Discussion Papers in Economics

ESTIMATION AND INFERENCE OF THE FORECAST ERROR VARIANCE DECOMPOSITION FOR SET-IDENTIFIED SVARS

By

Francesco Fusari
(Newcastle University Business School),

Joe Marlow
(University of Surrey)

&

Alessio Volpicella
(University of Surrey)

DP 04/24

School of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web https://www.surrey.ac.uk/school-economics

ISSN: 1749-5075

Estimation and Inference of the Forecast Error Variance Decomposition for Set-Identified SVARs*

Francesco Fusari †, Joe Marlow‡ and Alessio Volpicella§

This draft: September 2024.

Abstract

We study the Structural Vector Autoregressions (SVARs) that impose internal and external restrictions to set-identify the Forecast Error Variance Decomposition (FEVD). This object measures the importance of shocks for macroeconomic fluctuations and is therefore of first-order interest in business cycle analysis. We make the following contributions. First, we characterize the endpoints of the FEVD as the extreme eigenvalues of a symmetric reduced-form matrix. A consistent plug-in estimator naturally follows. Second, we use the perturbation theory to prove that the endpoints of the FEVD are differentiable. Third, we construct confidence intervals that are uniformly consistent in level and have asymptotic Bayesian interpretation. We also describe the conditions to derive uniformly consistent confidence intervals for impulse responses. A Monte-Carlo exercise demonstrates the approach properties in finite samples. An unconventional monetary policy application illustrates our toolkit.

Keywords: Forecast Error Variance Decomposition, Set-Identification, SVAR, Unconventional Monetary Policy.

JEL: C1, C32, E47, E52.

---

*We would like to thank Raffaella Giacomini, Jonas Arias, Atsushi Inoue, Christian Matthes, Henrique Castro-Pires and Stefano Soccorsi for suggestions. This paper also benefited from comments of seminar participants at Indiana University, University of Essex, Lancaster University and University of Pavia. We are grateful to the audience of the following conferences: 31^{st} Symposium of the Society for Nonlinear Dynamics and Econometrics, 2024 Annual Conference of the Scottish Economic Society, 8^{th} RCEA Time Series Econometrics Workshop, 2^{nd} UEA Time Series Workshop, 28^{th} EAYE Annual Meeting, North American Summer Meeting (NASM) of the Econometric Society 2024, 2024 Conference of the Bristol Econometrics Study Group, European Summer Meeting (ESEM) of the Econometric Society 2024 and 4^th Sailing the Macro Workshop. The authors thank the University of Surrey for financial support. Fusari thankfully acknowledges the Turing Scheme for funding his visit to Indiana University, where part of this research was conducted. This paper is the Chapter 3 of Fusari's PhD thesis.

†Newcastle University Business School, 5 Barrack Road, Newcastle upon Tyne NE1 4SE, UK. Email: francesco.fusari@newcastle.ac.uk

‡University of Surrey, Elizabeth Fry Building (AD), Guildford GU2 7XH, UK. Email: j.marlow@surrey.ac.uk

§Corresponding Author, University of Surrey, Elizabeth Fry Building (AD), Guildford GU2 7XH, UK. Email: a.volpicella@surrey.ac.uk

1 Introduction and Related Literature

A common practice in empirical macroeconomics is to set-identify the parameters of a Structural Vector Autoregression (SVAR). This relies on sign restrictions (Uhlig (2005) and subsequent literature) and Proxy SVARs, where set-identification naturally arises when the instrument is not fully exogenous (correlated with shocks other than that of interest -plausible proxy-), e.g. Piffer and Podstawski (2018), Braun and Brüggemann (2022), Jarociński and Karadi (2020), Ludvigson et al. (2021), Fusari (2023), Caggiano et al. (2021) and Caggiano and Castelnuovo (2023). In this setting, the Forecast Error Variance Decomposition (FEVD) is a pivotal object. It describes the contribution or importance of the identified shock to explain the fluctuations (volatility) of the variables of interest over time. It is reported in any empirical application along with the Impulse Response Functions (IRFs) and yields pivotal economic information, especially for business cycle analysis (Christiano et al., 1999; Smets and Wouters, 2007; Beaudry and Portier, 2006). Furthermore, it is also a crucial object of interest in the recent empirical research on the long-term ("hysteresis") effects of demand shocks (Benati and Lubik, 2022; Furlanetto et al., forthcoming). FEVD is also commonly used as a source of identifying information, the so-called *Max Share Identification*, popularized by Uhlig (2004). For example, this approach identifies technology shocks as those which explain the most of the FEV decomposition of labor productivity at 10-year period (Francis et al., 2014).[1]

Empirical practice for the estimation and inference of set-identified FEVD mostly employs Bayesian methods (Arias et al., 2018). The main concern about standard Bayesian analysis in set-identified frameworks is that the posterior distributions are influenced by the prior specification, even asymptotically (Poirier, 1998; Baumeister and Hamilton, 2015). Second, any selection of prior breaks down the asymptotic equivalence between Bayesian and frequentist inference, in the sense that the former asymptotically lies inside the true identified set (Moon and Schorfheide, 2012). This led to alternative methodologies that do not require the characterization of a prior specification over the set-identified structural parameters. Granziera et al. (2018) proposed a frequentist approach, where a moment-inequality-minimum-distance toolkit delivers estimation and inference for the IRFs. Gafarov et al. (2018) presented a delta-method interval for the IRFs. Giacomini and Kitagawa (2021) [GK21] delivered robust Bayes credible interval that achieves a given credibility level regardless of the prior specified over the set-identified structural parameters.[2] The common feature of those prior-robust contributions

[1]Other applications include DiCecio and Owyang (2010) (technology shocks), Barsky and Sims (2011) and Kurmann and Sims (2021) (news shocks), Mumtaz et al. (2018) (credit shocks), Mumtaz and Theodoridis (2023) (inflation target shocks), Caldara et al. (2016) (uncertainty and credit shocks), Levchenko and Pandalai-Nayar (2020) (sentiment shocks) and Angeletos et al. (2020) (a variety of supply and demand shocks). Volpicella (2022) and Carriero and Volpicella (2024) extended the setting allowing set-identification and multiple-shock identification, respectively.

[2]See Giacomini et al. (2022a) for the extension to plausible proxy SVARs.

is that they are specifically constructed for estimation and inference of IRFs. In principle, an exception allowing robust inference for the set-identified FEVD is GK21. However, the computational challenges of the toolkit prevent its wide use in the empirical practice.

We provide a computationally viable estimation and inference toolkit for the FEVD in set-identified SVARs without relying on any prior specification. The toolkit can be also used to quantify the sensitivity of the standard Bayesian inference to the choice of an unrevisable prior. First, we characterize the lower and upper bounds of the set of the FEVD as the extreme eigenvalues of a symmetric reduced-form matrix. We show that those endpoints correspond to the solutions of a quadratic constrained optimization problem involving the orthonormal matrix transforming reduced-form shocks into structural shocks. In particular, the problem can be solved iteratively by active (binding) set strategies or interior point methods where each iteration requires the solution of an equality constrained problem. Consistency of a plug-in estimator follows. A user-friendly algorithm is provided. Second, we prove the differentiability of the endpoints with respect to the reduced-form parameters. A by-product of the FEVD differentiability is that we can guarantee frequentist validity of the robust Bayesian inference in GK21 when applied to the FEVD. Third, we propose a delta-method interval adjusted by the length of the identified set that is both uniformly consistent in level and has asymptotic Bayesian interpretation. Furthermore, we illustrate that our machinery has computational advantages with respect to GK21 and is extremely user-friendly. We also show the adjustment under which a delta-method confidence interval for set-identified IRFs, i.e. Gafarov et al. (2018), is uniformly consistent. A simple Monte Carlo simulation displays the properties of the approach in finite samples and draws the attention to the improvements in coverage by using a set-length adjusted delta-method confidence interval with respect to alternatives. The empirical application, which relies on set-identification of unconventional monetary policy shocks, illustrates the toolkit. In our exercise, under a standard Bayesian approach, the findings about the FEVD are mostly driven by an unrevisable prior rather than identification itself. Computational efficiency of our framework makes it viable for SVARs with several variables and restrictions, unlike alternative robust frameworks, e.g. GK21.

While our coverage statements hold for the true value of the object of interest, i.e. FEVD, rather than for its true identified set, literature for inference over interval-defined parameters typically faces a well-known trade-off. On the one hand, valid confidence intervals - in the sense that the coverage is at least equal to the nominal confidence level- for the parameter of interest are not necessarily valid for the set. On the other hand, valid confidence intervals for the set tend to be unnecessarily conservative for the targeted parameter (Imbens and Manski, 2004; Stoye, 2009).[3] Thus, we also provide users with a further confidence interval for the identified

[3]Imbens and Manski (2004), Stoye (2009), Gafarov et al. (2018) and Granziera et al. (2018) focused on inference over parameter; Giacomini and Kitagawa (2021) provided a valid confidence interval for the set;

set.

Our proposal has two main limitations. First, while it can accommodate zero restrictions on more shocks under some conditions, sign constraints can restrict a single shock only. It is true that some applications sign-restrict more shocks. However, the empirical practice often sign-constrains a single shock: among many others, see Uhlig (2005), Scholl and Uhlig (2008), Baumeister and Hamilton (2018), Dedola et al. (2017) and Arias et al. (2019) for identification of monetary policy shocks; Fujita (2011) identifies the effects of labour market shocks; Beaudry et al. (2011) study the effects of sentiment shocks; the application of Amir-Ahmadi and Drautzburg (2021) identifies news shocks; Dedola and Neri (2007), Peersman and Straub (2009) and Mumtaz and Zanetti (2012) sign-restrict technology shocks; Mumtaz et al. (2018) analyse the effects of credit supply shocks. Furthermore, estimation and inference frameworks for sign-restricted impulse responses typically allow constraints on a single shock only; this applies to the frequentist machinery in Granziera et al. (2018) and the delta-method of Gafarov et al. (2018). In principle, robust Bayesian approach of GK21 allows sign restrictions on more shocks; in practice, the non-linear nature of the algorithm for inference makes the methodology cumbersome as the number of constraints increases. Classical Bayesian approach in Arias et al. (2018) accommodates restrictions on a multiplicity of shocks, but it is sensitive to the prior specification, even asymptotically.

Second, when our framework is applied to Proxy SVARs, the instrument cannot be weak. The intuition being that the latter would make the reduced-form gradient matrix of the constraints rank deficient, so that the Karush-Kuhn-Tucker conditions cannot be derived.

This article is also related to the literature providing estimation and inference for the shocks' contribution to volatility in dynamic models. Early literature includes Lütkepohl (1990), who presented delta-method intervals for the FEVD in point-identified SVARs. Our paper focuses on set-identified SVARs. Phillips (1998) illustrated the FEVD asymptotics for nonstationary VAR. Amisano and Giannini (1997) delivered the asymptotics for the FEVD in SVARs. Braun and Mittnik (1993) analysed the effect of some VAR misspecifications on the FEVD estimation, e.g. omitted variables, ignored moving average terms, incorrectly specified lag lengths, or incorrect orthogonalization of innovations. Lanne and Nyberg (2016) relied on the well-known generalized IRFs to propose generalized FEVD. Gorodnichenko and Lee (2020) presented an estimator in local projections. Plagborg-Møller and Wolf (2022) proposed a frequentist procedures for conducting inference in a general moving-average model with external instruments. The framework they consider also includes set-identification.

Furthermore, this paper shares the spirit of the delta-method in Gafarov et al. (2018). They provide inference for the set-identified IRFs, and the extension to the FEVD is practically

---

Plagborg-Møller and Wolf (2022) proposed inference over the set, but suggested - in the practical implementation- to adjust it for the parameter.

feasible, but not recommended. The reason being that the FEVD at a given horizon is a non-linear function of IRFs in previous periods. For each horizon, Gafarov et al. (2018) characterize the upper and lower bounds of the IRFs by finding the optimal rotation matrix, i.e. structural model. However, using their estimated IRFs as plug-in estimator for the FEVD implies that the latter comes from a multiplicity of structural models (rotation matrices), resulting in a loss of interpretability.[4] Furthermore, some by-products of our paper include i) conditions under which the toolkit in Gafarov et al. (2018) can accommodate zero restrictions on more shocks, and ii) showing that adjusting the IRFs confidence interval in Gafarov et al. (2018) by the set-length makes it uniformly, rather than point-wise, consistent.

The paper is organized as follows. Section 2 introduces the SVAR and econometric framework. Section 3 delivers estimation, differentiability and inference for the FEVD in set-identified SVARs. Section 4 presents the empirical application. Section 5 illustrates the Monte-Carlo simulations. Section 6 concludes. The appendices provide omitted proofs (Appendix A), further findings from the empirical application (Appendix B) and additional simulation results (Appendix C).

2 Econometric Framework

2.1 SVAR

Consider a SVAR(p) model

$$\boldsymbol{A}_0 \boldsymbol{y}_t = \boldsymbol{a} + \sum_{j=1}^{p} \boldsymbol{A}_j \boldsymbol{y}_{t-j} + \boldsymbol{\epsilon}_t \tag{2.1}$$

for $t = 1, \ldots, T$, where $\boldsymbol{y}_t$ is an $n \times 1$ vector of endogenous variables, $\boldsymbol{\epsilon}_t$ an $n \times 1$ vector white noise process, normally distributed with mean zero and variance-covariance matrix $\boldsymbol{I}_n$, $\boldsymbol{A}_j$ is an $n \times n$ matrix of structural coefficient for $j = 0, \ldots, p$. $\boldsymbol{A}_0$ has positive diagonal elements (with sign normalisations), and is invertible. The initial conditions $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_p$ are given. Let $\boldsymbol{x}_t = \left(\boldsymbol{y}'_{t-1}, \ldots, \boldsymbol{y}'_{t-p}\right)'$ and $\boldsymbol{A}_+ = (\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p)$; we can write the SVAR(p) as

$$\boldsymbol{A}_0 \boldsymbol{y}_t = \boldsymbol{a} + \boldsymbol{A}_+ \boldsymbol{x}_t + \boldsymbol{\epsilon}_t, \tag{2.2}$$

with $\boldsymbol{\theta} = (\boldsymbol{A}_0, \boldsymbol{A}_+)$ structural parameters. The reduced-form VAR is as follows:

$$\boldsymbol{y}_t = \boldsymbol{b} + \boldsymbol{B} \boldsymbol{x}_t + \boldsymbol{u}_t, \tag{2.3}$$

where $\boldsymbol{b} = \boldsymbol{A}_0^{-1} \boldsymbol{a}$ is an $n \times 1$ vector of constants, $\boldsymbol{B}_j = \boldsymbol{A}_0^{-1} \boldsymbol{A}_j$, $\boldsymbol{u}_t = \boldsymbol{A}_0^{-1} \boldsymbol{\epsilon}_t$ denotes the $n \times 1$ vector of reduced-form errors. $var(\boldsymbol{u}_t) = E(\boldsymbol{u}_t \boldsymbol{u}'_t) = \boldsymbol{\Sigma} = \boldsymbol{A}_0^{-1} (\boldsymbol{A}_0^{-1})'$ is the $n \times n$ variance-covariance matrix of reduced-form errors. Assume that the reduced-form VAR is invertible into an infinite-order moving average (VMA(∞)) process.

[4]The same critique applies to usage of the toolkit in Granziera et al. (2018) for estimation of the FEVD.

We can reparameterize the VAR and write

$$\boldsymbol{y}_t = \boldsymbol{b} + \boldsymbol{B}\boldsymbol{x}_t + \boldsymbol{\Sigma}_{tr}\boldsymbol{Q}\boldsymbol{\varepsilon}_t, \tag{2.4}$$

where $\boldsymbol{\Sigma}_{tr}$ denotes the lower triangular Cholesky matrix with non-negative diagonal coefficients of $\boldsymbol{\Sigma}$, $\boldsymbol{Q} \in \boldsymbol{\Theta}(n)$ is an $n \times n$ matrix and $\boldsymbol{\Theta}(n)$ characterises the set of all orthonormal such matrices. The mapping is the following: $\boldsymbol{B} = \boldsymbol{A}_0^{-1}\boldsymbol{A}_+$, $\boldsymbol{\Sigma} = \boldsymbol{A}_0^{-1}(\boldsymbol{A}_0^{-1})'$, and $\boldsymbol{Q} = \boldsymbol{\Sigma}_{tr}^{-1}\boldsymbol{A}_0^{-1}$; alternatively, $\boldsymbol{A}_0 = \boldsymbol{Q}'\boldsymbol{\Sigma}_{tr}^{-1}$ and $\boldsymbol{A}_+ = \boldsymbol{Q}'\boldsymbol{\Sigma}_{tr}^{-1}\boldsymbol{B}$. Sign normalisations on $\boldsymbol{A}_0$ correspond to $diag(\boldsymbol{Q}'\boldsymbol{\Sigma}_{\boldsymbol{tr}}^{-1}) \geq \boldsymbol{0}_{n\times 1}$. Thus, the VMA($\infty$) is

$$\boldsymbol{y}_t = \boldsymbol{b} + \sum_{h=0}^{\infty} \boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{u}_{t-h} = \boldsymbol{b} + \sum_{h=0}^{\infty} \boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{\Sigma}_{\boldsymbol{tr}}\boldsymbol{Q}\boldsymbol{\varepsilon}_{\boldsymbol{t}}, \tag{2.5}$$

with $\boldsymbol{C}_h(\boldsymbol{B})$ being the h-th term of $(\boldsymbol{I}_n - \sum_{h=1}^{p} \boldsymbol{B}_h L^h)^{-1}$. The impulse response function of variable i to the j-th shock at horizon h is

$$r_{ijh} = \boldsymbol{e}_i'\boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{\Sigma}_{\boldsymbol{tr}}\boldsymbol{Q}\boldsymbol{e}_j = \boldsymbol{c}_{ih}'(\boldsymbol{\phi})\boldsymbol{q}_j, \tag{2.6}$$

where $\boldsymbol{e}_i$ is the i-th column vector of $\boldsymbol{I}_n$, $\boldsymbol{q}_j$ is the j-th column of $\boldsymbol{Q}$, $\boldsymbol{c}_{ih}'(\boldsymbol{\phi})$ represents the i-th row vector of $\boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{\Sigma}_{tr}$ and $\boldsymbol{\phi}$ collects the reduced-form parameters: $\boldsymbol{\phi} \equiv (\boldsymbol{b}', vec(\boldsymbol{B})', vech(\boldsymbol{\Sigma})') \in \boldsymbol{\Xi} \subset \mathcal{R}^d$. Note that $vec(\bullet)$ stacks the columns of any $n \times z$ matrix $\bullet$ to form a $nz \times 1$ vector (we sometimes use $vech(\bullet)$, vectorizing a $n \times n$ symmetric matrix $\bullet$ into a $\frac{n(n+1)}{2} \times 1$ vector) and $\otimes$ is the Kronecker product. Finally, in this paper the notation "-1" refers to the inverse for square matrices and the pseudo-inverse for non-square and/or reduced-rank matrices.

