the functional form of $\eta_{i,j,h}(q_{11})$. Hence, if the true ϵ^* is unknown a more conservative choice of ϵ

guarantee a reduction of the bias under the condition described in Proposition 3.6. In practice one can also report results for different values of ϵ over a range of values one is confident that ϵ^* is contained. This however de-sharpens the identification the larger the range of ϵ values. Thus the augmented Max-Share approach is most useful when a rather strong intuition for the correct ϵ^* exists.

3.2.3 Estimation and Inference

The combination of the proxy VAR with the Max-Share framework can also be handled with the bayesian Algorithm 1 that was depicted above. In combination with the Max-Share framework, steps two and three are simply replaced with carrying out the Max-Share optimization. Note that this reduces the robust bayesian approach to conventional bayesian inference. Step three in Algorithm 1, the core of the robust bayesian inference, computes the boundaries of the identified set and in the point identification case the identified set is a singleton. Hence, computing the bounds reduces to the computation of the point estimate.

However, in point identification scheme bootstrap inference is popular. In this case, I propose the bootstrap by Jentsch and Lunsford (2019), which is based on the heteroskedasticity robust bootstrap by Brüggemann, Jentsch, and Trenkler (2016). This approach relies on estimating Z and Σ to get an estimate for B_{11}^c (see Piffer & Podstawski, 2018). The bootstrap confidence intervals are constructed in the conventional way.

4 Simulation Results

In the simulation studies of this section I simulate a trivariate system. I follow Piffer and Podstawski (2018) and use the New Keynesian model by An and Schorfheide (2007) and Komunjer and Ng (2011). The model contains interest rates r_t, output x_t and inflation π_t. TFP shocks w_t^z, government spending shocks w_t^g and monetary shocks w_t^r are the structural shocks that hit the system. As pointed out by Giacomini (2013), calibrating the parameter gives following DGP:

$$\begin{pmatrix} r_t \\ x_t \\ \pi_t \end{pmatrix} = \begin{pmatrix} 0.79 & 0 & 0.25 \\ 0.19 & 0.95 & -0.46 \\ 0.12 & 0 & 0.62 \end{pmatrix} \begin{pmatrix} r_{t-1} \\ x_{t-1} \\ \pi_{t-1} \end{pmatrix} + \begin{pmatrix} 0.61 & 0 & 0.69 \\ 1.49 & 1 & -1.16 \\ 1.49 & 0 & -0.75 \end{pmatrix} \begin{pmatrix} w_t^z \\ w_t^g \\ w_t^r \end{pmatrix}. \tag{30}$$

In contrast to Piffer and Podstawski (2018) I set the variance of the structural shocks to unity in order easily compute the actual contribution to the FEV decomposition. Hence, the structural shocks are drawn from a normal distribution with mean zero and unit variance and then used to simulate the data with equation (30). The instruments are constructed with:

$$m_{1t} = \tau_1 w_t^z + (1-\tau_1) w_t^g + \tau_2 \nu_{1t}$$
$$m_{2t} = (1-\tau_1) w_t^z + \tau_1 w_t^g + \tau_2 \nu_{2t},$$

where τ_1 governs the strength of the relation of the first to shocks with the instruments and τ_2 governs the effect of the white noise disturbances ν_{1t} and ν_{2t}. I set $\tau_1 = 0.55$ and $\tau_2 = 0.01$ which leads to the proxies being sufficiently strong.

The technology shock w_t^z is accountable for the most part of the FEV of the interest rate r_t. An interesting feature of this DGP is that the government spending shock does not contribute to the FEV of the interest rate and inflation at any horizon. This enables me to construct two scenarios:

Scenario A: The proxies are constructed such that they are related to the technology shock ^z_t and the government spending shock w_t^g. Out of this two shocks w_t^z contributes exclusively to the FEV of the interest rate r_t.

Scenario B: The proxies are constructed such that they are related to the technology shock w_t^z and the monetary policy shock w_t^r. Both of this two shocks contribute to the FEV of the interest rate r_t.

4.1 Simulation Results - Max-Share

This subsection presents the simulation results for the case when to shocks are disentangled in the proxy VAR. The first results compare Scenario A and B which were described above. In Scenario A out of the two identified shocks, only the technology shocks contributes to the FEV of the interest rate. In Scenario B both shocks contribute to the FEV of the interest rate. Up to horizon $H = 13$ the technology shock w_t^z contributes on average 82%, the government spending shock w_t^g does not contribute to the FEV and the monetary policy shock w_t^r contributes the remaining 18%. In this section the interest rate is always the target variable in the Max-Share framework. Hence, the underlying assumption is that the technology shock is the one that maximizes the contribution to the FEV of the interest rate.

Table 1 depicts the results for the identification via maximization of the FEV after incorporating the information of the instruments as in (26) and without further inequality restrictions on relative magnitudes. As this is more of a confirmation exercise in which cases the Max-Share approach succeeds and fails I choose a large sample size of $T = 1,000$ with $M = 1,000$ Monte-Carlo iterations. The first two columns of the table show the true structural parameters of the DGP, the next two columns the combination of proxy VAR and Max-Share and the last two the identification via the Cholesky decomposition as in e.g. Mertens and Ravn (2013), i.e. $B_{11} = B_{11}^c$ is lower triangular. Looking at the results for Scenario A shows that after incorporating the proxies the shock that is not related to them is purged from the maximization problem. Only the two shocks that are related to the instruments are factored in and as one of the shocks contributes exclusively to the FEV of the target variable the Max-Share approach is suitable to disentangle these two. Yet, this set-up with exclusive contribution implies a recursive structure for the contemporaneous impacts of the shock, and thus the Cholesky decomposition for the upper 2×2 block of the proxy impact matrix will also identify the true underlying structural parameters.

Hence, the potentially more interesting case is when both identified shocks contribute to the FEV of the target variable as in Scenario B. As seen in the panel for Scenario B of Table 1, the Cholesky decomposition fails to identify the true impact matrix as there is no recursive structure between the two identified shocks. However, as expected also the combination of the proxy VAR with the Max-Share approach yields biased results because both shocks contribute to the FEV of the target variable. As pointed out by Dieppe et al. (2019), the Max-Share framework will be biased due to the confounding shock.

To successfully disentangle the two shocks when both shocks contribute to the FEV, the Max-Share framework needs to be augmented. In this case I use an in-

Table 1: Two Shocks without Exclusive Contribution

DGP		Proxy + Max-Share		Proxy + Cholesky	
Scenario A					
0.61	0	0.609 (0.007)	0.004 (0.004)	0.609 (0.007)	0 (0)
1.49	1	1.487 (0.016)	1.014 (0.036	1.485 (0.012)	1.004 (0.03)
1.49	0	1.488 (0.008)	0.015 (0.008)	1.488 (0.008)	0.004 (0.023)
Scenario B					
0.61	0.69	0.766 (0.008)	0.512 (0.004)	0.921 (0.006)	0 (0)
1.49	-1.1	1.161 (0.021)	1.624 (0.019	0.163 (0.022)	1.845 (0.016)
1.49	-0.75	1.25 (0.016)	1.578 (0.014)	0.426 (0.017)	1.6131 (0.011)

The table depicts the average of the estimated structural parameters over the 1,000 Monte-Carlo simulation iterations derived with the proxy SVAR. The shocks are once disentangled with the baseline Max-Share approach and once with the Cholesky decomposition. The value in the bracket depicts the respective standard deviation.

equality constraint which restricts the difference in the contemporaneous responses of output $D2,0$. In this DGP, a common, one standard deviation, expansionary technology shock has a more pronounced positive effect (1.49) than a common expansionary monetary policy shock (1.1). Augmenting the maximization problem such that this inequality constraint holds helps to disentangle the two shocks.

If the conditions of Proposition 3.3 hold, identification of the true underlying shocks is guaranteed. Algorithm 3 is used to normalize the second column such that $D_{2,0}(q_{11})$ is a monotonic function at $q_{11} \in [0, 1]$, and thus a linear restriction on it implies a linear restriction on q_{11}.