2.2 Forecast Error Variance Decomposition

The h-step ahead Forecast Error (FE) for a SVAR, as in equation (2.1), given all the data up to $t-1$, is $\boldsymbol{FE}(h) \equiv \boldsymbol{y}_{t+h} - \boldsymbol{y}_{t+h|t-1}$. Thus, the FEV at horizon h is

$$\boldsymbol{FEV}(h) \equiv E\left[(\boldsymbol{y}_{t+h} - \boldsymbol{y}_{t+h|t-1})(\boldsymbol{y}_{t+h} - \boldsymbol{y}_{t+h|t-1})'\right]. \tag{2.7}$$

As a result, the contribution of shock j to the FEV of variable i at horizon h is

$$FEVD_{ijh} \equiv \frac{FEV_j^i(h)}{FEV^i(h)} = \frac{\sum_{\tilde{h}=0}^{h} r_{ij\tilde{h}}^2}{\sum_{j=1}^{n}\sum_{\tilde{h}=0}^{h} r_{ij\tilde{h}}^2}, \tag{2.8}$$

where $FEV_j^i(h)$ is the FEV of variable i due to shock j at horizon h, $FEV^i(h)$ denotes the total FEV of variable i at horizon h, and $0 \leq FEVD_j^i(h) \leq 1$ by definition. Faust (1998) and Uhlig (2004) showed that equation (2.8) can be written as

$$FEVD_{ijh} = (\boldsymbol{Q}\boldsymbol{e}_j)'\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})(\boldsymbol{Q}\boldsymbol{e}_j) = \boldsymbol{q}_j'\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})\boldsymbol{q}_j, \tag{2.9}$$

where $\mathbf{\Upsilon}^i_h(\boldsymbol{\phi}) = \frac{\sum_{\tilde{h}=0}^h \boldsymbol{c}_{i\tilde{h}}(\boldsymbol{\phi})\boldsymbol{c}'_{i\tilde{h}}(\boldsymbol{\phi})}{\sum_{\tilde{h}=0}^h \boldsymbol{c}'_{i\tilde{h}}(\boldsymbol{\phi})\boldsymbol{c}_{i\tilde{h}}(\boldsymbol{\phi})}$ is a symmetric positive semidefinite $n \times n$ real matrix.

The FEVD represents the contribution of the identified shock to explain the fluctuations of the variables of interest over time. It is reported in any empirical application along with IRFs and provides important economic information and can be also employed for identification (*Max Share* approach). For local projections, the measure of shock importance is typically the Forecast Variance Ratio (FVR) (Gorodnichenko and Lee, 2020; Plagborg-Møller and Wolf, 2022). While the two measures usually diverge (the FVR is based on the observables, while the FEVD relies on all structural shocks), they are equivalent if all SVAR shocks are invertible, e.g. there is no information asymmetry between agent and econometrician.[5] This is why the macroeconometric literature has mostly ignored the discrepancy between the two concepts. Furthermore, while we impose stationarity, if the deviations from some equilibrium relationship are stationary, e.g. Engle and Granger (1987), results of this paper apply to the FEVD of target variables to the disturbances of the equilibrium.

Standard econometric literature mostly focuses on estimation and inference for IRFs; on the other hand, this paper provides a toolkit targeting set-identified FEVD.

2.3 Set-Identification

Set-identification for structural parameters and their functions, such as $FEVD_{ijh}$, arises when reduced-form parameters $\boldsymbol{\phi}$ cannot pin down a unique $\boldsymbol{A}_0$. Any $\boldsymbol{A}_0 = \boldsymbol{Q}'\boldsymbol{\Sigma}_{tr}^{-1}$ satisfies $\boldsymbol{\Sigma} = \boldsymbol{A}_0^{-1}(\boldsymbol{A}_0^{-1})'$, so the identified set for $\boldsymbol{A}_0$ is $\{\boldsymbol{A}_0 = \boldsymbol{Q}'\boldsymbol{\Sigma}_{tr}^{-1} : \boldsymbol{Q} \in \boldsymbol{\Theta}(n)\}$, where $\boldsymbol{\Theta}(n)$ is the set of $n \times n$ orthonormal matrices. Identification therefore requires to placing a set of restrictions on $\boldsymbol{Q}$, i.e. reducing the set of feasible $\boldsymbol{Q}$s to a subspace $\boldsymbol{\mathcal{Q}} \in \boldsymbol{\Theta}(n)$. The identified set for the FEVD would be

$$IS_{FEVD}(\boldsymbol{\phi}) = \{FEVD_{ijh} : \boldsymbol{Q} \in \boldsymbol{\mathcal{Q}}\}. \tag{2.10}$$

The following subsections describe the common identifying restrictions used for set-identification.

2.3.1 Zero Restrictions

Set-identifying zero restrictions include constraints on some off-diagonal elements of $\boldsymbol{A}_0$, on the lagged coefficients $\boldsymbol{A}_l$ for $l = 1, \ldots, p$, on the contemporaneous responses $\boldsymbol{A}_0^{-1}$, and on the long-run responses ($\boldsymbol{LIR} = \sum_{h=0}^{\infty} \boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{\Sigma}_{tr}\boldsymbol{Q}$). All these constraints are linear restrictions

[5] For the formal definition of invertibility and related concepts, see Chapter 17 in Kilian and Lütkepohl (2017).

on the columns of $\boldsymbol{Q}$:

$$(i,j)\text{th element of } \boldsymbol{A}_0 = 0 \Leftrightarrow (\boldsymbol{\Sigma}_{tr}^{-1}\boldsymbol{e}_j)'\boldsymbol{q}_i = 0 \tag{2.11}$$

$$(i,j)\text{th element of } \boldsymbol{A}_l = 0 \Leftrightarrow (\boldsymbol{\Sigma}_{tr}^{-1}\boldsymbol{B}_l\boldsymbol{e}_j)'\boldsymbol{q}_i = 0 \tag{2.12}$$

$$(i,j)\text{th element of } \boldsymbol{A}_0^{-1} = 0 \Leftrightarrow (\boldsymbol{e}_i'\boldsymbol{\Sigma}_{tr}^{-1})\boldsymbol{q}_j = 0 \tag{2.13}$$

$$(i,j)\text{th element of } \boldsymbol{LIR} = 0 \Leftrightarrow \left[\boldsymbol{e}_i' \sum_{h=0}^{\infty} \boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{\Sigma}_{tr}\right]\boldsymbol{q}_j = 0. \tag{2.14}$$

Also, exogeneity conditions in Proxy SVARs can be characterized as exclusion restrictions on the columns of the orthonormal matrix: $E(\boldsymbol{m}_t\boldsymbol{\varepsilon}_t') = \boldsymbol{D}\boldsymbol{\Sigma}_{tr}\boldsymbol{Q} = \left[\boldsymbol{0}_{k\times(n-k)}, \boldsymbol{\Psi}\right]$, where $\boldsymbol{m}_t$ is the $k \times 1$ vector containing k instruments, $\boldsymbol{D}$ is a reduced-form $k \times n$ matrix coming from the regression of $\boldsymbol{m}_t$ on $\boldsymbol{y}_t$ (first-stage regression) and $\boldsymbol{\Psi}$ is the $k \times k$ matrix representing the strength of the instruments, i.e. the correlation between $\boldsymbol{m}_t$ and the k instrumented shocks.[6] [7] The exogeneity restrictions between $\boldsymbol{m}_t$ and a shock j can be expressed as

$$(1:k, j)\text{th elements of } E(\boldsymbol{m}_t\boldsymbol{\varepsilon}_t') = \boldsymbol{0} \Leftrightarrow \boldsymbol{D}\boldsymbol{\Sigma}_{tr}\boldsymbol{q}_j = \boldsymbol{0}. \tag{2.15}$$

Thus, we can collect the zero restrictions as follows:

$$\boldsymbol{F}(\boldsymbol{\phi}, \boldsymbol{Q}) \equiv \begin{pmatrix} \boldsymbol{F}_1(\boldsymbol{\phi})\boldsymbol{q}_1 \\ \vdots \\ \boldsymbol{F}_n(\boldsymbol{\phi})\boldsymbol{q}_n \end{pmatrix} = \boldsymbol{0}_{(\sum_{i=1}^n f_i)\times 1}, \quad \boldsymbol{F}_i(\boldsymbol{\phi})\text{: } f_i \times n, \tag{2.16}$$

where f_i denotes the number of zero restrictions on shock i. In other words, each row of $\boldsymbol{F}_i(\boldsymbol{\phi})$ collects the coefficient vector of an exclusion constraints that restricts $\boldsymbol{q}_i$ as in (2.11)-(2.15). If $f_i = 0$, there are no zero restrictions on $\boldsymbol{q}_i$ and $\boldsymbol{F}_i(\boldsymbol{\phi})$ does not exist.

2.3.2 Sign Restrictions

Sign constraints can be imposed alone or in addition to zero restrictions. Let s_{hj} denote the number of sign restrictions on impulse responses at horizon h. The sign restrictions on shock j are $\boldsymbol{S}_{hj}(\boldsymbol{\phi})\boldsymbol{q}_j \geq \boldsymbol{0}$, where $\boldsymbol{S}_{hj}(\boldsymbol{\phi}) \equiv \tilde{\boldsymbol{D}}_{hj}\boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{\Sigma}_{tr}$ is a $s_{hj} \times n$ matrix and $\tilde{\boldsymbol{D}}_{hj}$ is the $s_{hj} \times n$ selection matrix that selects the sign-restricted responses from the $n \times 1$ response vector $\boldsymbol{C}_h(\boldsymbol{B})\boldsymbol{\Sigma}_{tr}\boldsymbol{q}_j$. The nonzero elements of $\tilde{\boldsymbol{D}}_{hj}$ can be equal to 1 or to -1 depending on the sign of the restriction on the impulse response of interest. By considering multiple horizons, the whole set of sign restrictions placed on the $j-$th shock is $\boldsymbol{S}_j(\boldsymbol{\phi})\boldsymbol{q}_j \geq \boldsymbol{0}$. Specifically, $\boldsymbol{S}_j$ is a $\left(s_j = \sum_{h=0}^{\bar{h}_j} s_{hj}\right) \times n$ matrix defined by $\boldsymbol{S}_j(\boldsymbol{\phi}) = \left[\boldsymbol{S}_{0j}'(\boldsymbol{\phi}), \ldots, \boldsymbol{S}_{\bar{h}_j j}'(\boldsymbol{\phi})\right]'$. Let $\mathcal{I}_{\mathcal{S}} \subset \{1, 2, \ldots, n\}$

[6] The instruments are relevant if and only if $rank(\boldsymbol{\Psi}) = k$.

[7] With instruments, $\boldsymbol{\phi}$ includes the reduced-form parameters from the first-stage regression. We do not formalize this to avoid heavier notation.

be the set of indices such that $j \in \mathcal{I}_{\mathcal{S}}$ if some of the impulse responses to the j-th structural shock are sign-constrained. Thus, the set of sign restrictions on the shock j is

$$\boldsymbol{S}_j(\boldsymbol{\phi})\boldsymbol{q}_j \geq \mathbf{0}, \text{ for } j \in \mathcal{I}_{\mathcal{S}}. \tag{2.17}$$

With abuse of notation, let

$$\boldsymbol{S}(\boldsymbol{\phi}, \boldsymbol{Q}) \geq \mathbf{0} \tag{2.18}$$

collect all sign restrictions $\boldsymbol{S}_j(\boldsymbol{\phi})\boldsymbol{q}_j \geq \mathbf{0}$ for any $j \in \mathcal{I}_{\mathcal{S}}$. Equation (2.18) also nests sign constraints on structural objects other than impulse responses, e.g. Arias et al. (2019), ranking restrictions in Amir-Ahmadi and Drautzburg (2021) and sign-restricted factor models (Amir-Ahmadi and Uhlig, 2015; Korobilis, 2022; Stock and Watson, 2016).

2.3.3 The Identified Set for the FEVD

The identified set for the FEVD is

$$IS_{FEVD}(\boldsymbol{\phi}) = \{FEVD_{ijh} : \boldsymbol{Q} \in \boldsymbol{\mathcal{Q}}(\boldsymbol{\phi}|\boldsymbol{F}, \boldsymbol{S})\}, \tag{2.19}$$

where $\boldsymbol{\mathcal{Q}}(\boldsymbol{\phi}|\boldsymbol{F}, \boldsymbol{S})$ is the set of $\boldsymbol{Q}$s that satisfy the zero restrictions (2.16), sign restrictions (2.18) and sign normalizations:

$$\boldsymbol{\mathcal{Q}}(\boldsymbol{\phi}|\boldsymbol{F}, \boldsymbol{S}) = \{\boldsymbol{Q} \in \boldsymbol{\theta}(n) : \boldsymbol{S}(\boldsymbol{\phi}, \boldsymbol{Q}) \geq \mathbf{0}, \boldsymbol{F}(\boldsymbol{\phi}, \boldsymbol{Q}) = \mathbf{0}, diag(\boldsymbol{Q}'\boldsymbol{\Sigma}_{\boldsymbol{tr}}^{-1}) \geq \mathbf{0}\}. \tag{2.20}$$

Definition of the endpoints for $IS_{FEVD}(\boldsymbol{\phi})$ follows.

Definition 2.1 *Given a vector of the reduced-form parameters $\boldsymbol{\phi}$, a shock of interest j^*, $l_{ij^*h}(\boldsymbol{\phi})$ and $u_{ij^*h}(\boldsymbol{\phi})$ are the lower- and upper-bound of $IS_{FEVD}(\boldsymbol{\phi})$, respectively:*

$$l_{ij^*h}(\boldsymbol{\phi}) \equiv min_{\boldsymbol{Q}}(\boldsymbol{Q}\boldsymbol{e}_j)'\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})(\boldsymbol{Q}\boldsymbol{e}_j) \text{ s.t. } \boldsymbol{S}(\boldsymbol{\phi}, \boldsymbol{Q}) \geq \mathbf{0},\ \boldsymbol{F}(\boldsymbol{\phi}, \boldsymbol{Q}) = \mathbf{0},\ diag(\boldsymbol{Q}'\boldsymbol{\Sigma}_{\boldsymbol{tr}}^{-1}) \geq \mathbf{0} \tag{2.21}$$

and

$$u_{ij^*h}(\boldsymbol{\phi}) \equiv max_{\boldsymbol{Q}}(\boldsymbol{Q}\boldsymbol{e}_j)'\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})(\boldsymbol{Q}\boldsymbol{e}_j) \text{ s.t. } \boldsymbol{S}(\boldsymbol{\phi}, \boldsymbol{Q}) \geq \mathbf{0},\ \boldsymbol{F}(\boldsymbol{\phi}, \boldsymbol{Q}) = \mathbf{0},\ diag(\boldsymbol{Q}'\boldsymbol{\Sigma}_{\boldsymbol{tr}}^{-1}) \geq \mathbf{0}, \tag{2.22}$$

where $\boldsymbol{Q}'\boldsymbol{Q} = \boldsymbol{I}_n$. Recall that $\boldsymbol{q}_j = \boldsymbol{Q}\boldsymbol{e}_j$. From now on, $\underline{\boldsymbol{q}}_{j^*}$ and $\bar{\boldsymbol{q}}_{j^*}$ denote the optimizers corresponding to $l_{ij^*h}(\boldsymbol{\phi})$ and $u_{ij^*h}(\boldsymbol{\phi})$, respectively.

In this article, we consider constraints that make the model set-identified:

Assumption A1 *(Set-Identification) Without loss of generality, order the variables such that* $f_1 \geq \ldots f_n \geq 0$. *In case of ties, order the shock of interest first. Condition* $f_i \leq n - i$ *for all* $i = 1, \ldots, n$ *with strict inequality for at least one* $i \in \{1, \ldots, n\}$ *must hold (Rubio-Ramirez et al., 2010).*

This assumption guarantees the set-identification of any structural object and its functions, e.g. IRFs and FEVD. In a Proxy SVAR setting, with a single instrument ($dim(\boldsymbol{m}_t) = 1$) for a single shock, set-identification for the whole $\boldsymbol{Q}$ arises when the instrument is free to be correlated with two shocks at least for $n \geq 3$ (plausibly exogenous instrument).[8] With multiple instruments for multiple shocks, for $n \geq 3$ and $1 < k < n - 1$ set-identification naturally arises for all the columns of $\boldsymbol{Q}$ unless additional zero restrictions are imposed.[9] Set-identification through plausibly exogenous instruments, often combined with sign restrictions, is increasingly getting common (Piffer and Podstawski, 2018; Braun and Brüggemann, 2022; Jarociński and Karadi, 2020; Ludvigson et al., 2021; Fusari, 2023; Caggiano et al., 2021; Caggiano and Castelnuovo, 2023).