1. Carrying out the unrestricted Max-Share approach with $q_{11} \geq 0$ gives $q_{21} \geq 0$.
2. From the reduced form parameter one gets $DC_{2,0} \leq 0$ and $SC_{2,0} \geq 0$.

Table 2: Max-Share with Additional Inequality Constraint - Scenario B

DGP		Max-Share$^+$, T=1000		Max-Share$^+$, T=250	
Panel A: $\epsilon = 0.38$					
0.61	0.69	0.613 (0.029)	0.688 (0.026)	0.612 (0.057)	0.688 (0.052)
1.49	-1.1	1.486 (0.044)	-1.106 (0.044)	1.481 (0.087)	-1.101 (0.087)
1.49	-0.75	1.487 (0.031)	-0.755 (0.038)	1.483 (0.062)	-0.755 (0.075)
Panel B: $\epsilon = 0.2$					
0.61	0.69	0.659 (0.027)	0.644 (0.027)	0.658 (0.055)	0.643 (0.054)
1.49	-1.1	1.406 (0.044)	-1.206 (0.045)	1.405 (0.075)	-1.205 (0.074)
1.49	-0.75	1.432 (0.032)	-0.856 (0.037)	1.429 (0.062)	-0.858 (0.072)
Panel C: $\epsilon = 0$					
0.61	0.69	0.706 (0.026)	0.592 (0.027)	0.706 (0.053)	0.591 (0.055)
1.49	-1.1	1.311 (0.037)	-1.311 (0.037)	1.31 (0.075)	-1.309 (0.074)
1.49	-0.75	1.362 (0.032)	-0.963 (0.036)	1.36 (0.063)	-0.965 (0.07)

The table depicts the average of the estimated structural parameters over the 1,000 Monte-Carlo simulation iterations. The estimates are derived with the Max-Share$^+$ framework in the proxy VAR with different values of ϵ.

3.1 As the signs of $DC_{2,0}$ and $SC_{2,0}$ are the opposite the second column has to be normalized such that $q_{22} = q_{11} \geq 0$ and $q_{12} = -q_{12} \leq 0$.

With these normalizations the contemporaneous response of output to the technology shock is $\eta_{2,1,0} = 1.161$ while the response to the monetary shock is $\eta_{2,2,0} = 1.6124$ in the unrestricted Max-Share framework. Hence, $D_{2,0} = \eta_{2,1,0} - \eta_{2,2,0}$ is an actual difference as the signs of the two responses are the same. Knowing the DGP makes it easy also fulfil condition (b) of Proposition 3.3 by placing the constraint $D_{2,0} \geq \epsilon = 0.39$, which is a binding constraint. In order to impose such a constraint, one would need to argue that from prior knowledge or economic theory it is known that the common expansionary technology shock affects output more than the common expansionary monetary policy shock on impact. Without knowledge about the true DGP the true margin ϵ^* is typically unknown and needs to be gauged by the researcher. This simulation exercise reports results for different values of ϵ.

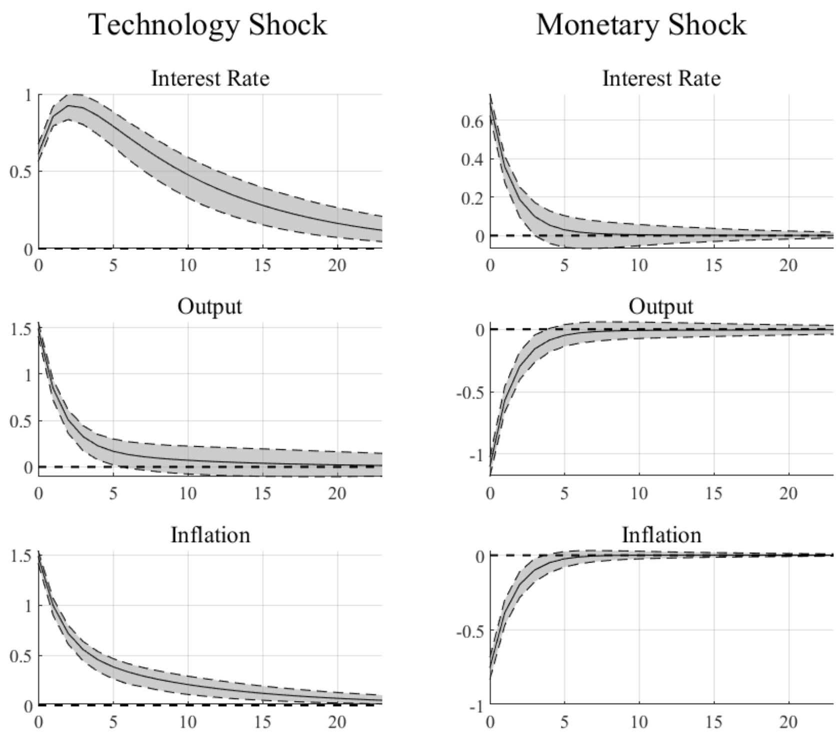
The full maximization problem for this particular simulation with the just mentioned restrictions is:

$$
\begin{aligned}
& q_1^* = \operatorname{argmax} \sum_{h=0}^{H} \Omega_{1,1}^{z}(h) \qquad (31) \\
& \text{s.t.} \\
& q_1' q_1 = 1, \\
& q_1' q_2 = 0, \\
& \eta_{2,1,0} - \eta_{2,2,0} \geq \epsilon, \\
& q_{11} \geq 0, \\
& q_{22} \geq 0.
\end{aligned}
$$

Table 2 depicts the results for this augmented Max-Share identification, denoted by Max-Share$^+$. The first two columns again depict the true DGP parameters, while columns three and four give the results of the Max-Share$^+$ framework with $T = 1000$ while the last two columns give the results for $T = 250$. Panel A shows the results for $\epsilon = 0.38$ which is very close to the true margin by which $\eta_{2,1,0}$ exceeds $\eta_{2,2,0}$ in absolute terms. Panel B shows the results for $\epsilon = 0.2$ and Panel C the results for $\epsilon = 0$. The latter represents the case when restriction boils down to a sign restriction on the relative magnitudes of the two shocks.

Comparing the estimates throughout the panels reveals that having the (almost) correct economic intuition with the inequality restriction ($\epsilon = 0.38$) removes almost all of the bias of the Max-Share approach. Yet, if the true margin is not met part of the bias remains and it is larger the less close the true margin is met. This is true for both the technology as well as the monetary policy shock. Comparing the results in terms of the sample sizes shows that also for smaller sample sizes like $T = 250$ the correct identification is achieved. Naturally, the bias also remains for smaller sample sizes when ϵ is not very close to the true margin.

Figure 1: 95% Point Estimate Bands, $\epsilon = 0.38$ - Scenario B



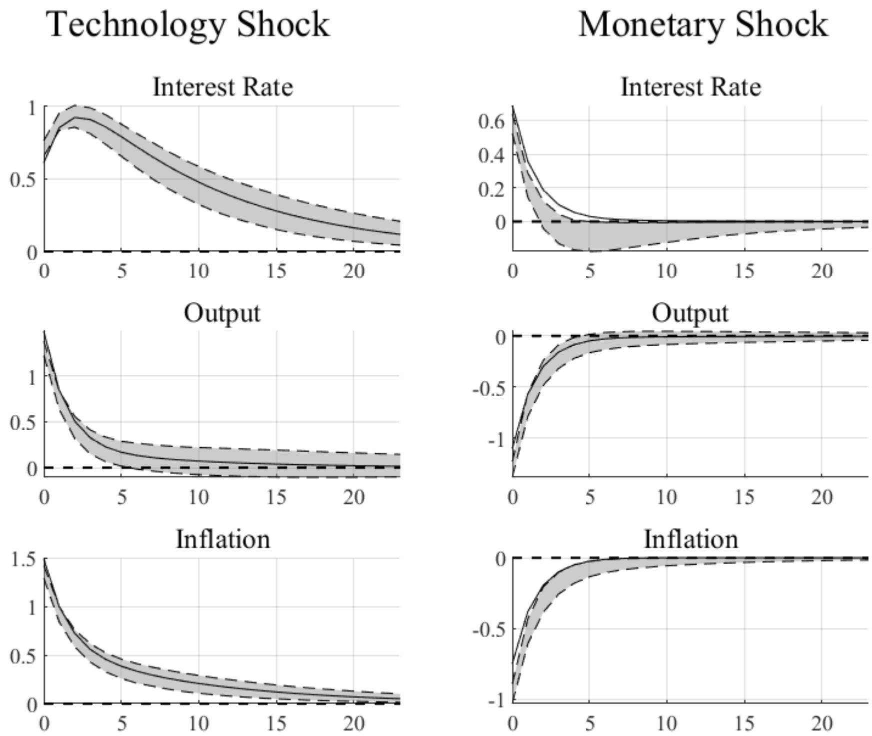
The solid lines depict the true impulse responses of the DGP. The dashed lines are the 2,5% and 97,5% quantile of the solutions found for the 1,000 Monte-Carlo simulation iterations.