The following assumption allows to derive the Karush-Kuhn-Tucker conditions when zero restrictions are imposed on more shocks. If the assumption failed, the results of this paper would still hold for settings where a single shock is zero-constrained.

Assumption A2 *(Zero Restrictions) Assume the order of variable in Assumption A1 and let* j^* *denote the shock of interest.* $j^* \geq 2$ *and* $f_i < n - 1$ *for all* $i = 1, \ldots, j^* - 1$ *must hold.*

Lemma 2.1 technically proves why, under the previous assumption, we can use the Karush-Kuhn-Tucker conditions; the intuition being that, if we are interested in shock j^*, Assumption A2 guarantees that the set of $FEVD_{ij^*h}$ is affected by restrictions on $\boldsymbol{q}_{j^*}$ only, i.e. the Karush-Kuhn-Tucker conditions are a function of the constraints on j^* only.

Lemma 2.1 *Suppose that Assumption A2 holds. Then zero restrictions on shocks* $1, \ldots, j^* - 1, j^* + 1, \ldots, n - 1$ *leave the set of* $FEVD_{ij^*h}$ *unchanged for* $i = 1, \ldots, n$ *and* $h = 0, \ldots$.

Zero constraints on more shocks are typically imposed in Proxy SVARs. Let us focus on the single instrument case -$k = 1$ for shock j^*-. Here, set-identification for the shock would imply that the instrument is correlated with two shocks at least (including j^*): recalling restrictions (2.15) delivers $f_1 = f_2 = \cdots = f_{j^*-1} = 1, f_{j^*} = f_n = 0$. Assumption A2 is satisfied for $n \geq 3$. Other than Proxy SVARs, it is hard to find applications with set-identification where more shocks are zero-restricted; for example, the popular exercise in Arias et al. (2019) zero-restricts monetary policy shock only. Thus, we believe that Assumption A2 covers most of the

[8] For $k = 1$ and exogenous instrument, $f_i = 1$ for $i = 1, \ldots, n - 1$ and $f_n = 0$: columns $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{n-1}$ are set-identified, while $\boldsymbol{q}_n$ is point-identified.

[9] For $k = n - 1$, $f_1 = n - 1$ and $f_i = 0$ for $i = 2, \ldots, n$. As a result, $\boldsymbol{q}_1$ is point-identified; the other columns are set-identified.

empirically relevant cases. A by-product of Lemma 2.1 is that the toolkit in Gafarov et al. (2018) can be employed with zero constraints on more shocks under Assumption A2.

Given a reduced-form parameters vector $\boldsymbol{\phi}$, we also assume that the identified set of the FEVD is non-empty, i.e. identifying assumptions do not contradict each other and are not rejected at $\boldsymbol{\phi}$.

Assumption A3 *(Non-Emptiness) $IS_{FEVD}(\boldsymbol{\phi})$ as defined in equation (2.19) is non-empty at $\boldsymbol{\phi}$.*

Relevant literature provides several frameworks to check for non-emptiness (Giacomini and Kitagawa, 2021; Giacomini et al., 2022b; Amir-Ahmadi and Drautzburg, 2021; Uhlig, 2005; Granziera et al., 2018; Arias et al., 2018).

Our characterization of the endpoints relies on using the Karush-Kuhn-Tucker conditions. The following assumption is therefore needed:

Assumption A4 *(Linear Independence) Given a constrained shock j^*, $\boldsymbol{F}_{j^*}(\boldsymbol{\phi})$ and $\boldsymbol{S}_{j^*}(\boldsymbol{\phi})$ are linearly independent at $\boldsymbol{\phi}$.*

This assumption is common in the relevant literature, e.g. Gafarov et al. (2018) and Giacomini and Kitagawa (2021). To the best of our knowledge, Proxy SVARs with weak instruments is the only empirically relevant setting where the assumption is not satisfied. The following example illustrates our reasoning.

Example 2.1 *(Weak Proxy) Suppose that there are no sign restrictions and the only zero restrictions come from set-identifying exogeneity constraints. In this case, we would have*

$$\boldsymbol{F}_{j^*}(\boldsymbol{\phi}) = \boldsymbol{D}\boldsymbol{\Sigma}_{tr}, \tag{2.23}$$

where $\boldsymbol{F}_{j^*}(\boldsymbol{\phi})$ is a $k \times n$ matrix. The relevance condition $rank(\boldsymbol{\Psi}) = k$ holds if and only if $rank(\boldsymbol{D}) = k$. Suppose that the instrument is weak. This leads to

$$rank(\boldsymbol{\Psi}) < k \Rightarrow rank(\boldsymbol{D}) < k \Rightarrow rank(\boldsymbol{D}\boldsymbol{\Sigma}_{tr}) < k. \tag{2.24}$$

because $rank(\boldsymbol{D}\boldsymbol{\Sigma}_{tr}) \leq min\{rank(\boldsymbol{D}), rank(\boldsymbol{\Sigma}_{tr})\}$. This implies that $\boldsymbol{F}_{j^*}(\boldsymbol{\phi})$ is rank deficient and Assumption A4 fails.

3 Estimation and Inference for Set-Identified FEVD

This section describes the main results. Section 3.1 provides the characterization of the FEVD endpoints and proposes a consistent estimator. Section 3.2 proves the differentiability of the endpoints. Section 3.3 (i) constructs a set-length adjusted delta-method confidence interval that is both uniformly consistent in level and has asymptotic Bayesian interpretation and (ii) proposes a further confidence interval for users mostly interested in the inference over the set of the FEVD (rather than the FEVD itself).

3.1 Estimation

Proposition 3.1 characterizes the set of the FEVD up to a set of active, i.e. binding, constraints. To put it another way, if we knew the active constraints, we would be able to characterize the set of the FEVD. As a result, evaluating the set at different active constraints and checking the primal feasibility leads to the optimizers and endpoints of the FEVD for problems in (2.21) and (2.22) (Theorem 3.1).

First, let us introduce some more definitions. Let $\boldsymbol{r}(\boldsymbol{\phi})$ be the $m \times n$ matrix collecting the gradient vectors of the m constraints that are active at an optimizer of problem (2.21)-(2.22). The m rows of $\boldsymbol{r}(\boldsymbol{\phi})$ consist of the f_{j^*} rows of the matrix $\boldsymbol{F}_{j^*}(\boldsymbol{\phi})$ and s_{rj^*} out of the s_{j^*} rows from $\boldsymbol{S}_{j^*}(\boldsymbol{\phi})$, i.e. $m = f_{j^*} + s_{rj^*}$. In particular, there are $\sum_{i=0}^{min(n-1-f_{j^*}, s_{j^*})} \frac{s_{j^*}!}{i!(s_{j^*}-i)!}$ possible combinations of active constraints, i.e. possible ways to construct $\boldsymbol{r}(\boldsymbol{\phi})$.

Definition 3.1 *Assume that a single shock j^* is sign-constrained and $\boldsymbol{r}(\boldsymbol{\phi})$ collects the gradient vectors of the active constraints. $l_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r})$ and $u_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r})$ are defined as follows:*

$$l_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) \equiv min_{\boldsymbol{q}_{j^*}} \boldsymbol{q}_{j^*}' \boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi}) \boldsymbol{q}_{j^*} \;\; s.t. \;\; \boldsymbol{r}(\boldsymbol{\phi}) \boldsymbol{q}_{j^*} = \boldsymbol{0}, \tag{3.1}$$

and

$$u_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) \equiv max_{\boldsymbol{q}_{j^*}} \boldsymbol{q}_{j^*}' \boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi}) \boldsymbol{q}_{j^*} \;\; s.t. \;\; \boldsymbol{r}(\boldsymbol{\phi}) \boldsymbol{q}_{j^*} = \boldsymbol{0}, \tag{3.2}$$

where $\boldsymbol{q}_{j^*}' \boldsymbol{q}_{j^*} = 1$. Let $\underline{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ and $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ denote the optimizer of problem (3.1) and (3.2), respectively.

The next Proposition solves the quadratic programming in (3.1) and (3.2).

Proposition 3.1 *Suppose that Assumptions A1-A4 hold and a single shock j^* is sign-constrained. Then*

$$u_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) = \lambda_{max}(\boldsymbol{\phi}, \boldsymbol{r}) = max\{\lambda_1(\boldsymbol{\phi}, \boldsymbol{r}), \ldots, \lambda_n(\boldsymbol{\phi}, \boldsymbol{r})\}, \tag{3.3}$$

with

$$det\left[\boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r}) - \lambda_\tau(\boldsymbol{\phi}, \boldsymbol{r}) \boldsymbol{I}_n\right] = 0 \; for \; all \; \tau = 1, \ldots, n \tag{3.4}$$

and

$$\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r}) = \pm eignvct(\lambda_{max}(\boldsymbol{\phi}, \boldsymbol{r}), \boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r})), \tag{3.5}$$

where $\boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r}) = [\boldsymbol{I}_n - \boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r})] \boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})$, $\boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r}) = \boldsymbol{r}(\boldsymbol{\phi})' \left[\boldsymbol{r}(\boldsymbol{\phi}) \boldsymbol{r}(\boldsymbol{\phi})'\right]^{-1} \boldsymbol{r}(\boldsymbol{\phi})$ and $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ is the eigenvector associated to $\lambda_{max}(\boldsymbol{\phi}, \boldsymbol{r})$. We can obtain the minimum as follows:
$l_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) = \lambda_{min}(\boldsymbol{\phi}, \boldsymbol{r}) = min\{\lambda_1(\boldsymbol{\phi}, \boldsymbol{r}), \ldots, \lambda_n(\boldsymbol{\phi}, \boldsymbol{r})\}$ for $\lambda_{min}(\boldsymbol{\phi}, \boldsymbol{r}) \neq 0$.

For $\lambda_{min}(\boldsymbol{\phi}, \boldsymbol{r}) = 0$*:*

$$l_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) = \begin{cases} 0 \ \textit{if} \ \exists \boldsymbol{q}_{j^*} | FEVD_{ij^*h} = 0 \\ \lambda^+_{min}(\boldsymbol{\phi}, \boldsymbol{r}) \ \textit{otherwise}, \end{cases} \tag{3.6}$$

where $\lambda^+_{min}(\boldsymbol{\phi}, \boldsymbol{r})$ *is the smallest non-zero eigenvalue.*

In other words, the bounds of the FEVD subject to active constraints are given by the maximum and minimum eigenvalues of the $n \times n$ matrix $\boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r})$. Note that, in absence of constraints, $\boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r}) \equiv \boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})$. Given the characterization of the bounds as eigenvalues, an associated feasible eigenvector can be found, i.e. both $eignvct(\lambda_{max}(\boldsymbol{\phi}, \boldsymbol{r}), \boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r}))$ and $-eignvct(\lambda_{max}(\boldsymbol{\phi}, \boldsymbol{r}), \boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r}))$ are associated to the same eigenvalue. For $\lambda_{min}(\boldsymbol{\phi}, \boldsymbol{r}) = 0$, the Karush-Kuhn-Tucker conditions are not always satisfied. As a result, if a column vector $\boldsymbol{q}_j$ delivering $FEVD_{ij^*h} = 0$ does not exist, the lower bound is given by the smallest non-zero eigenvalue. On the other hand, $\lambda_{max}(\boldsymbol{\phi}, \boldsymbol{r}) \neq 0$ from Assumption A1.

In the technical proof, we rely on Rao (1964) to show that the eigenvalues of $\boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r})$ are equivalent to those of the symmetric matrix

$$\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}, \boldsymbol{r}) = \boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})^{\frac{1}{2}} \left[\boldsymbol{I}_n - \boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r})\right] \boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})^{\frac{1}{2}'}. \tag{3.7}$$

In practice, we can use $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}, \boldsymbol{r})$ for the calculation of the eigenvalues of $\boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r})$. On top of obvious computational gains, this will allow us to derive the differentiability of the endpoints (Section 3.2). The relationship between the optimizer $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ and the eigenvector $\bar{\tilde{\boldsymbol{q}}}_{j^*}(\boldsymbol{r})$ of $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})$ is the following:

$$\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r}) = \{\bar{\tilde{\boldsymbol{q}}}'_{j^*}(\boldsymbol{r}) \left[\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\right]^{-1} \bar{\tilde{\boldsymbol{q}}}_{j^*}(\boldsymbol{r})\} \left[\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\right]^{-\frac{1}{2}} \bar{\tilde{\boldsymbol{q}}}_{j^*}(\boldsymbol{r}). \tag{3.8}$$

For $n \leq 4$, the extreme eigenvalues are analytically available by solving the characteristic polynomial for $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}, \boldsymbol{r})$. For $n > 4$, generally this is not the case; we can use the hypergeometric functions[10] to get an analytical solution or, more conveniently, we can approximate it by using well-established numerical methods for eigenvalues.

Given the Proposition 3.1, we can obtain the solution to the full problems (2.21) and (2.22) as follows: compute the maximum and minimum eigenvalues (and the associated eigenvectors) for all possible $\boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r})$ matrices, i.e. for all possible combination of active constraints. The endpoints of the FEVD are the largest and smallest eigenvalues,[11] while the associated eigenvector satisfying the inactive constraints is the optimizer. The next Theorem formalizes this result.

[10]See the survey in Erdélyi et al. (1953) and Daalhuis (2010).

[11]For $\lambda_{min}(\boldsymbol{\phi}, \boldsymbol{r}) = 0$, condition in (3.6) applies.

Theorem 3.1 *Suppose that Assumptions A1-A4 hold and a single shock j^* is sign-constrained. Then*

$$u_{ij^*h}(\boldsymbol{\phi}) = max_{\boldsymbol{r}(\boldsymbol{\phi})} u_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) \tag{3.9}$$

and

$$\bar{\boldsymbol{q}}_{j^*} = eignvct(u_{ij^*h}(\boldsymbol{\phi}), \boldsymbol{Z}(\boldsymbol{\phi})), \tag{3.10}$$

where $\bar{\boldsymbol{q}}_{j^*} \in \mathcal{Q}(\phi|\boldsymbol{F}, \boldsymbol{S})$, $\boldsymbol{Z}(\boldsymbol{\phi}) \equiv \boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r^*})$, $\boldsymbol{r^*}(\boldsymbol{\phi}) = argmax_{\boldsymbol{r}(\boldsymbol{\phi})} u_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r})$. $l_{ij^*h}(\boldsymbol{\phi})$ and $\underline{\boldsymbol{q}}_{j^*}$ can be obtained similarly.

Previous results are expressed in population values, e.g. $\boldsymbol{\phi}$. We now turn our attention to the estimation. Let $\rightarrow_p$ denote the standard convergence in probability; let $\widehat{\ }$ represent the estimated values; let P denote a data-generating process (DGP).

Assumption A5 *(Simple eigenvalues) The algebraic multiplicity of the eigenvalues delivering $u_{ij^*h}(\boldsymbol{\phi})$ and $l_{ij^*h}(\boldsymbol{\phi})$ is equal to 1.*

The assumption above guarantees that $u_{ij^*h}(\boldsymbol{\phi})$ is a simple, i.e. unique, eigenvalue. The same applies to $l_{ij^*h}(\boldsymbol{\phi})$. Finding non-simple eigenvalues is usually very uncommon. Let us stress that we require the simplicity of the extreme eigenvalues only, without excluding the possibility of multiplicities for the $n-2$ eigenvalues of each of all the other possible combinations of active constraints.

Proposition 3.2 *If Assumptions A1-A5 hold, a single shock j^* is sign-constrained and $\widehat{\boldsymbol{\phi}} \rightarrow_p \boldsymbol{\phi}(P)$, then*

$$u_{ij^*h}(\widehat{\boldsymbol{\phi}}) \rightarrow_p u_{ij^*h}(\boldsymbol{\phi}(P)). \tag{3.11}$$

The same applies to $l_{ij^*h}(\widehat{\boldsymbol{\phi}})$.

The proof in the Appendix A relies on basic statistics as long as the reduced-form estimator is consistent and $T \ggg n$. The latter is fairly uncontroversial in SVARs, so we did not explicit it as a formal assumption. Should not this be the case, results from Principal Component Analysis literature can be evoked to get consistency. It is also possible to show that, if Assumption A5 failed, Proposition 3.2 would still hold.[12] However, that assumption simplifies the proof and is needed for the differentiability result.

[12] We would like to thank Henrique Castro-Pires for pointing that out.

3.2 Differentiability

This section presents the differentiability of the FEVD endpoints. It is a novel result in the literature; it is interesting per se and instrumental to construct a delta-method confidence interval. Also, a by-product of this finding is that the robust Bayesian approach in GK21 has frequentist validity when applied to the FEVD. In order to derive the differentiability of FEVD endpoints, we require one more condition.

Assumption A6 *(Differentiability) $F_{j^*}(\boldsymbol{\phi})$ and $S_{j^*}(\boldsymbol{\phi})$ are differentiable at $\boldsymbol{\phi}$.*

This is a very standard smoothness assumption we share with, among others, Gafarov et al. (2018) and GK21. We are unaware of any identifying constraints that do not meet this condition.

Our differentiability result is based on the intuition of Magnus (1985): under uniqueness of eigenvalues, the implicit function theorem states that there is a neighborhood of $\boldsymbol{\phi}$ where $u_{ij^*h}(\boldsymbol{\phi})$ and $l_{ij^*h}(\boldsymbol{\phi})$ and the associated optimizers exist and are continuously differentiable. This consideration, combined with the application of the chain rule, delivers the result in Theorem 3.2, which holds for $FEVD_{ij^*h} \in (0, 1)$.