Figure 1 and 2 show the 2,5% and 97,5% quantiles (dashed lines) of the estimated impulse responses identified by the augmented Max-Share throughout the 1,000 Monte-Carlo simulation rounds. The true impulse responses are depicted by the solid lines. Figure 1 presents the responses obtained with the $\epsilon = 0.38$, close to the true margin while Figure 2 presents the results for $\epsilon = 0$. The impulse responses for $\epsilon = 0.2$ can be found in the Appendix B.

The figures are in line with the results of Table 2. The estimated responses in Figure 1 closely identify the true structural impulse response parameters while the responses in Figure 2 reflect the bias of the impact matrix parameters in the bias of the impulse responses especially at earlier horizons.

The bias of the structural impulse responses at different horizon and for different values of ϵ is depicted in Table 3 for the technology shock and in Table 4 for the monetary policy shock. Comparing the results along the columns shows that the bias is typically larger the further away ϵ is from ϵ^* and for earlier horizons of the

Figure 2: 95% Point Estimate Bands, $\epsilon = 0$ - Scenario B



The solid lines are the true impulse responses of the DGP. The dashed lines are the 2,5% and 97,5% quantile of the solutions found for the 1,000 Monte-Carlo simulation iterations.

responses. These patterns are consistent throughout both tables, and therefore both identified shocks. Results for $T = 250$ are presented to the Appendix B.

The patters found in Tables 3 and 4 are visualized in Figures 3 and 4. The coloured lines depict the bias of the structural impulse responses at different horizons and for different values of ϵ. Comparing the results throughout the panels the bias is larger when ϵ is farther away from the true margin. This indicates that the condition of Proposition 3.6 is most likely fulfilled for most of the responses throughout the variables, shocks and horizons. Only the responses of output to the monetary policy shock $\eta_{2,2,h}$ for some horizons hint towards some non-monotonic functions. The same visualization for $T = 250$ is again relegated to the Appendix B.

Lastly, the additional constraints in the maximization problem (29) or (31) can also serve a pure inequality restrictions in order to set identify the shocks. Hence, one could also try to disentangle the shocks in the Proxy VAR with this inequality restrictions. The resulting sets of impulse responses are depicted in Figure 13 of the Appendix. The picture shows that the use of the Max-Share framework helps to estimate the structural impulse responses more precisely, as the simulated sets are rather wide compared to the range of simulated point estimates of the Max-Share approach for the majority of the impulse responses.

Table 3: Bias of the IRFs to the Technology Shock, $T = 1000$ - Scenario B

Variable	$H = 0$	$H = 6$	$H = 12$	$H = 18$
Panel A: $\epsilon = 0.38$				
r_t	0.0025	-0.0178	-0.0198	-0.0136
x_t	-0.0045	-0.0108	-0.0055	-0.0018
π_t	-0.0026	-0.0146	-0.0113	-0.0065
Panel B: $\epsilon = 0.2$				
r_t	0.0486	-0.0186	-0.0207	-0.014
x_t	-0.0844	-0.0131	-0.0061	-0.0022
π_t	-0.0584	-0.0163	-0.0117	-0.0067
Panel C: $\epsilon = 0$				
r_t	0.0958	-0.0233	-0.0235	-0.0154
x_t	-0.1795	-0.016	-0.0068	-0.0025
π_t	-0.1273	-0.0199	-0.0129	-0.0072

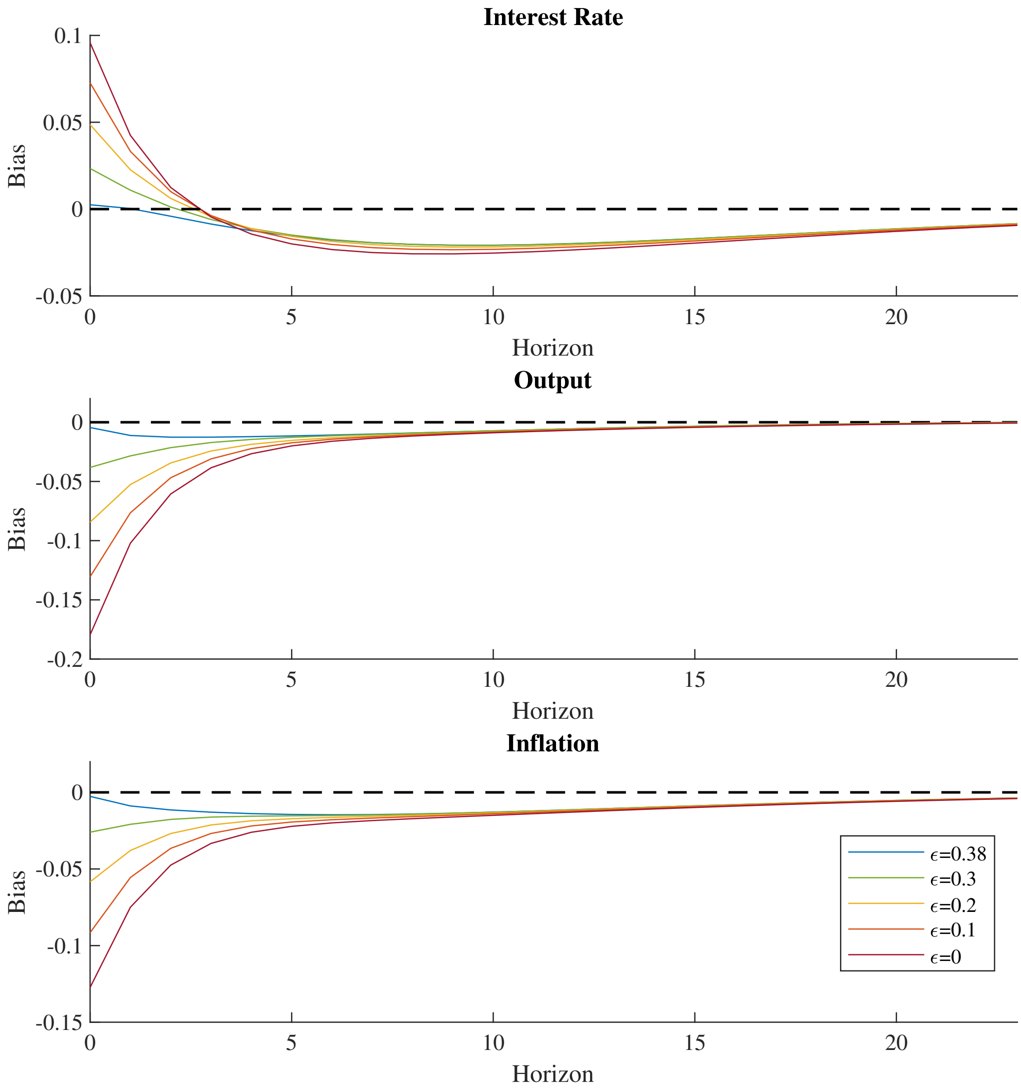
The table depicts the average of the estimated structural parameters over the 1,000 Monte-Carlo simulation iterations. The estimates are derived with the Max-Share^{+} framework in the proxy VAR with different values of ϵ.

Table 4: Bias of the Impulse Response Functions to the Monetary Shock, $T = 1000$.

Variable	$H = 0$	$H = 6$	$H = 12$	$H = 18$
Panel A: $\epsilon = 0.38$				
r_t	-0.0021	-0.0059	-0.0025	-0.0008
x_t	-0.0055	-0.0005	0.0009	0.001
π_t	-0.0051	-0.0029	-0.0007	-0.0001
Panel B: $\epsilon = 0.2$				
r_t	-0.0462	-0.0546	-0.0278	-0.0138
x_t	-0.1056	-0.0088	-0.0025	-0.0007
π_t	-0.1062	-0.0249	-0.0115	-0.0057
Panel C: $\epsilon = 0$				
r_t	-0.0982	-0.1084	-0.0557	-0.0282
x_t	-0.2105	-0.0184	-0.0067	-0.0029
π_t	-0.2135	-0.0491	-0.0234	-0.0119

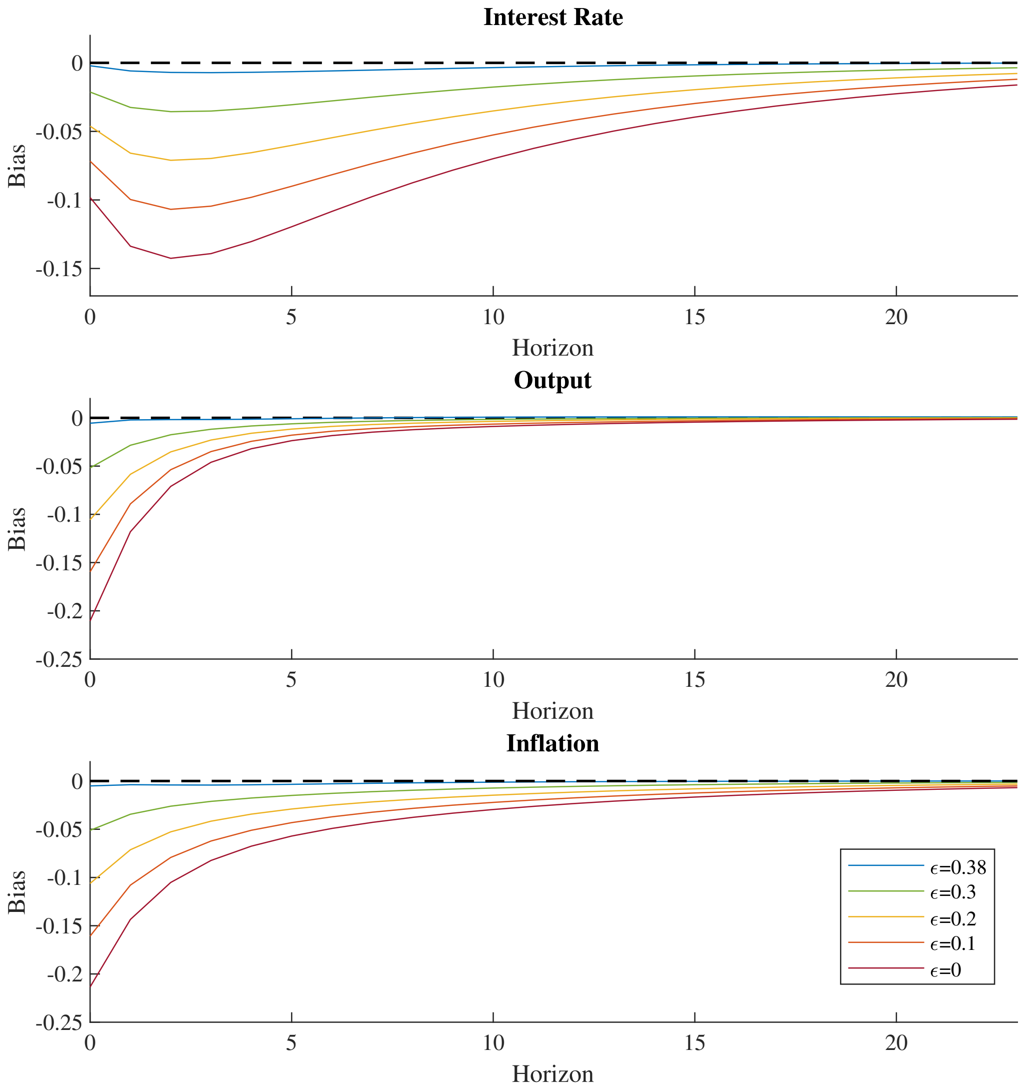
The table depicts the average of the estimated structural parameters over the 1,000 Monte-Carlo simulation iterations. The estimates are derived with the Max-Share^{+} framework in the proxy VAR with different values of ϵ.

Figure 3: Bias of the IRFs to the Technology Shock, $T = 1000$ - Scenario B



The coloured lines depict the average bias of the impulse response functions over 1,000 Monte-Carlo simulations for different values of ϵ.

Figure 4: Bias of the IRFs to the Monetary Shock, $T = 1000$ - Scenario B



The coloured lines depict the average bias of the impulse response functions over 1,000 Monte-Carlo simulations for different values of ϵ.

5 Conclusion

Identifying restrictions on the FEV can be used in multiple ways to disentangle the shocks in the proxy VAR framework. Firstly, bounds the FEV as introduced by Volpicella (2021) replace or accompany other inequality restrictions to disentangle the shocks in the proxy VAR. This paper provides the general framework of the bounds applied to the proxy VAR, how to conduct robust bayesian inference in this setting.

Second, the structural parameters are sharply point identified in the proxy VAR with the help of the Max-Share framework. In the highly relevant case of two proxies shocks can be identified by the baseline Max-Share framework if out of the shocks related to the proxies one exclusively contributes to the FEV of the target variable. Confounding shocks not related to the proxies are avoided. Without exclusive contribution, the bias due to confounding shocks in Max-Share approach can be tackled by augmenting the Max-Share approach with an inequality constraint. This paper provides sufficient conditions under which the inequality constraint yields identification of the true underlying structural parameters. Further, it assesses when these conditions are met for simple inequality restrictions on an impulse response and on the difference of the impulse responses of the two identified shocks.

A simulation study illustrates how to assess the conditions behind the inequality constraint and showcases successful identification. Further it illustrates the behaviour of the bias of Max-Share approach if the conditions for successful identification are not entirely met. For the future an empirical illustration is planned to highlight the usefulness in practice.

References

An, S., & Schorfheide, F. (2007). Bayesian Analysis of DSGE models. *Econometric Reviews*, *26*(2-4), 113-172.

Barsky, R. B., & Sims, E. R. (2011). News shocks and business cycles. *Journal of Monetary Economics*, *58*(3), 273-289.

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

Ben Zeev, N., & Pappa, E. (2017). Chronicle of a War Foretold: The Macroeconomic Effects of Anticipated Defence Spending Shocks. *The Economic Journal*, *127*(603), 1568-1597.

Brüggemann, R., Jentsch, C., & Trenkler, C. (2016). Inference in VARs with conditional heteroskedasticity of unknown form. *Journal of Econometrics*, *191*(1), 69-85.

Dieppe, A., Neville, F., & Kindberg-Hanlon, G. (2019). *New Approaches to the Identification of Low-Frequency Drivers: An Application to Technology Shocks.* The World Bank.

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

Francis, N., Owyang, M. T., Roush, J. E., & DiCecio, R. (2014). A Flexible Finite-Horizon Alternative to Long-Run Restrictions with an Application to Technology Shocks. *The Review of Economics and Statistics*, *96*(4), 638-647.

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

Giacomini, R. (2013, 12). The relationship between DSGE and VAR models. In *Advances in econometrics* (Vol. 32, p. 1-25).

Giacomini, R., Kitagawa, T., & Read, M. (2021). Robust bayesian inference in proxy svars. *Journal of Econometrics*.

Jentsch, C., & Lunsford, K. G. (2019, May). *Asymptotically Valid Bootstrap Inference for Proxy SVARs* (Working Papers No. 201908).