Theorem 3.2 *Suppose that the Assumptions A1-A6 hold and a single shock j^* is sign-constrained. Then $u_{ij^*h}(\boldsymbol{\phi})$ is differentiable, with*

$$\frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}} = \bar{\bar{\boldsymbol{q}}}'_{j^*} \otimes \bar{\bar{\boldsymbol{q}}}'_{j^*} \frac{\partial vec(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))}{\partial \boldsymbol{\phi}}, \tag{3.12}$$

where $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}) \equiv \tilde{\boldsymbol{Z}}(\boldsymbol{\phi}, \boldsymbol{r^*})$.

Section 3.3 builds upon Theorem 3.2 to present a confidence interval for $FEVD_{ij^*h}$. Also, the theorem above leads to prove that the robust Bayesian approach in GK21 has frequentist validity when applied to the FEVD.

Giacomini and Kitagawa (2021). *Under some regularity conditions, frequentist validity of the robust Bayesian toolkit in GK21 requires $IS_{FEVD}(\boldsymbol{\phi})$ to be convex, continuous and differentiable (at $\boldsymbol{\phi}$). Under the restrictions considered in this paper, $IS_{FEVD}(\boldsymbol{\phi})$ is convex as long as is non-empty, i.e. $IS_{FEVD}(\boldsymbol{\phi})$ is convex whenever the set for r_{ij^*h}, i.e. IRF, is convex (formally, see Section 3.1 in Giacomini et al. (2022a)). Continuity requires mild and easily verifiable conditions on the reduced-form matrix of zero and sign restrictions (Proposition B2 in the Appendix of GK21 and the consideration that, given $\boldsymbol{\phi}$, the set for $FEVD_{ij^*h}$ is continuous whenever the set for r_{ij^*h} is continuous). We provide the missing piece, i.e. differentiability, under which the robust credible interval in GK21 is an asymptotically valid confidence set for the true identified set. The next corollary formalizes the result. Let $\rightarrow_{as}$ and $\rightarrow_{d}$ denote the almost sure convergence and convergence in distribution; CI_{α}^{GK} is the credible region with credibility α proposed in GK21.*

Corollary 3.1 *Suppose that Assumptions A1-A6 hold,* $\frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ *and* $\frac{\partial l_{ij^*h}(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ *are different from zero, a single shock* j^* *is sign-constrained,* $\widehat{\boldsymbol{\phi}} \to_{as} \boldsymbol{\phi}(P)$*, and the posterior of* $\boldsymbol{\phi}$ *and the sampling distribution of* $\widehat{\boldsymbol{\phi}}$ *are* $\sqrt{T}$*-asymptotically normal with an identical covariance matrix* $\boldsymbol{\Omega}$*:*

$$\sqrt{T}(\boldsymbol{\phi}(P) - \widehat{\boldsymbol{\phi}})|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T \to_d \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Omega}(P)) \text{ as } T \to \infty,$$
$$\sqrt{T}(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}(P))|\boldsymbol{\phi}(P) \to_d \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Omega}(P)) \text{ as } T \to \infty.$$

Then CI_α^{GK} *is an asymptotically valid frequentist confidence set for* $IS_{FEVD}(\boldsymbol{\phi}(P))$*:*

$$lim_{T\to\infty} Pr(IS_{FEVD}(\boldsymbol{\phi}(P)) \subset CI_\alpha^{GK}) = 1 - \alpha. \tag{3.13}$$

Unlike Imbens and Manski (2004), Stoye (2009), Gafarov et al. (2018), Gafarov et al. (2016), Granziera et al. (2018) and this paper, the frequentist coverage in the approach of GK21 is for the true identified set rather than for the true value of the parameter, i.e. $FEVD_{ij^*h}$*. This implies that* CI_α^{GK} *is asymptotically wider than our proposal of confidence interval for* $FEVD_{ij^*h}$*. Also, Corollary 3.1 shows point-wise coverage for* CI_α^{GK}*, while the next section establishes uniform coverage for our confidence interval.*

3.3 Inference

This section proposes a delta-method interval for $FEVD_{ij^*h}$, showing that it has asymptotically frequentist validity and Bayesian equivalence:

$$CI_\alpha \equiv \left[l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) - c_\alpha \widehat{\sigma}_{lij^*h}/\sqrt{T}, u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) + c_\alpha \widehat{\sigma_{uij^*h}}/\sqrt{T}\right], \tag{3.14}$$

where $\widehat{\boldsymbol{\phi}}_{OLS} \equiv (\widehat{\boldsymbol{b}}'_{OLS}, vec(\widehat{\boldsymbol{B}}_{OLS})', vech(\widehat{\boldsymbol{\Sigma}}_{OLS})')$ is the OLS estimator of $\boldsymbol{\phi}$. In particular, $\widehat{\boldsymbol{B}}_{OLS} \equiv \left(\frac{1}{T}\sum_{t=1}^T \boldsymbol{y}_t \boldsymbol{x}'_t\right)\left(\frac{1}{T}\sum_{t=1}^T \boldsymbol{x}_t \boldsymbol{x}'_t\right)^{-1}$, $\widehat{\boldsymbol{\Sigma}}_{OLS} \equiv \frac{1}{T-np-1}\sum_{t=1}^T \widehat{\boldsymbol{u}}_t \widehat{\boldsymbol{u}}'_t$, $\widehat{\boldsymbol{b}}_{OLS} = \bar{\boldsymbol{y}}_t - \widehat{\boldsymbol{B}}_{OLS}\bar{\boldsymbol{x}}_t$, with $\widehat{\boldsymbol{u}}_t \equiv \boldsymbol{y}_t - \widehat{\boldsymbol{b}}_{OLS} - \widehat{\boldsymbol{B}}_{OLS}\boldsymbol{x}_t$, and $\bar{\boldsymbol{y}}_t$ and $\bar{\boldsymbol{x}}_t$ being the sample means. Our formula for the standard errors is the following:

$$\widehat{\sigma}_{uij^*h} = \left[\left(\frac{\partial u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})}{\partial \widehat{\boldsymbol{\phi}}_{OLS}}\right) \widehat{\boldsymbol{\Omega}} \left(\frac{\partial u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})}{\partial \widehat{\boldsymbol{\phi}}_{OLS}}\right)'\right]^{\frac{1}{2}}, \tag{3.15}$$

where $\widehat{\boldsymbol{\Omega}}$ is the estimated variance-covariance matrix of $\boldsymbol{\phi}$[13] and $\widehat{\sigma}_{lij^*h}$ is defined similarly. We propose c_α solving the following:

$$\Phi\left(c_\alpha + \frac{\sqrt{T}\widehat{\Delta}_{ij^*h}}{max\{\widehat{\sigma}_{lij^*h}, \widehat{\sigma}_{uij^*h}\}}\right) - \Phi(-c_\alpha) = 1 - \alpha, \tag{3.17}$$

[13] A formula for $\widehat{\boldsymbol{\Omega}}$ under non-serial correlation (but heteroskedasticity) of $\boldsymbol{u}_t$ is

$$\widehat{\boldsymbol{\Omega}} \equiv \frac{1}{T}\sum_{t=1}^T \left[vec(\hat{\boldsymbol{u}}_t \boldsymbol{x}'_t)', vech(\hat{\boldsymbol{u}}_t \hat{\boldsymbol{u}}'_t - \hat{\boldsymbol{\Sigma}})'\right]' \left[vec(\hat{\boldsymbol{u}}_t \boldsymbol{x}'_t)', vech(\hat{\boldsymbol{u}}_t \hat{\boldsymbol{u}}'_t - \hat{\boldsymbol{\Sigma}})'\right]. \tag{3.16}$$

The formula above can be adjusted to take into account serial correlation.

where $\Phi(\bullet)$ is the standard normal cumulative distribution evaluated at $\bullet$ and $\widehat{\Delta}_{ij^*h} = u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ is the estimated length of the identified set.

In studying the asymptotic properties of the confidence interval (Section 3.3.1 and 3.3.2), we rule out the case where the partial derivatives are zero. This corresponds to degenerate zero-covariance Normal distribution, e.g. bounds of the set, where the coverage can be non-optimal. This is common for any interval-based inference rather than being a specific feature of our toolkit.

3.3.1 Frequentist Coverage

Frequentist coverage result requires the asymptotic normality of $\widehat{\boldsymbol{\phi}}_{OLS}$ and the consistency of its variance-covariance matrix.

Assumption A7 *(Asymptotic Normality) OLS estimators uniformly satisfy*

$$\sqrt{T}(\widehat{\boldsymbol{\phi}}_{OLS} - \boldsymbol{\phi}(P)) \to_d \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Omega}(P)), \tag{3.18}$$

$$\widehat{\boldsymbol{\Omega}} \to_p \boldsymbol{\Omega}(P). \tag{3.19}$$

In practice, we are proposing an adjusted (by the length of the set) delta-method interval for a parameter bounded by extreme eigenvalues.[14] Given the arguments in Imbens and Manski (2004) and Stoye (2009), the identified set length adjustment is the key to show that, under some conditions, CI_α is uniformly consistent, i.e. $lim_{T\to\infty}\, inf\; inf_{FEVD_{ij^*h}(P)\in IS(\boldsymbol{\phi}(P))} Pr(FEVD_{ij^*h}(P) \in CI_\alpha) = 1-\alpha$.

We focus on the structural object (parameter) of interest, i.e. $FEVD_{ij^*h}$, not its set. GK21 derived asymptotic point-wise coverage for the identified set rather than for the structural object. As a result, their frequentist coverage is asymptotically more conservative for the FEVD than our confidence interval. On the other hand, in finite samples, the Monte-Carlo exercise shows the coverage of their approach for the FEVD is problematic.

For set-identified impulse responses, Gafarov et al. (2018) generated delta-method point-wise consistent confidence intervals. In principle, users could take advantage of the relationship between IRFs and FEVD, e.g. equation (2.8), and construct a plug-in estimator and confidence interval for the FEVD from Gafarov et al. (2018). However, this is not to recommend: the FEVD at a given horizon is a non-linear function of IRFs in previous periods. For each horizon, Gafarov et al. (2018) characterize the endpoints of the IRFs by finding the optimal rotation matrix, i.e. structural model. The plug-in estimator for the FEVD would come from

[14] Rao (1973) was the first one to present the delta-method for eigenvalues analysis. Further references include Carter et al. (1986), who produced a conservative Scheffé-type interval; Carter et al. (1990) introduced the delta-method for stationary points of a quadratic response surface; Peterson (1993) discussed the inference for eigenvalues subject to constraints; Bisgaard and Ankenman (1996) introduced a 2-regression approach delivering confidence intervals asymptotically equivalent to the delta-method inference proposed by Carter et al. (1990).

a multiplicity of structural models (rotation matrices), resulting in a loss of interpretability. The same would apply to the usage of the toolkit in Granziera et al. (2018) for estimation of the FEVD.

The following theorem formalizes the uniform consistency of CI_α:

Theorem 3.3 *In a compact subset of* $\boldsymbol{\Xi}$*, suppose that Assumptions A1-A7 hold and* $\frac{\partial u_{ij^*h}(\widehat{\phi}_{OLS})}{\partial \widehat{\phi}_{OLS}}$ *and* $\frac{\partial l_{ij^*h}(\widehat{\phi}_{OLS})}{\partial \widehat{\phi}_{OLS}}$ *are different from zero. We obtain*

$$lim_{T\to\infty}\ inf\ inf_{FEVD_{ij^*h}(P)\in IS(\phi(P))} Pr(FEVD_{ij^*h}(P)\in CI_\alpha) = 1-\alpha. \tag{3.20}$$

The proof builds upon the uniform convergence in distribution of the delta-method under the conditions in Kasy (2018) and the arguments in Imbens and Manski (2004) and Stoye (2009).

Our toolkit focuses on inference over single scalars. However, Inoue and Kilian (2022) stressed that this approach can be invalid because it ignores the mutual dependence of the structural object of interest (in this case, FEVD) across variables and horizons. Applying a standard Bonferroni correction, e.g. Bisgaard and Ankenman (1996), would easily provide joint inference.

Gafarov et al. (2018). *While Gafarov et al. (2018) showed that a standard delta-method confidence interval for response functions* r_{ijh} *is point-wise consistent, we argue that the following adjusted delta-method confidence interval for response functions is uniformly consistent:*

$$CI_\alpha^r \equiv \left[l_{ij^*h}^r(\widehat{\phi}_{OLS}) - c_\alpha^r \widehat{\sigma}_{lij^*h}^r/\sqrt{T}, u_{ij^*h}^r(\widehat{\phi}_{OLS}) + c_\alpha^r \widehat{\sigma}_{uij^*h}^r/\sqrt{T} \right], \tag{3.21}$$

where $l_{ij^*h}^r(\widehat{\phi}_{OLS})$*,* $u_{ij^*h}^r(\widehat{\phi}_{OLS})$*,* $\widehat{\sigma}_{lij^*h}^r$*, and* $\widehat{\sigma}_{uij^*h}^r$ *are the bounds of* r_{ijh} *and the standard errors as derived in Gafarov et al. (2018). While they consider a standard critical value, we propose to take set-length into account:*

$$\Phi\left(c_\alpha^r + \frac{\sqrt{T}\widehat{\Delta}_{ij^*h}^r}{max\{\widehat{\sigma}_{lij^*h}^r, \widehat{\sigma}_{uij^*h}^r\}}\right) - \Phi\left(-c_\alpha^r\right) = 1-\alpha, \tag{3.22}$$

where $\widehat{\Delta}_{ij^*h}^r$ *is the estimated length of the identified set for the impulse response function* r_{ij^*h}*. Under a standard critical value, the confidence interval is point-wise consistent. The next corollary formalizes that the adjusted critical value delivers a uniformly consistent* CI_α^r*.*

Corollary 3.2 *In a compact subset of* $\boldsymbol{\Xi}$*, suppose that Assumptions A1, A2, A3, A4, A6, A7 hold and the derivatives of the endpoints of* r_{ij^*h} *are non-zero. We obtain*

$$lim_{T\to\infty}\ inf\ inf_{r_{ij^*h}(P)\in IS^r(\phi(P))} Pr(r_{ij^*h}(P)\in CI_\alpha) = 1-\alpha, \tag{3.23}$$

where $IS^r(\boldsymbol{\phi}(P))$ *is the identified set for* $r_{ij^*h}(P)$*.*

3.3.2 Bayesian Interpretation

We now turn our attention to the Bayesian interpretation of the proposed confidence interval: under some conditions over the prior distribution of the reduced-form VAR, the posterior bounds of the FEVD are asymptotically centered at $u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ and $l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$. Empirical macroeconomists usually employ Bayesian methods, so we believe that a Bayesian reading of our toolkit can be useful to practitioners.

Let $P^*_{\boldsymbol{\theta}}$ denote a prior distribution for the structural parameters $\boldsymbol{\theta}$. It is well-established that $P^*_{\boldsymbol{\theta}} \equiv P^*_{\boldsymbol{\phi}} P^*_{\boldsymbol{Q}|\boldsymbol{\phi}}$, where $P^*_{\boldsymbol{\phi}}$ and $P^*_{\boldsymbol{Q}|\boldsymbol{\phi}}$ are the prior specification for $\boldsymbol{\phi}$ and $\boldsymbol{Q}|\boldsymbol{\phi}$, respectively. $\mathcal{P}(P^*_{\boldsymbol{\phi}})$ represents the class of prior distributions such that $\boldsymbol{\phi}^* \sim P^*_{\boldsymbol{\phi}}$.

We assume that $P^*_{\boldsymbol{\phi}}$ and the DGP P satisfy the Bernstein-von Mises Theorem in probability (Ghosal et al., 1995).

Assumption A8 *(Bernstein-von Mises Theorem)*

$$sup_{\boldsymbol{BO} \in \boldsymbol{\mathcal{BO}}(\mathcal{R}^d)} \left\{ P^*_{\boldsymbol{\phi}} \left(\sqrt{T}(\boldsymbol{\phi}^* - \widehat{\boldsymbol{\phi}}_{OLS}) \in \boldsymbol{BO} | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_T \right) - \mathcal{P}(\boldsymbol{\xi}(P) \in \boldsymbol{BO}) \right\} \to_p 0, \tag{3.24}$$

where $\boldsymbol{\xi}(P) \sim N(\mathbf{0}, \boldsymbol{\Omega}(P))$ *and* $\boldsymbol{\mathcal{BO}}(\mathcal{R}^d)$ *is the set of all Borel measurable sets in* $\mathcal{R}^d$*.*

This assumption is fairly unrestrictive: in a VAR setting, if the reduced-form errors are i.i.d. Gaussian and $P^*_{\boldsymbol{\phi}}$ is continuous at $\boldsymbol{\phi}$, Assumption A8 is met.[15] For example, researchers often use a Normal-Inverse Wishart prior on $\boldsymbol{\phi}$ with Gaussian i.i.d. errors. While we fully recognize that non-Gaussian SVAR is an interesting literature, that is mostly employed for achieving point-identification.

The next Theorem formally establishes the asymptotic equivalence between our frequentist setting and a Bayesian framework:

Theorem 3.4 *Suppose that Assumptions A1-A8 hold. Then*

$$sup_{\boldsymbol{BO} \in \boldsymbol{\mathcal{BO}}(\mathcal{R}^d)} \left\{ P^*_{\boldsymbol{\phi}} \left(\sqrt{T} \left(\begin{bmatrix} l_{ij^*h}(\boldsymbol{\phi}^*) \\ u_{ij^*h}(\boldsymbol{\phi}^*) \end{bmatrix} - \begin{bmatrix} l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \\ u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \end{bmatrix} \right) \in \boldsymbol{BO} | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_T \right) - \mathcal{P}(\tilde{\boldsymbol{\xi}}(P) \in \boldsymbol{BO}) \right\} \to_p 0, \tag{3.25}$$

where $\tilde{\boldsymbol{\xi}}(P) \sim \boldsymbol{N} \left(\mathbf{0}, \begin{bmatrix} \sigma^2_{lij^*h}(P) & 0 \\ 0 & \sigma^2_{uij^*h}(P) \end{bmatrix} \right)$.

3.3.3 Inference Over the Identified Set

Note that the confidence interval (3.14) is not necessarily valid if the user is interested in inference over the set. As a result, this section proposes an alternative confidence interval,

[15] See Theorem 1 and 2 in Ghosal et al. (1995).

which is asymptotically valid for the set of FEVD (but conservative for the FEVD itself). This naively corresponds to consider the critical value from a standard Normal distribution, i.e. critical value without any adjustment:

$$CI_{\alpha}^{*} \equiv \left[l_{ij^*h}(\widehat{\phi}_{OLS}) - c_{\alpha}^{*} \widehat{\sigma}_{lij^*h}/\sqrt{T}, u_{ij^*h}(\widehat{\phi}_{OLS}) + c_{\alpha}^{*} \widehat{\sigma_{uij^*h}}/\sqrt{T} \right], \tag{3.26}$$

where $\Phi(c_{\alpha}^{*}) - \Phi(-c_{\alpha}^{*}) = 1 - \alpha$. A simple Bonferroni argument shows that the confidence interval in (3.26) (i) is asymptotically valid for the set, i.e. it has asymptotic coverage of at least $1 - \alpha$, and (ii) is conservative for the parameter. We refer the technical reader to the argument in Imbens and Manski (2004) and Stoye (2009) for the formal statement of the result. The Monte-Carlo simulation provides some evidence that those features tend to hold in finite samples as well.

4 Empirical Application

We illustrate our toolkit with an empirical application based on a combination of sign and zero restrictions on the IRFs. In particular, we identify an expansionary unconventional monetary policy (QE, Quantitative Easing) shock in the US following the application in Gafarov et al. (2018), based on monthly data sourced from Gertler and Karadi (2015) that span the period July 1979 until August 2008. The variables included in the model, along with the identifying restrictions, are listed in Table 1. All variables are in first-differences, and Consumer Price Index and Industrial Production are included in logs.[16]

Variable	**Impact Restrictions**	**Notation**
Consumer Price Index (CPI_t)	≥ 0	$(\boldsymbol{e}_1'\boldsymbol{\Sigma}_{tr}^{-1})\boldsymbol{q}_1 \geq 0$
Industrial Production (IP_t)	≥ 0	$(\boldsymbol{e}_2'\boldsymbol{\Sigma}_{tr}^{-1})\boldsymbol{q}_1 \geq 0$
Fed Funds Rate (FF_t)	$= 0$	$(\boldsymbol{e}_3'\boldsymbol{\Sigma}_{tr}^{-1})\boldsymbol{q}_1 = 0$
2 Year Treasury Yield ($2YT_t$)	≤ 0	$(\boldsymbol{e}_4'\boldsymbol{\Sigma}_{tr}^{-1})\boldsymbol{q}_1 \leq 0$

Table 1: Set-Identification of an Unconventional Monetary Policy Shock

We impose that an expansionary QE shock generates a non-positive response for 2 year Treasury yields. On the contrary, CPI inflation and Industrial production are assumed to respond non-negatively, while the Fed Funds rate is assumed to not respond at all on impact. This identification scheme yields 7 different possible combinations of binding restrictions. Using these to derive $\boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r})$ never results in an empty set when taking the eigen-decomposition, satisfying Assumption A3. Additionally, the matrices $\boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r})$ are found to always be full

[16] The data can be obtained from the Gertler and Karadi (2015) replication package, available online. As in Gertler and Karadi (2015) and Gafarov et al. (2018), the number of lags is set equal to $p = 12$.

rank, implying that the binding restrictions are always linearly independent at ϕ. As a result, Assumption A4 also holds.

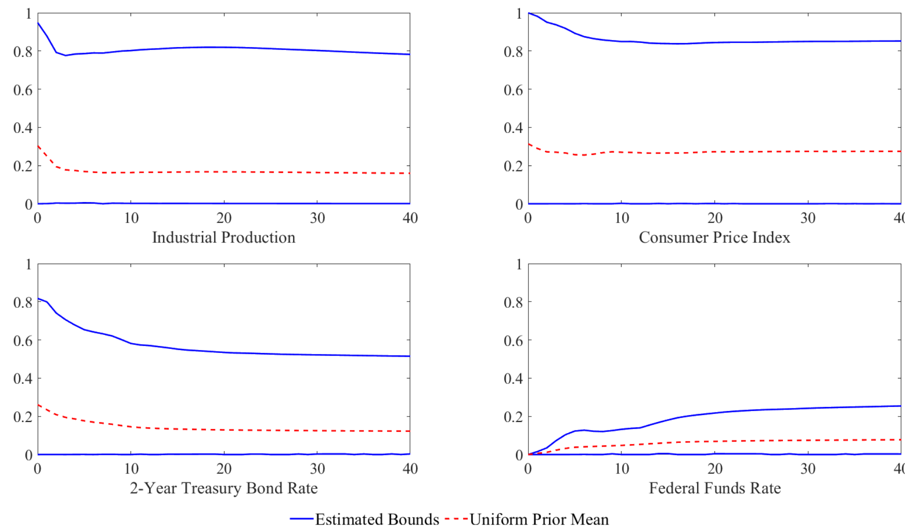

Figure 1: Estimated Bounds *vs.* Mean Under Uniform Prior on $\boldsymbol{Q}$.

Figure 1 compares our approach with the standard Bayesian method, i.e. uniform prior on the rotation matrices $\boldsymbol{Q}$.[17] Overall, under our tooklit the estimated intervals are quite large and range from 0 to almost 1 on impact, with the exception of for the Fed Funds rate, which is zero by assumption. While the lower bounds remain close to zero for all horizons, the upper bounds decrease over time in most cases. After 40 periods, an unconventional monetary policy shock explains between 0% and 90% of the FEVD of industrial production and prices, while its contribution to the FEVD of financial variables is in general smaller.

As stressed by Baumeister and Hamilton (2015), the use of a uniform prior on $\boldsymbol{Q}$ does not imply that its elements are uniformly distributed over the identified set. The likelihood does not in fact depend on $\boldsymbol{Q}$ and this prior is thus not updated by the data. Although uniform, it might be informative for objects of interest as FEVD, even asymptotically. Unsurprisingly, our confidence intervals (Figure 2) are significantly larger. To put it another way, the gap between the confidence sets of the two approaches in Figure 2 can be viewed as a way of quantifying the sensitivity of the standard Bayesian inference to the choice of unrevisable prior. We therefore conclude that, under the standard Bayesian approach, the estimation of the FEVD is mostly driven by the unrevisable prior on $\boldsymbol{Q}$ rather than identifying constraints.

[17] For the reduced-form parameters, an uninformative Normal-Inverse Wishart prior is used.

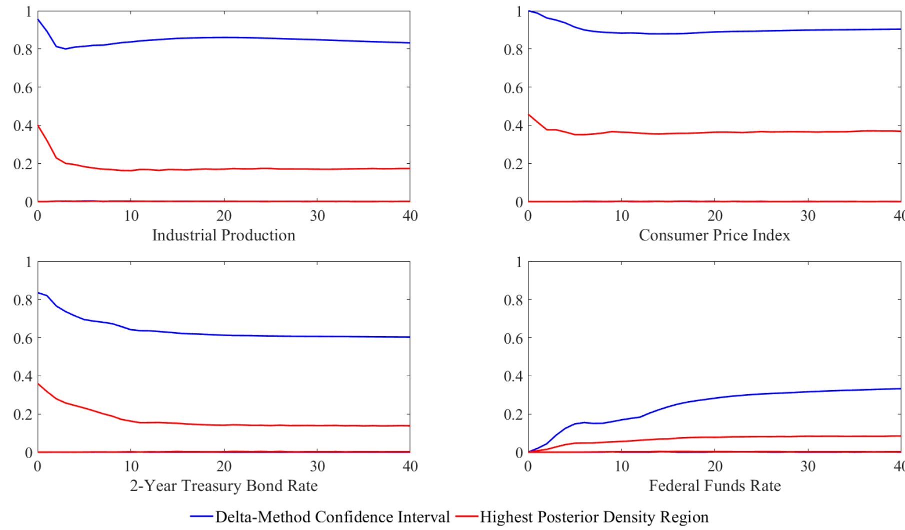

Figure 2: 68% Delta-Method Confidence Interval *vs* 68% Highest Posterior Density Region Under Uniform Prior on $\boldsymbol{Q}$.

In Table 2, we present the precise estimates of the objects plotted in Figure 1 and 2, for selected horizons. In particular, the left panel shows the estimated bounds, while the right panel shows the 68% delta-method interval.

	Estimated Bounds			Delta-Method 68% CI		
	$h = 0$	$h = 12$	$h = 40$	$h = 0$	$h = 12$	$h = 40$
CPI_t	[0.00,0.95]	[0.00,0.81]	[0.00,0.78]	[0.00,0.96]	[0.00,0.85]	[0.00,0.83]
IP_t	[0.00,1.00]	[0.00,0.85]	[0.00,0.85]	[0.00,1.00]	[0.00,0.88]	[0.00,0.90]
FF_t	[0.00,0.82]	[0.00,0.57]	[0.00,0.51]	[0.00,0.84]	[0.00,0.64]	[0.00,0.60]
$2YT_t$	[0.00,0.00]	[0.00,0.14]	[0.00,0.25]	[0.00,0.00]	[0.00,0.18]	[0.00,0.33]

Table 2: Estimated Bounds and Delta-Method 68% Confidence Interval.

When comparing our framework with the robust Bayesian approach in GK21, the results become relatively more similar as we increase the number of draws of $\boldsymbol{Q}$ for the framework in GK21. This is expected as our toolkit has an asymptotic Bayesian interpretation. However, in order to reduce the high computational costs of the procedure in GK21, in practice researchers tend to limit significantly the number of draws, introducing substantial bias. Constructing the bounds using the delta-method approach outlined in this paper takes 19s, in comparison to $4,086$s for the GK21 algorithm with $10,000$ draws for the rotation matrix, and 375s for $1,000$ draws.[18] We stressed that using a plug-in estimator for the FEVD from Gafarov et al. (2018)

[18]We implement it by using $1,000$ draws from the reduced-form posterior. Computational times are measured

is unwise; Figure B.10 in the Appendix shows the dramatic difference with our approach.

Finally, the large identification uncertainty implied by the identification strategy in this application is confirmed by the identified sets for the IRFs, reported in Figure B.11 in the Appendix.

5 Monte-Carlo Evidence

We conduct a Monte-Carlo exercise to illustrate the coverage of our delta-method interval in finite samples. We draw $\boldsymbol{\phi}$ directly from a multivariate normal distribution when assessing frequentist coverage, which has the advantage of directly enforcing Assumption A7. The moments of the distribution are set at their estimated values from the empirical application, which is our DGP. T is set to be 341 to mirror the number of periods (months) in the application. For the robust Bayesian credibility, we use an uninformative Normal-Inverse Wishart prior for $\boldsymbol{\phi}$, and draw from the posterior distribution. In both cases, we set $1 - \alpha = 0.68$, and compute 10,000 draws of $\widehat{\boldsymbol{\phi}}$. In addition, to account for any small sample bias, we implement the bias correction used in Gorodnichenko and Lee (2020).[19]

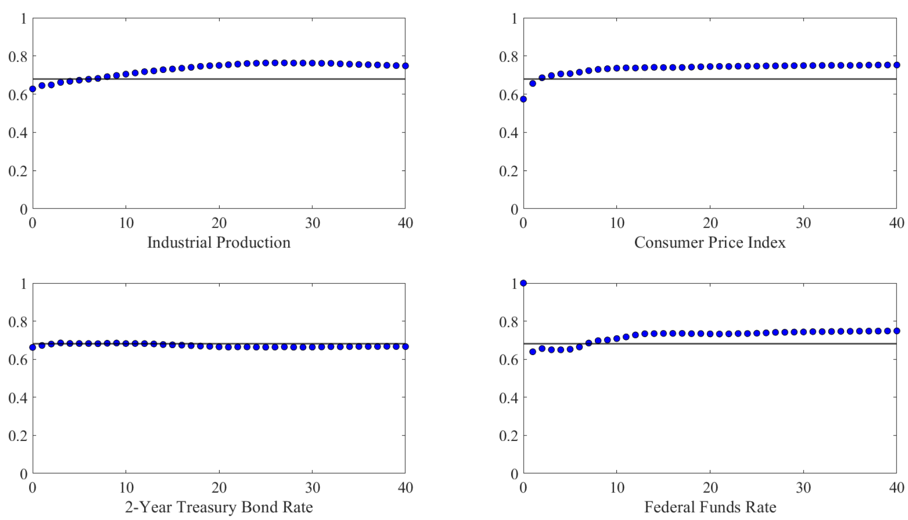

Figure 3: Monte-Carlo Exercise: Frequentist Coverage.

For each draw of $\widehat{\boldsymbol{\phi}}$, the delta-method interval is constructed as detailed in equation (3.14).

by running Matlab R2023b on a standard quad-core laptop @2.40GHz IntelCore i5. Also, GK21 provide two algorithms for their approach; we rely on the quicker of the two, so it provides a fairer comparison.

[19] Accounting for the small sample bias helps illustrate the coverage properties of our approach. However, as shown in Figure B.9 in the Appendix, this only marginally alters the estimated bounds. For this reason, in Section 4, we use the non-bias corrected bounds when comparing them with alternative approaches.

Coverage is then assessed to be the proportion of these intervals which contain the estimated bounds from the empirical application. In the main text we focus on the coverage of the estimated upper bound. As the FEVD is bounded between zero and one, the distribution is truncated. Given that in the DGP (the application in Section 4) the lower bound is always very close to zero, its coverage is less interesting. Also, the coverage at the boundaries is expected to be affected by this truncation. It is well-known in the literature about interval-defined parameters that the coverage close to the boundaries should be interpreted with a pinch of salt (see, for example, Gorodnichenko and Lee (2020)). Thus, we store illustration of coverage for the lower bound in Appendix C.

Frequentist coverage, reported in Figure 3, is shown to be correct. The exception is for the Federal Funds Rate on impact, where the imposed zero restriction ensures a coverage of one by definition. Note that in the very few periods where the coverage is slightly below the nominal level, the bounds in the DGP are extremely close to the boundaries. Figure 4 shows that the set-length correction is the key to obtain the correct non-conservative coverage.

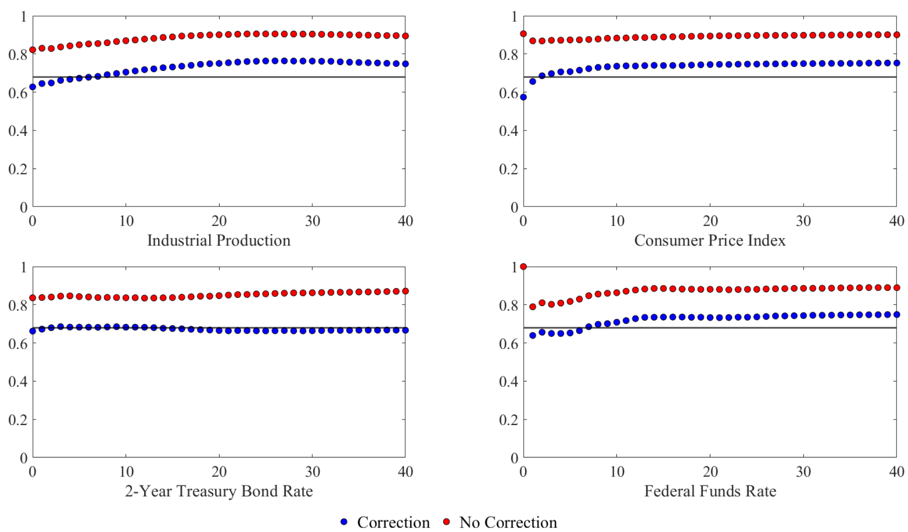

Figure 4: Monte-Carlo Exercise: Frequentist Coverage With and Without Set-Length Correction to the Critical Value (FEVD).

In addition, for the upper bound of the IRFs, Figure 5 demonstrates that the set-length correction is also the key to obtaining good coverage for IRFs by adjusting the confidence interval in Gafarov et al.'s (2018). The red line represents the coverage for the IRFs induced by the confidence interval in Gafarov et al.'s (2018); the blue line displays the coverage of our proposed interval for the IRFs (equation 3.21). Figure C.14 in the Appendix reports the coverage for the lower bound of the IRFs.

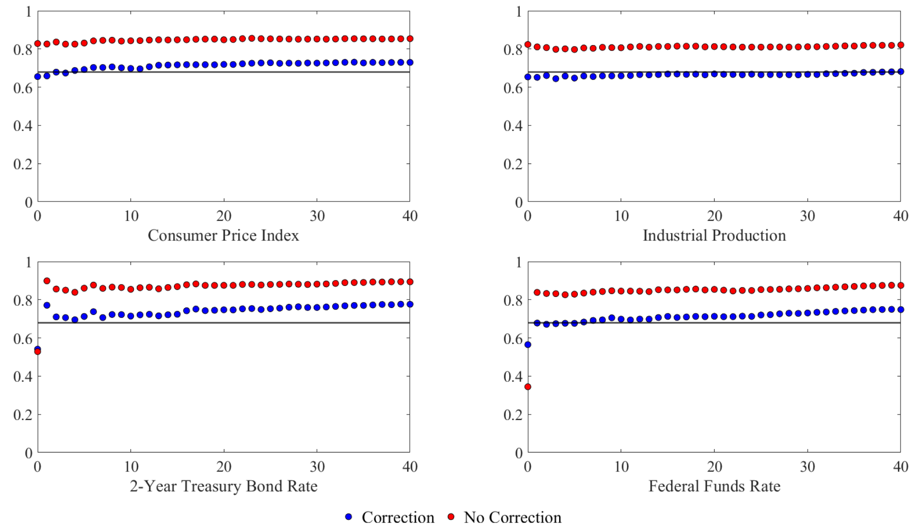

Figure 5: Monte-Carlo Exercise: Frequentist Coverage With and Without (Gafarov et al.'s (2018)) Set-Length Correction to the Critical Value (IRFs).