Kilian, L., & Lütkepohl, H. (2017). *Structural vector autoregressive analysis.* Cambridge University Press.

Komunjer, I., & Ng, S. (2011). Dynamic identification of dynamic stochastic general equilibrium models. *Econometrica*, *79*(6), 1995–2032.

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

Review, *103*(4), 1212-47.

Montiel Olea, J. L., & Nesbit, J. (2021). (machine) learning parameter regions. *Journal of Econometrics*, *222*(1, Part C), 716-744.

Piffer, M., & Podstawski, M. (2018). Identifying Uncertainty Shocks Using the Price of Gold. *The Economic Journal*, *128*(616), 3266-3284.

Stock, J., & Watson, M. (2012). Disentangling the Channels of the 2007-2009 Recession. *Brookings Papers on Economic Activity*, *Spring 2012*, 81-135.

Uhlig, H. (2004a). Do technology shocks lead to a fall in total hours worked? *Journal of the European Economic Association*, *2*(2-3), 361-371.

Uhlig, H. (2004b). *What moves GNP?* (Econometric Society 2004 North American Winter Meetings No. 636).

Volpicella, A. (2021). Svars identification through bounds on the forecast error variance. *Journal of Business & Economic Statistics*, *0*(0), 1-11.

A Appendix A

Proof of Proposition 3.1.

The proof of Proposition 3.1 closely follows the ideas of the proof for Proposition 3.1 in the article by Volpicella (2021). The contribution of shock j to the FEV of variable i at horizon h is given by:

$$\Omega_{i,j}(h) = q_j' R_{i,h} q_j, \tag{32}$$

where $R_{i,h}$ is a positive semidefinite symmetric $l \times l$ real matrix. As $R_{i,h}$ is symmetric it can be diagonalized such that:

$$P' R_{i,h} P = D, \tag{33}$$

where P is an orthogonal matrix and D a diagonal matrix with the real eigenvalues λ_m^{ih} of matrix $R_{i,h}$ as entries on the diagonal for $m = 1, ..., l$.

For the $l \times 1$ orthogonal eigenvector q_m it holds that:

$$R_{i,h} q_m = \lambda_m^{ih} q_m, \tag{34}$$

and thus:

$$q_m' R_{i,h} q_m = \lambda_m^{ih} q_m' q_m = \lambda_m^{ih}. \tag{35}$$

The bound restrictions on the FEV are collected by:

$$\underline{\tau}_{i,j,h} \leq q_j' R_{i,h} q_j \leq \overline{\tau}_{i,j,h}, \text{ for } i \in \mathcal{I}_j \text{ and } h \in \mathcal{H}_{ij}. \tag{36}$$

Hence, if there exists an eigenvalue for which $\underline{\tau}_{i,j,h} \leq \lambda_m^{ih} \leq \overline{\tau}_{i,j,h}$ it holds that:

$$\underline{\tau}_{i,j,h} \leq q_m' R_{i,h} q_m \leq \overline{\tau}_{i,j,h}, \tag{37}$$

and condition (a) is satisfied for $q_j = q_m$. Condition (b) states that q_m satisfies also the remaining bound restrictions for all $i \in \mathcal{I}_j$ and $h \in \mathcal{H}_{ij}$ and the all the other identifying restrictions that are imposed. It follow that there exists an orthogonal matrix $Q = [q_1, ..., q_m, ..., q_l]$ for which all restrictions are satisfied and the identified set is non-empty.

Given that the identified set for the structural impulse responses is non-empty $\eta_{i,j,h} = e_i' C_m B_z q_j$ exists. Due to the restriction that the reduced-form VAR process is invertible in holds that $||e_i' C_m B_z|| < \infty$. Thus, it holds that $|\eta_{i,j,h}| \leq ||e_i' C_m B_z|| < \infty$ and the identified set for the impulse responses is bounded.

■

To aid with the proof of Proposition 3.2 I first define under which conditions a shock does not contribute to the FEV of a specified variable.

Definition 1: Let c^h_{ij} be the ij-th element of the moving average coefficient matrix C_h and b^z_{ij} the ij-th element of the proxy impact matrix B_z. Shock j does not contribute to the FEV of variable n at horizon $\overline{h}$ if the following conditions hold:

a) $\overline{h} = 0$: The nj-th element of the impact matrix $b^z_{nj} = 0$.

b) $\overline{h} > 0$: Additionally to condition a), $c^h_{ni} = 0$ if $b_{ij} \neq 0$ for $i = 1, ..., k$, $i \neq n$ and for all $h \leq \overline{h}$.

Condition b) of Definition 1 states that all variables affected by shock j are not allowed to affect the target variable n in the moving average process over all horizons lower than $\overline{h}$.

Example: Consider the case with two instruments in a three variable system. The second shock does not contribute to the FEV of the first variable. Conditions a) and b), for example, require:

$$B_z = \begin{pmatrix} * & 0 \\ * & * \\ * & * \end{pmatrix} \quad \text{and} \quad C_h = \begin{pmatrix} * & 0 & 0 \\ * & * & * \\ * & * & * \end{pmatrix} \quad \text{for all} \quad h \leq \overline{h},$$

or

$$B_z = \begin{pmatrix} * & 0 \\ * & * \\ * & 0 \end{pmatrix} \quad \text{and} \quad C_h = \begin{pmatrix} * & 0 & * \\ * & * & * \\ * & 0 & * \end{pmatrix} \quad \text{for all} \quad h \leq \overline{h}.$$

Proof of Proposition 3.2.

Suppose the two proxies z_1 and z_2 are valid and satisfy conditions (9) of the main text. Hence, the proxy VAR alone will identify the true proxy impact matrix B_z up to a rotation with an orthonormal matrix Q. For simplicity and w.l.o.g. assume that the initial rotation matrix Q is just the identity matrix such that proxy impact matrix in the maximization problem is just the true proxy impact matrix. Furthermore, w.l.o.g assume that the first two shocks of the system are related to the proxy variable and that the second shock does not contribute to the FEV of the first variable over the horizon $\underline{h}$ up to $\overline{h}$. Condition a) of Definition 1 requires structure of the proxy impact matrix to be:

$$B_z = \begin{pmatrix} * & 0 \\ * & * \\ \vdots & \vdots \\ * & * \end{pmatrix}.$$

As the first variable contributes exclusively to the FEV of the first variable, the Max-Share approach maximizes the contribution of the first shock to the FEV of the first variable.

Recall that the contribution of the first shock to the FEV of the first variable at horizon h can be written as:

$$\Omega_{1,1}(h) = q_1' R_{1,h} q_1. \tag{38}$$

Uhlig (2004b) shows that the sum over all $\Omega_{1,1}(h)$ from horizons $\underline{h}$ up to $\overline{h}$ can be expressed by:

$$\sum_{h=\underline{h}}^{\overline{h}} \Omega_{1,1}^z(h) = q_1' S q_1, \tag{39}$$

where

$$S = \sum_{h=0}^{\overline{h}} (\overline{h} + 1 - \max(\underline{h}, h))(e_1 C_h B_z)'(e_1 C_h B_z) = \sum_{h=0}^{\overline{h}} (\overline{h} + 1 - \max(\underline{h}, h)) S_h, \tag{40}$$

with e_1 being the first row of the identity matrix I_k. Hence, $e_1 C_h B_z$ is just the first row of $C_h B_z$.

For $h = 0$, C_0 is just the identity matrix, and $e_1 C_0 B_z = (* \ 0)$, and thus

$$S_0 = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix}.$$

If $h > 0$, due to condition b) of Definition 1, $e_1 C_h B_z = (* \ 0)$ for $h < \overline{h}$, and thus

$$S_h = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix} \quad \forall \ h < \overline{h}.$$

As S is just the weighted sum over all S_h,

$$S = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix}.$$

Uhlig (2004b) shows that finding the rotation vector q_1^* that maximizes the sum over the contributions

$$q_1^* = \text{argmax } q_1' S q_1 \quad \text{s.t.} \quad q_1' q_1 = 1 \tag{41}$$

amounts to finding the eigenvector associated with the largest eigenvalue of the matrix S. Due to the particular structure of the matrix S induced by the exclusive contribution it can easily be seen that the eigenvector associated with the non-zero eigenvalue is just the first column of the identity matrix I_2. Hence, the rotation matrix Q^* is just the identity matrix I_2 up to a sign normalization of the columns. Thus, the maximum of the contribution to the FEV is achieved at the true proxy impact matrix parameters up to a sign normalization of the columns.