For those readers interested in the inference over the identified set, Figure 6 confirms the discussion in Section 3.3.3: i) our proposed confidence interval for the inference over the set (equation 3.26, red line in Figure 6) is valid and ii) confidence intervals targeting parameters are likely to be invalid for the set (blue line).

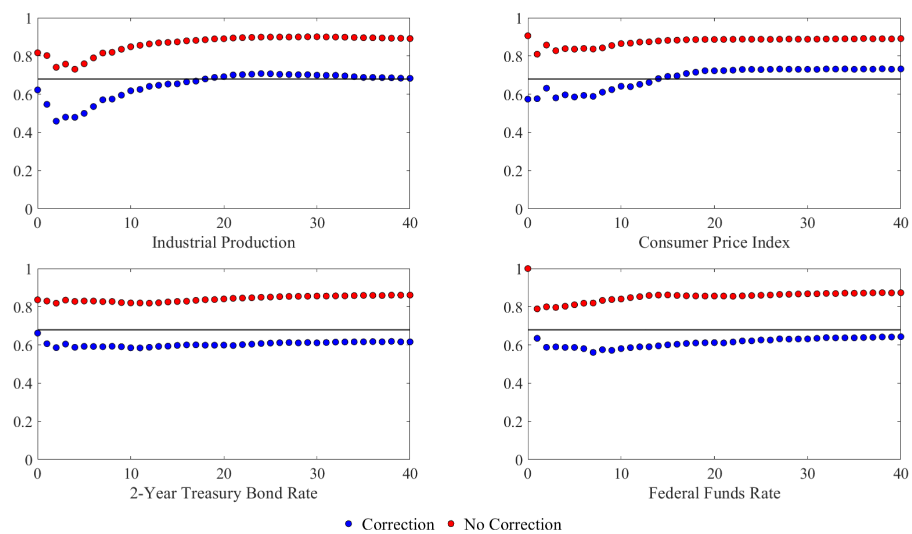

Figure 6: Monte-Carlo Exercise: Frequentist Coverage *over the set* With and Without Set-Length Correction to the Critical Value (FEVD).

Furthermore, Figure 7 displays that the frequentist coverage for the upper bound of the Robust Bayesian interval in GK21, which is valid for the set of FEVD (see Corollary 3.1 above), is mostly invalid when the object of interest is the FEVD itself. Appendix C i) provides details about the computation of the coverage for the approach in GK21 and ii) shows the coverage for the lower bound (Figure C.15).

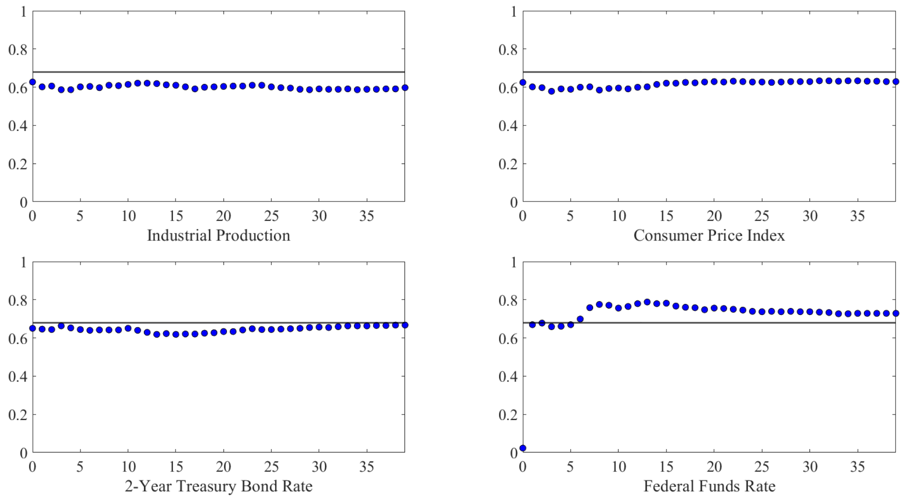

Figure 7: Monte Carlo Exercise: Coverage of GK's Robust Credible Interval.

Finally, we compute the robust Bayesian credibility of our delta-method interval based on an uninformative Normal-Inverse Wishart prior (Figure 8). The former is shown to be larger than the nominal level in most of the cases. Thus, Theorem 3.4, which states the asymptotic equivalence between our method and a Bayesian approach, also has some validity in finite samples.

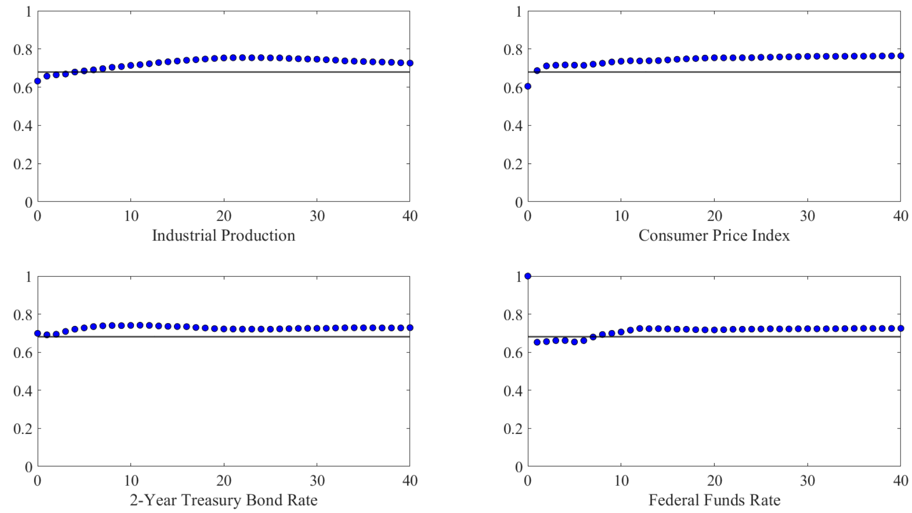

Figure 8: Monte-Carlo Exercise: Robust Bayesian Credibility.

6 Conclusion

This paper provides a toolkit for estimation and inference of the FEVD for set-identified SVARs, while the literature mostly focuses on IRFs. It overcomes the well-known problem of having a prior distribution that cannot be updated in the standard Bayesian approach. We derive the bounds as the extreme eigenvalues of a symmetric reduced-form matrix coming from quadratic programming. We prove the differentiability of the endpoints and construct a delta-method confidence interval adjusted by the length of the identified set. This is uniformly consistent in level and recovers asymptotic Bayesian credibility. We also provide conditions under which a similar adjustment can be applied to the IRFs. In finite samples, a Monte-Carlo exercise demonstrates that our approach has good coverage and outperforms alternatives. An application based on the identification of unconventional monetary policy shocks illustrates our toolkit and its computational convenience.

References

Amir-Ahmadi, P. and T. Drautzburg (2021): "Identification and Inference with Ranking Restrictions," *Quantitative Economics*, 12, 1–39.

Amir-Ahmadi, P. and H. Uhlig (2015): "Sign Restrictions in Bayesian FAVARs with an Application to Monetary Policy Shocks," *unpublished manuscript*.

Amisano, G. and C. Giannini (1997): "Impulse Response Analysis and Forecast Error Variance Decomposition in SVAR Modelling," *Topics in Structural VAR Econometrics*, 60–77.

Angeletos, G.-M., F. Collard, and H. Dellas (2020): "Business-cycle Anatomy," *American Economic Review*, 110, 3030–70.

Arias, J. E., D. Caldara, and J. F. Rubio-Ramirez (2019): "The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure," *Journal of Monetary Economics*, 101, 1–13.

Arias, J. E., J. F. Rubio-Ramirez, and D. Waggoner (2018): "Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications," *Econometrica*, 86, 685–720.

Barsky, R. B. and E. R. Sims (2011): "News Shocks and Business Cycles," *Journal of monetary Economics*, 58, 273–289.

Baumeister, C. and J. D. Hamilton (2015): "Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information," *Econometrica*, 83, 1963–1999.

——— (2018): "Inference in Structural Vector Autoregressions when the Identifying Assumptions Are Not Fully Believed: Re-evaluating the Role of Monetary Policy in Economic Fluctuations," *Journal of Monetary Economics*, 100, 48–65.

Beaudry, P., D. Nam, and J. Wang (2011): "Do Mood Swings Drive Business Cycles and Is It Rational?" *National Bureau of Economic Research Working Paper*.

Beaudry, P. and F. Portier (2006): "Stock Prices, News, and Economic Fluctuations," *American economic review*, 96, 1293–1307.

Benati, L. and T. A. Lubik (2022): "Searching for Hysteresis," *Federal Reserve Bank of Richmond Working Paper*.

Bisgaard, S. and B. Ankenman (1996): "Standard Errors for the Eigenvalues in Second-Order Response Surface Models," *Technometrics*, 38, 238–246.

Braun, P. A. and S. Mittnik (1993): “Misspecifications in Vector Autoregressions and their Effects on Impulse Responses and Variance Decompositions,” *Journal of econometrics*, 59, 319–341.

Braun, R. and R. Brüggemann (2022): “Identification of SVAR models by combining sign restrictions with external instruments,” *Journal of Business & Economic Statistics*, 1–13.

Caggiano, G. and E. Castelnuovo (2023): “Global Financial Uncertainty,” *Journal of Applied Econometrics*, 38, 432–449.

Caggiano, G., E. Castelnuovo, S. Delrio, and R. Kima (2021): “Financial Uncertainty and Real Activity: The good, the Bad, and the Ugly,” *European Economic Review*, 136, 103750.

Caldara, D., C. Fuentes-Albero, S. Gilchrist, and E. Zakrajšek (2016): “The Macroeconomic Impact of Financial and Uncertainty Shocks,” *European Economic Review*, 88, 185–207.

Carriero, A. and A. Volpicella (2024): “Max Share Identification of Multiple Shocks: An Application to Uncertainty and Financial Conditions,” *Journal of Business & Economic Statistics*, 1–20.

Carter, W. H., V. M. Chinchilli, and E. D. Campbell (1990): “A Large-Sample Confidence Region Useful in Characterizing the Stationary Point of a Quadratic Response Surface,” *Technometrics*, 32, 425–435.

Carter, W. H., V. M. Chinchilli, R. H. Myers, and E. D. Campbell (1986): “Confidence Intervals and An Improved Ridge Analysis of Response Surfaces,” *Technometrics*, 28, 339–346.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): “Monetary Policy Shock: What Have We Learned and to What End?” in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Elsevier.

Daalhuis, A. O. (2010): “Hypergeometric Function,” in *NIST Handbook of Mathematical Functions*, ed. by F. W. J. Olver, D. M. Lozier, R. F. Boisvert, and C. Clark, Cambridge University Press.

Dedola, L. and S. Neri (2007): “What Does a Technology Shock Do? A VAR Analysis with Model-Based Sign Restrictions,” *Journal of Monetary Economics*, 54, 512–549.

Dedola, L., G. Rivolta, and L. Stracca (2017): “If the Fed Sneezes, Who Catches a Cold?” *Journal of International Economics*, 108, S23–S41.

DiCecio, R. and M. Owyang (2010): "Identifying Technology Shocks in the Frequency Domain," *Federal Reserve Bank of St. Louise Working Paper No 2010-025A*.

Engle, R. F. and C. W. Granger (1987): "Co-integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 251–276.

Erdélyi, A., W. Magnus, F. Oberhettinger, and F. Tricomi (1953): *Higher Transcendental Functions*, McGraw–Hill.

Faust, J. (1998): "The Robustness of Identified VAR Conclusions about Money," *Carnegie-Rochester Conference Series on Public Policy*, 48, 207–244.

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

Fujita, S. (2011): "Dynamics of Worker Flows and Vacancies: Evidence from the Sign Restriction Approach," *Journal of Applied Econometrics*, 26, 89–121.

Furlanetto, F., A. Lepetit, Ørjan Robstad, J. F. Rubio-Ramirez, and P. Ulvedal (forthcoming): "Estimating Hysteresis Effects," *American Economic Journal: Macroeconomics*.

Fusari, F. (2023): "Identifying Monetary Policy Shocks Through External Variable Constraints," *Discussion Papers in Economics, University of Surrey*, 01/23.

Gafarov, B., M. Meier, and J.-L. Montiel-Olea (2016): "Projection inference for set-identified SVARs," *unpublished manuscript*.

Gafarov, B., M. Meier, and J. L. M. Olea (2018): "Delta-Method Inference for a Class of Set-Identified SVARs," *Journal of Econometrics*, 203, 316–327.

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

Ghosal, S., J. K. Ghosh, and T. Samanta (1995): "On Convergence of Posterior Distributions," *The Annals of Statistics*, 23, 2145–2152.

Giacomini, R. and T. Kitagawa (2021): "Robust Bayesian Inference for Set-Identified Models," *Econometrica*, 89, 1519–1556.

Giacomini, R., T. Kitagawa, and M. Read (2022a): "Robust Bayesian Inference in Proxy SVARs," *Journal of Econometrics*, 228, 107–126.

Giacomini, R., T. Kitagawa, and A. Volpicella (2022b): "Uncertain Identification," *Quantitative Economics*, 13, 95–123.

Gorodnichenko, Y. and B. Lee (2020): "Forecast Error Variance Decompositions with Local Projections," *Journal of Business & Economic Statistics*, 38, 921–933.

Granziera, E., H. R. Moon, and F. Schorfheide (2018): "Inference for VARs Identified with Sign Restrictions," *Quantitative Economics*, 9, 1087–1121.

Imbens, G. W. and C. F. Manski (2004): "Confidence intervals for partially identified parameters," *Econometrica*, 72, 1845–1857.

Inoue, A. and L. Kilian (2022): "Joint Bayesian Inference About Impulse Responses in VAR Models," *Journal of Econometrics*, 231, 457–476.

Jarociński, M. and P. Karadi (2020): "Deconstructing Monetary Policy Surprises—The Role of Information Shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Kasy, M. (2018): "Uniformity and the Delta Method," *Journal of Econometric Methods*, 8, 20180001.

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

Korobilis, D. (2022): "A New Algorithm for Structural Restrictions in Bayesian Vector Autoregressions," *European Economic Review*, 148, 104241.

Kurmann, A. and E. Sims (2021): "Revisions in Utilization-Adjusted TFP and Robust Identification of News Shocks," *The Review of Economics and Statistics*, 103, 216–235.

Lanne, M. and H. Nyberg (2016): "Generalized Forecast Error Variance Decomposition for Linear and Nonlinear Multivariate Models," *Oxford Bulletin of Economics and Statistics*, 78, 595–603.

Levchenko, A. A. and N. Pandalai-Nayar (2020): "TFP, News, and "sentiments": The International Transmission of Business Cycles," *Journal of the European Economic Association*, 18, 302–341.

Ludvigson, S. C., S. Ma, and S. Ng (2021): "Uncertainty and Business Cycles: Exogenous Impulse or Endogenous Response?" *American Economic Journal: Macroeconomics*, 13, 369–410.

Lütkepohl, H. (1990): "Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models," *The review of economics and statistics*, 116–125.

Magnus, J. R. (1985): "On Differentiating Eigenvalues and Eigenvectors," *Econometric theory*, 1, 179–191.

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

Mumtaz, H., G. Pinter, and K. Theodoridis (2018): "What do Vars tell us about the Impact of a Credit Supply Shock?" *International Economic Review*, 59, 625–646.

Mumtaz, H. and K. Theodoridis (2023): "The Federal Reserve's Implicit Inflation Target and Macroeconomic Dynamics. A SVAR Analysis," *Forthcoming, International Economic Review*.

Mumtaz, H. and F. Zanetti (2012): "Neutral Technology Shocks and the Dynamics of Labor Input: Results from an Agnostic Identification," *International Economic Review*, 53, 235–254.

Peersman, G. and R. Straub (2009): "Technology Shocks and Robust Sign Restrictions in a Euro Area SVAR," *International Economic Review*, 50, 727–750.

Peterson, J. J. (1993): "A General Approach to Ridge Analysis with Confidence Intervals," *Technometrics*, 35, 204–214.

Phillips, P. C. (1998): "Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VARs," *Journal of econometrics*, 83, 21–56.

Piffer, M. and M. Podstawski (2018): "Identifying Uncertainty Shocks Using the Price of Gold," *The Economic Journal*, 128, 3266–3284.

Plagborg-Møller, M. and C. K. Wolf (2022): "Instrumental Variable Identification of Dynamic Variance Decompositions," *Journal of Political Economy*, 130, 2164–2202.

Poirier, D. J. (1998): "Revising Beliefs in Nonidentified Models," *Econometric Theory*, 14, 483–509.

Rao, C. R. (1964): "The Use and Interpretation of Principal Component Analysis in Applied Research," *Sankhyā: The Indian Journal of Statistics, Series A*, 329–358.

——— (1973): *Linear Statistical Inference and Its Applications*, vol. 2, Wiley New York.

Rubio-Ramirez, J., D. Waggoner, and T. Zha (2010): "Structural Vector Autoregressions: Theory of Identification and Algorithm for Inference," *The Review of Economic Studies*, 77, 665–696.

Scholl, A. and H. Uhlig (2008): "New Evidence on the Puzzles: Results from Agnostic Identification on Monetary Policy and Exchange Rates," *Journal of International Economics*, 76, 1–13.

Smets, F. and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American economic review*, 97, 586–606.

Stock, J. H. and M. W. Watson (2016): "Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics," in *Handbook of macroeconomics*, Elsevier, vol. 2, 415–525.

Stoye, J. (2009): "More on Confidence Intervals for Partially Identified Parameters," *Econometrica*, 77, 1299–1315.

Uhlig, H. (2004): "What Moves GNP?" in *Econometric Society 2004 North American Winter Meetings*, Econometric Society, 636.

——— (2005): "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 52, 381–419.

Volpicella, A. (2022): "SVARs Identification Through Bounds on the Forecast Error Variance," *Journal of Business & Economic Statistics*, 40, 1291–1301.

Appendix to: "Estimation and Inference of the Forecast Error Variance Decomposition for set-identified SVARs"

The Appendix contains omitted proofs (Appendix A), further findings from the empirical application (Appendix B) and additional simulation results (Appendix C).

A Omitted Proofs

Proof of Lemma 2.1.

Conditions of Proposition 2.1 are equivalent to the assumptions of Proposition B1, (ii) in the Appendix of GK21. In turn, this implies that restrictions on shocks $1,\dots,j^*-1,j^*+1,\dots,n-1$ leave the set of r_{ij^*h} unchanged for $i=1,\dots,n$ and $h=0,\dots$ (Corollary B2 in the Appendix of GK21). Since $FEVD_{ij^*h}$ is a continuous function of $\boldsymbol{q}_j^*$, its set is unaffected by constraints on shocks $1,\dots,j^*-1,j^*+1,\dots,n-1$. ■

Proof of Proposition 3.1.

This proof builds upon Faust (1998) and Rao (1964). While Uhlig (2004) provides a solution to a quadratic unconstrained problem, this proof can be seen as a generalization to a quadratic constrained problem.

The maximization problem is provided by (3.2). $\boldsymbol{r}(\boldsymbol{\phi})$ is of full rank by Assumption A4; as a result, there are between 0 and $n-1$ active constraints ($0\le m\le n-1$) at a maximum.

By employing the Karush-Kuhn-Tucker theory, we obtain the following first order conditions (FOCs):

$$\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})-\pi\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})-\boldsymbol{r}(\boldsymbol{\phi})'\mu=0 \tag{A.1}$$

$$\boldsymbol{r}(\boldsymbol{\phi})\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})=0 \tag{A.2}$$

$$\bar{\boldsymbol{q}}_{j^*}'(\boldsymbol{r})\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})=1, \tag{A.3}$$

where π and μ are positive Lagrange multipliers. We are going to show that $u_{ij^*h}(\boldsymbol{\phi},\boldsymbol{r})$ and $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ are the maximum eigenvalue and the associated eigenvector of the matrix $\boldsymbol{Z}(\boldsymbol{\phi},\boldsymbol{r})=[\boldsymbol{I}_n-\boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})]\,\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})$, where $\boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})=\boldsymbol{r}(\boldsymbol{\phi})'\left[\boldsymbol{r}(\boldsymbol{\phi})\boldsymbol{r}(\boldsymbol{\phi})'\right]^{-1}\boldsymbol{r}(\boldsymbol{\phi})$. Pre-multiplying the FOCs by $\boldsymbol{I}_n-\boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})$ yields

$$[\boldsymbol{I}_n-\boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})]\,\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})-\pi\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})=0. \tag{A.4}$$

This is satisfied if and only if $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ is an eigenvector of $[\boldsymbol{I}_n-\boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})]\,\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})$. We need to show that $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ is the eigenvector associated with the largest eigenvalue. Rao (1964) (Section 2, part iii) observes that the eigenvalues of $[\boldsymbol{I}_n-\boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})]\,\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})$ are equivalent to the eigenvalues of $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi},\boldsymbol{r})=\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})^{\frac{1}{2}}\{\boldsymbol{I}_n-\boldsymbol{r}(\boldsymbol{\phi})'\left[\boldsymbol{r}(\boldsymbol{\phi})\boldsymbol{r}(\boldsymbol{\phi})'\right]^{-1}\boldsymbol{r}(\boldsymbol{\phi})\}\boldsymbol{\Upsilon}_h^i(\boldsymbol{\phi})^{\frac{1}{2}'}$. This provides some computational gains, since symmetric eigenvalue problems are well-understood. Also, Rao (1964) derives the following relationship between the eigenvector of $\boldsymbol{Z}(\boldsymbol{\phi},\boldsymbol{r})$ (among others, $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$) and $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi},\boldsymbol{r})$ (say $\bar{\tilde{\boldsymbol{q}}}_{j^*}(\boldsymbol{r})$):

$$\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r}) = \{\bar{\tilde{\boldsymbol{q}}}'_{j^*}(\boldsymbol{r}) \left[\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi}\right]^{-1} \bar{\tilde{\boldsymbol{q}}}_{j^*}(\boldsymbol{r})\} \left[\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\right]^{-\frac{1}{2}} \bar{\tilde{\boldsymbol{q}}}_{j^*}(\boldsymbol{r}). \tag{A.5}$$

This applies to any eigenvector of $\boldsymbol{Z}(\boldsymbol{\phi},\boldsymbol{r})$ and $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi},\boldsymbol{r})$.
Note that $\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})^{\frac{1}{2}}\{\boldsymbol{I}_n - \boldsymbol{r}(\boldsymbol{\phi})' \left[\boldsymbol{r}(\boldsymbol{\phi})\boldsymbol{r}(\boldsymbol{\phi})'\right]^{-1} \boldsymbol{r}(\boldsymbol{\phi})\}\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})^{\frac{1}{2}'}$ is positive semidefinite, since $[\boldsymbol{I}_n - \boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})]$ is idempotent and $\left[\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})^{\frac{1}{2}}\right]$ is full rank. $[\boldsymbol{I}_n - \boldsymbol{P}(\boldsymbol{\phi},\boldsymbol{r})]\,\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})$ is therefore positive semi-definite and the eigenvectors can be selected to be orthogonal. Let us prove the result by contradiction and assume that $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ is not the eigenvecor associated with the largest eigenvalue:

$$\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r}) = \omega \boldsymbol{q}_{1j^*}(\boldsymbol{r}) + \sqrt{(1-\omega^2)}\tilde{\boldsymbol{q}}_{j^*}(\boldsymbol{r}), \tag{A.6}$$

where $\boldsymbol{q}_{1j^*}(\boldsymbol{r})$ and $\tilde{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ satisfy the FOCs, are orthogonal and $\boldsymbol{q}_{1j^*}(\boldsymbol{r})$ is associated to the largest eigenvalue. Let us parametrize $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ as follows:

$$\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r},\delta) = (1+\delta)\omega \boldsymbol{q}_{1j^*}(\boldsymbol{r}) + \sqrt{(1-(1+\delta)^2\omega^2)}\tilde{\boldsymbol{q}}_{j^*}(\boldsymbol{r}), \tag{A.7}$$

which is $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r})$ for $\delta = 0$ and satisfies the active constraints for small δ. The eigenvectors are orthogonal, so the value of the criterion function can expressed as

$$\bar{\boldsymbol{q}}'_{j^*}(\boldsymbol{r},\delta)\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r},\delta) = (1+\delta)^2\omega^2\boldsymbol{q}'_{1j^*}(\boldsymbol{r})\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\boldsymbol{q}_{1j^*}(\boldsymbol{r}) + (1-(1+\delta)^2\omega^2)\tilde{\boldsymbol{q}}'_{j^*}(\boldsymbol{r})\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\tilde{\boldsymbol{q}}_{j^*}(\boldsymbol{r}). \tag{A.8}$$

The first derivative of the right hand side of the equation above with respect to δ is $2(1+\delta)\omega^2\left[\boldsymbol{q}'_{1j^*}(\boldsymbol{r})\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\boldsymbol{q}_{1j^*}(\boldsymbol{r}) - \tilde{\boldsymbol{q}}'_{j^*}(\boldsymbol{r})\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\tilde{\boldsymbol{q}}_{j^*}(\boldsymbol{r})\right]$. It is positive for small δ because $\boldsymbol{q}'_{1j^*}(\boldsymbol{r})\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})\boldsymbol{q}_{1j^*}(\boldsymbol{r})$ maximizes the objective function. Let us conclude the proof: i) since all the constraints holding with equality at $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r},0)$ are also satisfied (with equality) for small δ; ii) since $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r},0)$ satisfies the inactive constraints by definition $\Rightarrow$ by continuity there must be some small $\delta > 0$ such that the restrictions are satisfied at $\bar{\boldsymbol{q}}_{j^*}(\boldsymbol{r},\delta)$.

For $\lambda_{min}(\boldsymbol{\phi},\boldsymbol{r}) \neq 0$, we can obtain the minimum analogously. For $\lambda_{min}(\boldsymbol{\phi},\boldsymbol{r}) = 0$, the Karush-Kuhn-Tucker conditions are not always satisfied. In that case, there are two possibilities: $l_{ij^*h}(\boldsymbol{\phi},\boldsymbol{r}) = 0$ or $l_{ij^*h}(\boldsymbol{\phi},\boldsymbol{r}) = \lambda^+_{min}(\boldsymbol{\phi},\boldsymbol{r}) \neq 0$, where $\lambda^+_{min}(\boldsymbol{\phi},\boldsymbol{r})$ is the smallest non-zero eigenvalue. ■

Proof of Theorem 3.1.

The maximization problem is provided by (2.22). First, we need to introduce an auxiliary function $\tilde{u}_{ij^*h}(\boldsymbol{\phi},\boldsymbol{r})$:

$$\tilde{u}_{ij^*h}(\boldsymbol{\phi},\boldsymbol{r}) = u_{ij^*h}(\boldsymbol{\phi},\boldsymbol{r}) - c\{1 - \mathbf{1}\left[\tilde{\boldsymbol{S}}_{j^*}(\boldsymbol{\phi})\boldsymbol{q}_{j^*} \geq \mathbf{0}\right]\}, \tag{A.9}$$

with $c > 1$ and where $\mathbf{1}(\bullet)$ denotes the indicator function. $\tilde{\boldsymbol{S}}_{j^*}(\boldsymbol{\phi})\boldsymbol{q}_{j^*} \geq \mathbf{0}$ is the set of inactive constraints, with $\tilde{\boldsymbol{S}}_{j^*}(\boldsymbol{\phi})$ being a $(s_{j^*} - s_{rj^*}) \times n$ matrix.

Given a set of active constraints $\boldsymbol{r}(\boldsymbol{\phi})$, if the optimizer satisfies the inactive constraints characterized by $\tilde{\boldsymbol{S}}_{j^*}(\boldsymbol{\phi})$, then $\mathbf{1}\left[\tilde{\boldsymbol{S}}_{j^*}(\boldsymbol{\phi})\boldsymbol{q}_{j^*} \geq \mathbf{0}\right] = 1$ and $\tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) = u_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r})$; otherwise, $\mathbf{1}\left[\tilde{\boldsymbol{S}}_{j^*}(\boldsymbol{\phi})\boldsymbol{q}_{j^*} \geq \mathbf{0}\right] = 0$ and the auxiliary function is negative (recall that $c > 1$) and, as such, cannot be considered an endpoint for the FEVD. In other words, we are making explicit that, if some constraints are not satisfied, the FEVD cannot be defined. Proving the theorem corresponds to show the following

$$u_{ij^*h}(\boldsymbol{\phi}) = max_{\boldsymbol{r}(\boldsymbol{\phi})}\tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}). \tag{A.10}$$

- CASE 1: consider any $\boldsymbol{r}(\boldsymbol{\phi})$.

  CASE 1.1 Suppose that the inactive constraints are not satisfied, i.e. $\tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}) < 0$. This implies that $u_{ij^*h}(\boldsymbol{\phi}) > \tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r})$.

  CASE 1.2 Suppose that the active constraints are satisfied, i.e. $\mathbf{1}\left[\tilde{\boldsymbol{S}}_{j^*}(\boldsymbol{\phi})\boldsymbol{q}_{j^*} \geq \mathbf{0}\right] = 1$. As a result, $u_{ij^*h}(\boldsymbol{\phi}) \geq \tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r})$. Case 1 therefore delivers

$$u_{ij^*h}(\boldsymbol{\phi}) \geq max_{\boldsymbol{r}(\boldsymbol{\phi})}\tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}). \tag{A.11}$$

- CASE 2: consider $\boldsymbol{r}^*(\boldsymbol{\phi})$ as the set of active constraints at the optimizer of the problem (2.22). As such, inactive constraints are satisfied. Proof of Proposition 3.1 yields:

$$\tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}^*) = u_{ij^*h}(\boldsymbol{\phi}) \leq max_{\boldsymbol{r}(\boldsymbol{\phi})}\tilde{u}_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r}). \tag{A.12}$$

Equations (A.11) and (A.12) deliver the desired result.

■

Proof of Proposition 3.2.

If the reduced-form estimator is consistent, i.e $\widehat{\boldsymbol{\phi}} \to_p \boldsymbol{\phi}(P)$, then we obtain

$$\boldsymbol{r^*}(\widehat{\boldsymbol{\phi}}) \to_p \boldsymbol{r^*}(\boldsymbol{\phi}(P)) \tag{A.13}$$

and

$$\boldsymbol{P}(\widehat{\boldsymbol{\phi}}, \boldsymbol{r^*}) \to_p \boldsymbol{P}(\boldsymbol{\phi}(P), \boldsymbol{r^*}(P)). \tag{A.14}$$

. This implies that

$$\boldsymbol{Z}(\widehat{\boldsymbol{\phi}}, \boldsymbol{r^*}) \to_p \boldsymbol{Z}(\boldsymbol{\phi}(P), \boldsymbol{r^*}(P)). \tag{A.15}$$

Since $u_{ij^*h}(\boldsymbol{\phi}) = u_{ij^*h}(\boldsymbol{\phi}, \boldsymbol{r^*})$, if $u_{ij^*h}(\boldsymbol{\phi})$ is simple and $T \ggg n$, it is suffice to recall basic statistics to obtain $u_{ij^*h}(\widehat{\boldsymbol{\phi}}) \to_p u_{ij^*h}(\boldsymbol{\phi}(P))$. Proof for $l_{ij^*h}(\widehat{\boldsymbol{\phi}}) \to_p l_{ij^*h}(\boldsymbol{\phi}(P))$ is obtained similarly. ■

Proof of Theorem 3.2.

From the chain rule, we obtain:

$$\frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}} = \frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial vec(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))}\frac{\partial vec(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))}{\partial \boldsymbol{\phi}}. \tag{A.16}$$

Recall that $\boldsymbol{Z}(\boldsymbol{\phi}) \equiv \boldsymbol{Z}(\boldsymbol{\phi}, \boldsymbol{r^*}) = [\boldsymbol{I}_n - \boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r^*})]\, \boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})$,
$\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}) \equiv \tilde{\boldsymbol{Z}}(\boldsymbol{\phi}, \boldsymbol{r^*}) = \boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})^{\frac{1}{2}}\, [\boldsymbol{I}_n - \boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r^*})]\, \boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi})^{\frac{1}{2}'}$,
$\boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r^*}) = \boldsymbol{r}^*(\boldsymbol{\phi})'\, [\boldsymbol{r}^*(\boldsymbol{\phi})\boldsymbol{r}^*(\boldsymbol{\phi})']^{-1}\, \boldsymbol{r}^*(\boldsymbol{\phi})$ and $\boldsymbol{\Upsilon}^i_h(\boldsymbol{\phi}) = \frac{\sum_{\tilde{h}=0}^{h} \boldsymbol{c}_{i\tilde{h}}(\boldsymbol{\phi})\boldsymbol{c}'_{i\tilde{h}}(\boldsymbol{\phi})}{\sum_{\tilde{h}=0}^{h} \boldsymbol{c}'_{i\tilde{h}}(\boldsymbol{\phi})\boldsymbol{c}_{i\tilde{h}}(\boldsymbol{\phi})}$; as a result, $\frac{\partial vec(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))}{\partial \boldsymbol{\phi}}$ exists because of Assumption A6 (that makes $\boldsymbol{r}^*(\boldsymbol{\phi})$ and $\boldsymbol{I}_n - \boldsymbol{P}(\boldsymbol{\phi}, \boldsymbol{r^*})$ differentiable) and the fact that $\boldsymbol{c}_{i\tilde{h}}(\boldsymbol{\phi})$ and $\boldsymbol{c}'_{i\tilde{h}}(\boldsymbol{\phi})$ are differentiable with respect to $\boldsymbol{\phi}$.