■

Hence, with the appropriate ordering of the variables in the system choosing the the lower triangular Cholesky decomposition B_{11}^c as a solution for $B_{11} = \Sigma_{11} - B_{12} B_{12}'$ immediately identifies the true structural parameters. Yet, if a different solution for $B_{11} = \Sigma_{11} - B_{12} B_{12}'$ is chosen the Max-Share maximization approach will revert this initial rotation as the maximum is achieved at the true parameters.

This can be illustrated for the case where $\underline{h} = \overline{h} = 0$. Again assume that the instruments are valid and that the first two shocks are the target shocks, from which the first shocks exclusively contributes to the FEV of the first variable. Now a different initial solution to $B_{11} = \Sigma_{11} - B_{12} B_{12}'$ is obtained for the proxy impact matrix. This solution is simply a rotation of the true proxy impact matrix with an orthogonal matrix Q_{init}:

$$B_z^q = B_z Q_{init} = \begin{pmatrix} b_{11} & 0 \\ b_{21} & b_{22} \\ \vdots & \vdots \\ b_{k1} & b_{k2} \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}.$$

Hence, the first column of the rotated proxy impact matrix is $(b_{11} q_{11} \; b_{11} q_{12})$ and

$$S = \begin{pmatrix} b_{11}^2 q_{11}^2 & b_{11}^2 q_{11} q_{12} \\ b_{11}^2 q_{11} q_{12} & b_{11}^2 q_{12}^2 \end{pmatrix}.$$

It can be shown that the eigenvalues of S are $[b_{11}^2, 0]$ and a corresponding eigenvector to the non-zero eigenvalue is

$$q_1^* = \begin{pmatrix} q_{11} \\ q_{12} \end{pmatrix}.$$

Hence, the optimal rotation matrix due to the Max-Share identification is Q'_{init}, up to a sign normalization of the columns. Thus, the initial rotation is reverted as

$$B_z Q_{init} Q'_{init} = B_z \tag{42}$$

and the true impact matrix parameters are identified. Again, up to a sign normalization of the columns.

Proof of Proposition 3.3.

Let $q_{11} \in [0, 1]$. Without loss of generality assume that the first shock is the target shock in the maximization of the Max-Share approach. In general the maximization problem of the Max-Share framework can be written as:

$$q_1^* = \text{argmax } q_1' S q_1 \quad \text{s.t.} \quad q_1' q_1 = 1, \tag{43}$$

where S is defined above and is a symmetric and positive semi definite matrix. Let:

$$S = \begin{pmatrix} s_{11} & s_{21} \\ s_{21} & s_{22} \end{pmatrix} \quad \text{and} \quad q_1 = \begin{pmatrix} q_{11} \\ q_{21} \end{pmatrix},$$

where $q_{21} = \sqrt{1 - q_{11}^2}$. Hence, the maximization problen in (43) can be written as:

$$q_1^* = \text{argmax } q_{11}^2 s_{11} + 2 q_{11} q_{21} s_{21} + q_{21}^2 s_{22} \quad \text{s.t.} \quad q_1' q_1 = 1. \tag{44}$$

As S is positive semi definite $q_{11}^2 s_{11} \geq 0$ and $q_{21}^2 s_{22} \geq 0$. To achieve the maximum the middle term also has to be positive. As $q_{11} \geq 0$ this term is positive if q_{21} and s_{21} have the same sign. Hence, $q_{21} = \sqrt{1 - q_{11}^2}$ is implicitly normalized to be the same sign as s_{21} due to the maximization in the Max-Share framework.

With the normalization of q_{21} the contribution to the FEV, $\Omega_{1,1}^z(q_{11}) = q_1' S q_1$ has one extreme value at q_{11}^{max} and $\Omega_{1,1}^z(q_{11})$ increases at $q_{11} \in [0, q_{11}^{max}]$ and decreases at $q_{11} \in [q_{11}^{max}, 1]$.

Let the true rotation, that yields the true structural parameters, be q_{11}^*. Suppose $q_{11}^* \leq q_{11}^{max}$. Suppose the inequality restriction $IQ_{h,i,1}(q_{11}) \lesseqgtr \epsilon$ implies a binding linear restriction on q_{11} such that $q_{11} < q_{11}^*$. As $\Omega_{1,1}^z(q_{11})$ is a strictly increasing function at $q_{11} \in [0, q_{11}^{max}]$ the maximum of the maximization problem (43) augmented with $IQ_{h,i,1}(q_{11}) \lesseqgtr \epsilon$ is at q_{11}^*. The very same logic holds when $q_{11} > q_{11}^*$.
■

Proof of Proposition 3.4.

Let $q_{11} \in [0, 1]$. Recall that the impulse response of variable i at horizon h to the first shock is $\eta_{i,1,h} = e_i' C_h B_z q_1$, with e_i being the first column of I_k and

$$C_h B_z = \begin{pmatrix} CB_{1,h,1} & CB_{1,h,2} \\ CB_{2,h,1} & CB_{2,h,2} \\ \vdots & \vdots \\ CB_{k,h,1} & CB_{k,h,2} \end{pmatrix}.$$

Hence, the impulse response function $\eta_{i,1,h}(q_{11})$ is given by:

$$\eta_{i,1,h} = CB_{i,h,1}q_{11} + CB_{i,h,2}q_{21} = \underbrace{CB_{i,h,1}q_{11}}_{\text{A}} + \underbrace{CB_{i,h,2}\sqrt{1-q_{11}^2}}_{\text{B}}. \tag{45}$$

If the signs of A and B are the opposite $\eta_{i,1,h}(q_{11})$ is monotonic at $q_{11} \in [0,1]$, and thus a linear restriction on $\eta_{i,1,h}(q_{11})$ implies a linear restriction on q_{11}.
■

Note that the signs of q_{11} and $q_{21} = \sqrt{1-q_{11}^2}$ are fixed due to the normalization that $q_{11} \geq 0$ and the maximization of the Max-Share framework (see. proof of Proposition 3.3). Hence, in practice one can determine the signs of terms A and B in the proof of Proposition 3.4 easily by checking whether the baseline Max-Share framework gives you a positive or negative q_{21} when fixing q_{11} to be positive.

Further note, that a linear restriction on $\eta_{i,1,h}(q_{11})$ can also imply a linear restriction on q_{11} if the sufficient condition of Proposition 3.4 is not fulfilled. Let q_{11}^{res} be the value of...(to be continued)

Proof of Proposition 3.5.

Let $q_{11} \in [0,1]$. The difference in the impulse responses of variable i at horizon h to the two identified shocks $D_{i,h}$ is given by:

$$D_{i,h} = \eta_{i,1,h} - \eta_{i,2,h} = e_i' C_h B_z q_1 - e_i' C_h B_z q_2, \tag{46}$$

where e_i is again the first column of I_k. This difference can also be written as:

$$\begin{aligned} D_{i,h} &= (CB_{i,h,1}q_{11} + CB_{i,h,2}q_{21}) - (CB_{i,h,1}q_{12} + CB_{i,h,2}q_{22}) \\ &= CB_{i,h,1}(q_{11} - q_{12}) + CB_{i,h,2}(q_{21} - q_{22}). \end{aligned}$$

Due to the properties of 2×2 orthonormal matrices $q_{22} = \pm q_{11}$, $q_{12} = \pm q_{21}$ and $q_{21} = \sqrt{1-q_{11}^2}$. Yet, q_{12} and q_{22} have opposite signs if the sings of q_{11} and q_{21} are the same and vice versa. Hence, $D_{i,h}(q_{11})$ is a function of q_{11}. Let $\eta_{i,h,2}$ be the response that corresponds to $q_{22} \geq 0$ and $-\eta_{i,h,2}$ the one that corresponds to $q_{22} \leq 0$.

Due to the properties of orthonormal matrices:

$$D_{i,h}(q_{11}) = \underbrace{(CB_{i,h,1} - CB_{i,h,2})}_{\text{A}} q_{11} + \underbrace{(CB_{i,h,1} + CB_{i,h,2})}_{\text{B}} \sqrt{1-q_{11}^2} \quad \text{if } q_{22} \geq 0 \quad (47)$$

$$D_{i,h}^{-}(q_{11}) = \underbrace{(CB_{i,h,1} + CB_{i,h,2})}_{\text{B}} q_{11} + \underbrace{(CB_{i,h,2} - CB_{i,h,1})}_{\text{-A}} \sqrt{1-q_{11}^2} \quad \text{if } q_{22} \leq 0. \quad (48)$$

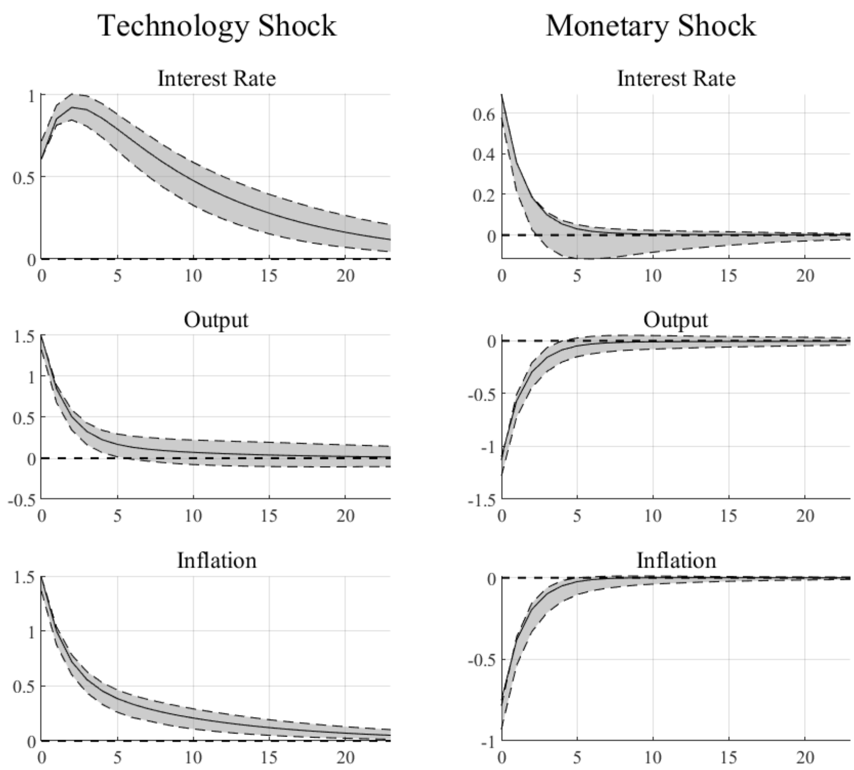
The signs of q_{11} and $\sqrt{1-q_{11}^2}$ are determined in the maximization process of the Max-Share approach (see proof of Proposition 3.3) and they are either the same or the opposite. From equations (47) and (48) you see that either $D_{i,h}(q_{11})$ or $D_{i,h}^{-}(q_{11})$ is a monotonic function over the domain $q_{11} \in [0,1]$ depending on the sign normalization of the second column of Q. Hence, you can always express the difference as a monotonic function at $q_{11} \in [0,1]$ and a linear restriction on either $D_{i,h}(q_{11})$ or $D_{i,h}^{-}(q_{11})$ implies a linear restriction on q_{11} at $q_{11} \in [0,1]$.

■

Proof of Proposition 3.6. Let $q_{11} \in [0,1]$. Let $\delta_{i,j,h} = \eta_{i,j,h}(q_{11}^*) - \eta_{i,1,h}(q_{11}$ be the bias of the structural impulse response of variable i at horizon h to shock j. Suppose $q_{11}^{max} < q_{11}^{res} < q_{11}^*$. If $\eta_{i,j,h}(q_{11})$ is strictly monotonically increasing at $q_{11} \in [q_{11}^{max}, q_{11}^*]$, then $\eta_{i,j,h}(q_{11}^{max}) < \eta_{i,j,h}(q_{11}^{res}) < \eta_{i,j,h}(q_{11}^*)$ and $\delta_{i,j,h}^{res} < \delta_{i,j,h}^{max}$. The same logic holds for $q_{11}^{max} > q_{11}^{res} > q_{11}^*$ and/or for strictly decreasing functions. Hence, if q_{11}^{res} approaches q_{11}^* from q_{11}^{max}, $|\delta_{i,j,h}|$ decreases. ■

B Appendix B

Figure 5: 95% Point Estimate Bands, $\epsilon = 0.2$



The solid lines are the true impulse responses and the dashed lines are the 2,5% and 97,5% quantile of the solutions found for the 1,000 simulation iterations.

Table 5: Bias of the IRFs to the Technology Shock, $T = 250$ - Scenario B

Variable	$H = 0$	$H = 6$	$H = 12$	$H = 18$
Panel A: $\epsilon = 0.38$				
r_t	0.0023	-0.0745	-0.0736	-0.0473
x_t	-0.0084	-0.0381	-0.0173	-0.0059
π_t	-0.0072	-0.0576	-0.0398	-0.0219
Panel B: $\epsilon = 0.2$				
r_t	0.0484	-0.0748	-0.0739	-0.0474
x_t	-0.0846	-0.0394	-0.0171	-0.0058
π_t	-0.0607	-0.059	-0.0399	-0.0219
Panel C: $\epsilon = 0$				
r_t	0.0956	-0.0795	-0.0764	-0.0485
x_t	-0.1803	-0.0421	-0.0176	-0.006
π_t	-0.1302	-0.0624	-0.0409	-0.0224

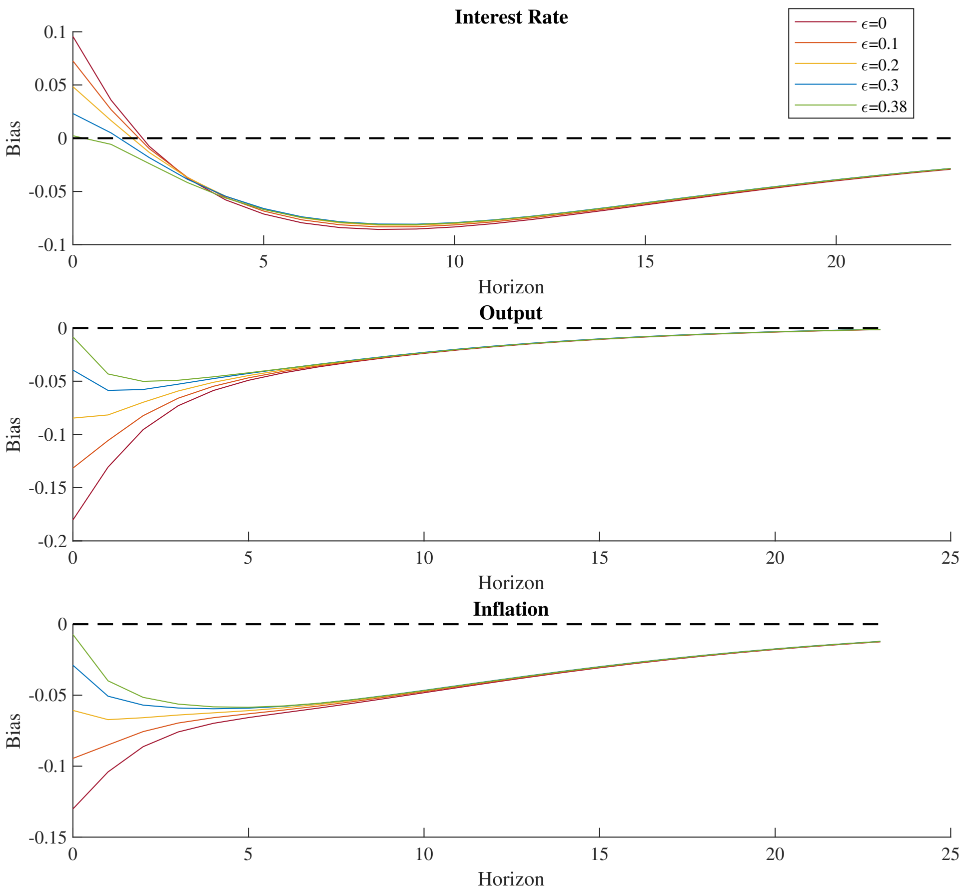
The table depicts the average of the estimated structural parameters over the 1,000 Monte-Carlo simulation iterations. The estimates are derived with the Max-Share$^+$ framework in the proxy VAR with different values of ϵ.

Table 6: Bias of the IRFs to the Monetary Shock, $T = 250$ - Scenario B

Variable	$H = 0$	$H = 6$	$H = 12$	$H = 18$
Panel A: $\epsilon = 0.38$				
r_t	-0.0026	-0.0105	-0.003	-0.0001
x_t	-0.0016	0.0032	0.0049	0.0039
π_t	-0.005	-0.0046	-0.0003	0.0004
Panel B: $\epsilon = 0.2$				
r_t	-0.0465	-0.0561	-0.0249	-0.011
x_t	-0.1053	-0.0041	0.0017	0.0021
π_t	-0.1082	-0.024	-0.0093	-0.0042
Panel C: $\epsilon = 0$				
r_t	-0.0986	-0.1052	-0.0486	-0.0228
x_t	-0.2086	-0.011	-0.0012	0.0005
π_t	-0.2146	-0.0449	-0.0189	-0.0092

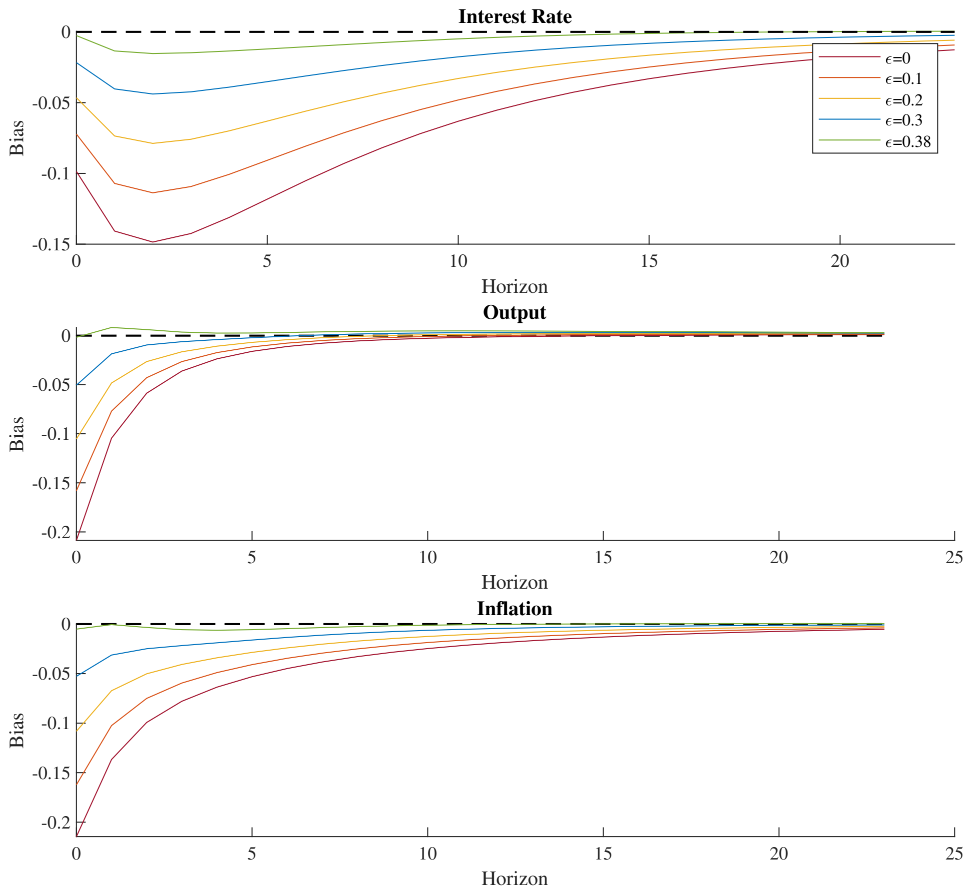
The table depicts the average of the estimated structural parameters over the 1,000 Monte-Carlo simulation iterations. The estimates are derived with the Max-Share$^+$ framework in the proxy VAR with different values of ϵ.

Figure 6: Bias of the IRFs to the Technology Shock, $T = 250$ - Scenario B



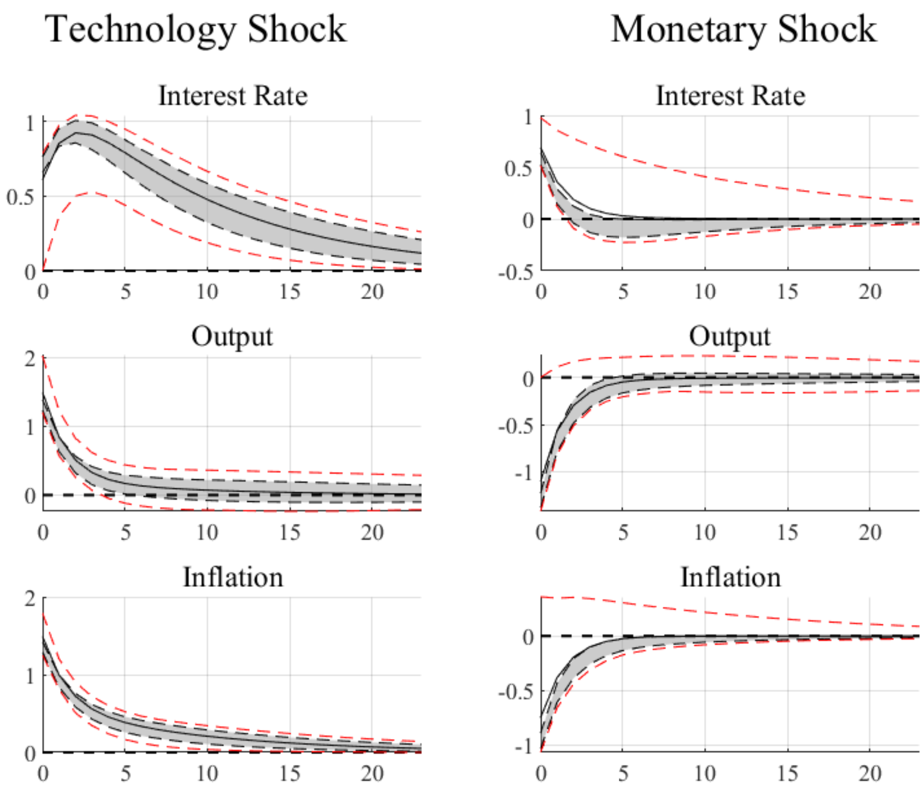
The coloured lines depict the average bias of the impulse response functions over 1,000 Monte-Carlo simulations for different values of ϵ.

Figure 7: Bias of the IRFs to the Monetary Shock, $T = 250$ - Scenario B



The coloured lines depict the average bias of the impulse response functions over 1,000 Monte-Carlo simulations for different values of ϵ.

Figure 8: Point Estimate Bands and Sign Restriction Sets, $\epsilon = 0$.



The solid lines are the true impulse responses and the dashed lines are the 2,5% and 97,5% quantile of the solutions found for the 1,000 simulation iterations. The red dashed line are the maximum and minimum responses of the identified set of the proxy SVAR disentangled via pure sign restrictions. The sign restrictions correspond to the constraints of the maximization problem.