We now need to show that $u_{ij^*h}(\boldsymbol{\phi})$ is differentiable with respect to $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})$. Let λ and $\boldsymbol{q}$ be a real-valued function and a vector function (respectively) defined for all $\boldsymbol{\Gamma}(\boldsymbol{\phi})$ in some neighborhood of $N(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))$. Consider the vector function $g : \mathcal{R}^{n+1} \times \mathcal{R}^{n\times n} \to \mathcal{R}^{n+1}$ defined as follows

$$g(\boldsymbol{q}, \lambda; \boldsymbol{\Gamma}(\boldsymbol{\phi})) = \begin{bmatrix} (\lambda \boldsymbol{I}_n - \boldsymbol{\Gamma}(\boldsymbol{\phi}))\boldsymbol{q} \\ \boldsymbol{q}'\boldsymbol{q} - 1 \end{bmatrix}. \tag{A.17}$$

Note that g is differentiable on $\mathcal{R}^{n+1} \times \mathcal{R}^{n\times n}$. $\left(\bar{\tilde{\boldsymbol{q}}}_{j^*}, u_{ij^*h}(\boldsymbol{\phi}); \tilde{Z}(\boldsymbol{\phi})\right)$ in $\mathcal{R}^{n+1} \times \mathcal{R}^{n\times n}$ satisfies the following conditions:

$$g(\bar{\tilde{\boldsymbol{q}}}_{j^*}, u_{ij^*h}(\boldsymbol{\phi}); \tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) = 0 \tag{A.18}$$

and

$$det \begin{bmatrix} u_{ij^*h}(\boldsymbol{\phi})\boldsymbol{I}_n - \tilde{\boldsymbol{Z}}(\boldsymbol{\phi}) & \bar{\tilde{\boldsymbol{q}}}_{j^*} \\ 2\bar{\tilde{\boldsymbol{q}}}'_{j^*} & 0 \end{bmatrix} \neq 0. \tag{A.19}$$

The determinant above is non-singular if and only if the eigenvalue $u_{ij^*h}(\boldsymbol{\phi})$ is simple. The implicit function theorem is therefore satisfied; there must exist a neighborhood $N(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) \in \mathcal{R}^{n\times n}$ of $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})$, a unique real-valued function $\lambda : N(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) \to \mathcal{R}$, and vector function (defined up to a sign) $\boldsymbol{q} : N(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) \to \mathcal{R}^n$, such that

- λ and $\boldsymbol{q}$ are differentiable on $N(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))$,
- $\lambda(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) = u_{ij^*h}(\boldsymbol{\phi}),\ \boldsymbol{q}(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) = \bar{\tilde{\boldsymbol{q}}}_{j^*}$,
- $\boldsymbol{\Gamma}(\boldsymbol{\phi})\boldsymbol{q} = \lambda\boldsymbol{q},\ \boldsymbol{q}'\boldsymbol{q} = 1$ for every $\boldsymbol{\Gamma}(\boldsymbol{\phi}) \in N(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))$.

This completes the proof about the differentiability of $\frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$. We now derive an explicit equation for $\frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$. We start from $\boldsymbol{\Gamma}(\boldsymbol{\phi})\boldsymbol{q} = \lambda\boldsymbol{q}$:

$$(d\boldsymbol{\Gamma}(\boldsymbol{\phi}))\bar{\tilde{\boldsymbol{q}}}_{j^*} + \tilde{\boldsymbol{Z}}(\boldsymbol{\phi})(d\boldsymbol{q}) = (d\lambda)\bar{\tilde{\boldsymbol{q}}}_{j^*} + u_{ij^*h}(\boldsymbol{\phi})(d\boldsymbol{q}), \tag{A.20}$$

with $d\boldsymbol{q}$ and $d\lambda$ being the differentials defined at $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})$. Let us pre-multiply the equation above by $\bar{\tilde{\boldsymbol{q}}}'_{j^*}$:

$$\bar{\tilde{\boldsymbol{q}}}'_{j^*}(d\boldsymbol{\Gamma}(\boldsymbol{\phi}))\bar{\tilde{\boldsymbol{q}}}_{j^*} + \bar{\tilde{\boldsymbol{q}}}'_{j^*}\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})(d\bar{\tilde{\boldsymbol{q}}}_{j^*}) = (d\lambda)\bar{\tilde{\boldsymbol{q}}}'_{j^*}\bar{\tilde{\boldsymbol{q}}}_{j^*} + u_{ij^*h}(\boldsymbol{\phi})\bar{\tilde{\boldsymbol{q}}}'_{j^*}(d\boldsymbol{q}). \tag{A.21}$$

Recall that $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})$ is symmetric, i.e. $\bar{\tilde{\boldsymbol{q}}}'_{j^*}\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}) = u_{ij^*h}(\boldsymbol{\phi})\bar{\tilde{\boldsymbol{q}}}'_{j^*}$, and $\bar{\tilde{\boldsymbol{q}}}'_{j^*}\bar{\tilde{\boldsymbol{q}}}_{j^*} = 1$:

$$(d\lambda) = \bar{\tilde{\boldsymbol{q}}}'_{j^*}(d\boldsymbol{\Gamma}(\boldsymbol{\phi}))\bar{\tilde{\boldsymbol{q}}}_{j^*}. \tag{A.22}$$

The equation above can be written as

$$(d\lambda) = (\bar{\tilde{\boldsymbol{q}}}'_{j^*} \otimes \bar{\tilde{\boldsymbol{q}}}'_{j^*})vec(d\boldsymbol{\Gamma}(\boldsymbol{\phi})). \tag{A.23}$$

Recalling that $d\lambda$ is evaluated at $\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})$, $\lambda(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) = u_{ij^*h}(\boldsymbol{\phi})$ and $\boldsymbol{q}(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})) = \bar{\tilde{\boldsymbol{q}}}_{j^*}$ yields

$$(du_{ij^*h}(\boldsymbol{\phi})) = (\bar{\tilde{\boldsymbol{q}}}'_{j^*} \otimes \bar{\tilde{\boldsymbol{q}}}'_{j^*})vec(d\tilde{\boldsymbol{Z}}(\boldsymbol{\phi})). \tag{A.24}$$

Thus, we obtain

$$\frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial vec(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))} = \bar{\tilde{\boldsymbol{q}}}'_{j^*} \otimes \bar{\tilde{\boldsymbol{q}}}'_{j^*}. \tag{A.25}$$

Combining the equation above with (A.16) delivers

$$\frac{\partial u_{ij^*h}(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}} = \bar{\tilde{\boldsymbol{q}}}'_{j^*} \otimes \bar{\tilde{\boldsymbol{q}}}'_{j^*}\frac{\partial vec(\tilde{\boldsymbol{Z}}(\boldsymbol{\phi}))}{\partial \boldsymbol{\phi}}. \tag{A.26}$$

■

Proof of Corollary 3.1.

Note that $IS_{FEVD}(\boldsymbol{\phi})$ is closed and bounded. The former comes from the continuity of the map between $\boldsymbol{\phi}$ and $FEVD$; the latter is due to the restriction on $\boldsymbol{\phi}$ such that the reduced-form VAR is invertible into $VMA(\infty)$. Thus, closedness and boundedness make sure that Assumption 3 in GK21 is met. On the other hand, the conditions in the corollary are consistent with Assumption 5 in GK21. Thus, we can apply Proposition 2 in GK21 to $IS_{FEVD}(\boldsymbol{\phi})$. ■

Proof of Theorem 3.3.

We introduce the following notation. P identifies a DGP, e.g. $\Delta(P) = u_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\boldsymbol{\phi}(P)), FEVD_{ij^*h}(P)$. For ease of notation, we suppress P from estimators $\widehat{\bullet}$. a_T indicates a sequence.

This proof shows point-wise limits; however, they imply uniformity because they are taken over sequences. Specifically, they apply along least favorable sequences. Proof presents two arguments: one for the case that $\Delta(P)$ is "small" and one for the case that it is "large" in a sense that will be delimited. Any sequence P can be decomposed into one "large" and one "small" subsequence.

Let $\Delta(P) \leq a_T$. Note the following.

(i) Given Assumption A7, the delta-method delivers a Normal distribution for $u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ and $l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$:

$$\sqrt{T}\begin{bmatrix} l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \\ u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \end{bmatrix} \rightarrow_d \boldsymbol{N}\left(\begin{bmatrix} l_{ij^*h}(\boldsymbol{\phi}(P)) \\ u_{ij^*h}(\boldsymbol{\phi}(P)) \end{bmatrix}, \begin{bmatrix} \sigma^2_{lij^*h}(P) & 0 \\ 0 & \sigma^2_{uij^*h}(P) \end{bmatrix}\right), \tag{A.27}$$

where we recall that $l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ and $l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ are independent eigenvalues because they come from a symmetric matrix. The convergence in distribution is uniform (formally, see Theorem 3 in Kasy (2018)).

(ii) For any P, $\sigma^2_{lij^*h}(P)$ and $\sigma^2_{uij^*h}(P)$ are positive (recall the non-zero condition) and finite and, by definition, $u_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\boldsymbol{\phi}(P)) < \infty$.

(iii) Since $u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ and $l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ are the extreme eigenvalues from a real symmetric matrix, they can be ordered such that $u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \geq l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ for all P.

We therefore obtain $\sqrt{T}(\widehat{\Delta} - \Delta(P)) \rightarrow_p 0$ for all P by Lemma 3 in Stoye (2009). As a result, we get

$$\sqrt{T}\left(u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) - u_{ij^*h}(\boldsymbol{\phi}(P))\right) = \sqrt{T}\left(l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) + \widehat{\Delta} - l_{ij^*h}(\boldsymbol{\phi}(P)) - \Delta(P)\right) \tag{A.28}$$

$$\rightarrow_p \sqrt{T}\left(l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) - l_{ij^*h}(\boldsymbol{\phi}(P))\right). \tag{A.29}$$

Relationship above in combination with (i) and (ii) yields $\sigma_{lij^*h}(P) = \sigma_{uij^*h}(P)$, implying that $\widehat{\sigma}_{uij^*h} - \widehat{\sigma}_{lij^*h} \rightarrow_p 0$. Employing Lemma 3 in Stoye (2009) delivers

$$\Phi\left(c_\alpha + \sqrt{T}\frac{\widehat{\Delta}}{max\{\widehat{\sigma}_{uij^*h}, \widehat{\sigma}_{lij^*h}\}}\right) \rightarrow_p \Phi\left(c_\alpha + \sqrt{T}\frac{\Delta(P)}{\widehat{\sigma}_{lij^*h}}\right). \tag{A.30}$$

We obtain uniform consistency by invoking the same argument as in the proof of Lemma 2 in Stoye (2009).

Let $\Delta(P) > a_T$. This leads to $\sqrt{T}\Delta(P) \rightarrow \infty$; in turn, this yields
$lim\ sup_{T\rightarrow\infty}\sqrt{T}\left(FEVD_{ij^*h}(P) - l_{ij^*h}(\boldsymbol{\phi}(P))\right) = \infty$
and $lim\ sup_{T\rightarrow\infty}\sqrt{T}\left(u_{ij^*h}(\boldsymbol{\phi}(P)) - FEVD_{ij^*h}(P)\right) = \infty$. Further steps get us to

$$Pr(FEVD_{ij^*h}(P) \in CI(\alpha)) \tag{A.31}$$

$$= Pr\{-c_\alpha\widehat{\sigma}_{lij^*h} \leq \sqrt{T}\left[FEVD_{ij^*h}(P) - l_{ij^*h}(\boldsymbol{\phi}(P))\right] + \sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})] \leq \sqrt{T}\widehat{\Delta} + c_\alpha\widehat{\sigma}_{uij^*h}\} \tag{A.32}$$

$$= Pr\{-c_\alpha\widehat{\sigma}_{lij^*h} \leq \sqrt{T}\left[FEVD_{ij^*h}(P) - l_{ij^*h}(\boldsymbol{\phi}(P))\right] + \sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})]\} \tag{A.33}$$

$$- Pr\{\sqrt{T}\left[FEVD_{ij^*h}(P) - l_{ij^*h}(\boldsymbol{\phi}(P))\right] + \sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})] > \sqrt{T}\widehat{\Delta} + c_\alpha\widehat{\sigma}_{uij^*h}\}. \tag{A.34}$$

Suppose that $lim\ sup_{T\to\infty}\sqrt{T}(FEVD_{ij^*h} - l_{ij^*h}(\boldsymbol{\phi}(P))) < \infty$. Given the consistency of $\widehat{\Delta}$, divergence of $\sqrt{T}\Delta(P)$ implies divergence (in probability) of $\sqrt{T}\widehat{\Delta}$. As a result,

$$Pr\{\sqrt{T}\,[FEVD_{ij^*h}(P) - l_{ij^*h}(\boldsymbol{\phi}(P))] + \sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})] > \sqrt{T}\widehat{\Delta} + c_\alpha \widehat{\sigma}_{uij^*h}\} \leq \tag{A.35}$$

$$Pr\{\sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})] > \sqrt{T}\widehat{\Delta} - \sqrt{T}\,[FEVD_{ij^*h}(P) - l_{ij^*h}(\boldsymbol{\phi}(P))]\} \to 0, \tag{A.36}$$

where $c_\alpha \widehat{\sigma}_{uij^*h} \geq 0$ by construction and $\sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})]$ converges by previous results in this paper. We therefore derive the following

$$lim_{T\to\infty} Pr(FEVD_{ij^*h}(P) \in CI(\alpha)) \tag{A.37}$$

$$= lim_{T\to\infty} Pr\{-c_\alpha \widehat{\sigma}_{lij^*h} \leq \sqrt{T}\,[FEVD_{ij^*h}(P) - l_{ij^*h}(\boldsymbol{\phi}(P))] + \sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})]\} \tag{A.38}$$

$$\geq lim_{T\to\infty} Pr\{-c_\alpha \widehat{\sigma}_{lij^*h} \leq \sqrt{T}[l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})]\} = 1 - \Phi(c_\alpha) \geq 1 - \alpha. \tag{A.39}$$

Note that the first inequality comes from $\sqrt{T}(FEVD_{ij^*h} - l_{ij^*h}(\boldsymbol{\phi}(P))) \geq 0$; the last inequality relies on the definition of c_α and convergence of $\widehat{\sigma}_{lij^*h}$ and $\sqrt{T}\frac{l_{ij^*h}(\boldsymbol{\phi}(P)) - l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})}{\sigma_{lij^*h}(P)}$. For any subsequence of P where $\sqrt{T}(FEVD_{ij^*h} - u_{ij^*h}(\boldsymbol{\phi}(P)))$ does not diverge, the same argument applies. If both diverge, coverage probability converges to 1. To see that we can obtain a coverage probability of $1 - \alpha$, consider $\Delta(P) = 0$. ■

Proof of Corollary 3.2.

For notation, see the previous theorem. Here we show that CI^r_α satisfies the conditions (i) to (iii) in the proof of Theorem 3.3; the result follows.

Let $\Delta(P) \leq a_T$. Note the following.

(i) Given Assumption A7, the delta-method delivers a Normal distribution for $u^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ and $l^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$

$$\sqrt{T}\begin{bmatrix} l^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \\ u^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \end{bmatrix} \to_d \boldsymbol{N} \tag{A.40}$$

centered at $\begin{bmatrix} l^r_{ij^*h}(\boldsymbol{\phi}(P)) \\ u^r_{ij^*h}(\boldsymbol{\phi}(P)) \end{bmatrix}$. The convergence in distribution is uniform (formally, see Theorem 3 in Kasy (2018)).

(ii) For any P, $\sigma^{r2}_{lij^*h}(P)$ and $\sigma^{r2}_{uij^*h}(P)$ are positive (recall the non-zero condition) and finite and $u^r_{ij^*h}(\boldsymbol{\phi}(P)) - l^r_{ij^*h}(\boldsymbol{\phi}(P)) < \infty$. The latter comes from the fact that $|r_{ij^*h}| \leq ||c_{ih}(\boldsymbol{\phi})|| < \infty$ for any $i \in 1, \ldots, n$, $j^* \in 1, \ldots, n$ and $h = 0, 1, \ldots,$ where $||c_{ih}(\boldsymbol{\phi})||$ is bounded to the restriction on $\boldsymbol{\phi}$ such that the reduced-form VAR is invertible into $VMA(\infty)$.

(iii) Since $u^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ and $l^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ are the bounds, they can be ordered such that $u^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \geq l^r_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS})$ for all P.

The argument in Theorem 3.3 delivers the result. ■

Proof of Theorem 3.4.

This comes from applying the delta-method to the posterior distribution of $\boldsymbol{\phi}$ subject to the class of priors $\boldsymbol{\phi}^* \sim P^*_{\boldsymbol{\phi}}$ in Assumption A8. The posterior distribution for $\boldsymbol{\phi}^*$ is

$$\sqrt{T}(\boldsymbol{\phi}^*|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T) \rightarrow_d \boldsymbol{N}(\widehat{\boldsymbol{\phi}}_{OLS}, \boldsymbol{\Omega}(P)). \tag{A.41}$$

We apply the delta-method to the distribution above:

$$\sqrt{T}\begin{bmatrix} l_{ij^*h}(\widehat{\boldsymbol{\phi}}^*) \\ u_{ij^*h}(\widehat{\boldsymbol{\phi}}^*) \end{bmatrix} \Bigg| \boldsymbol{y}_1, \ldots, \boldsymbol{y}_T \rightarrow_d \boldsymbol{N}\left(\begin{bmatrix} l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \\ u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \end{bmatrix}, \begin{bmatrix} \sigma^2_{lij^*h}(P) & 0 \\ 0 & \sigma^2_{uij^*h}(P) \end{bmatrix} \right). \tag{A.42}$$

Applying the result above to the class of priors $\boldsymbol{\phi}^* \sim P^*_{\boldsymbol{\phi}}$ disciplined by Assumption A8 (and recalling the definition of a Borel set) yields

$$sup_{\boldsymbol{BO} \in \boldsymbol{\mathcal{BO}}(\mathcal{R}^d)} \left\{ P^*_{\boldsymbol{\phi}} \left(\sqrt{T} \left(\begin{bmatrix} l_{ij^*h}(\boldsymbol{\phi}^*) \\ u_{ij^*h}(\boldsymbol{\phi}^*) \end{bmatrix} - \begin{bmatrix} l_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \\ u_{ij^*h}(\widehat{\boldsymbol{\phi}}_{OLS}) \end{bmatrix} \right) \in \boldsymbol{BO} | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_T \right) - \mathcal{P}(\tilde{\boldsymbol{\xi}}(P) \in \boldsymbol{BO}) \right\} \rightarrow_p 0, \tag{A.43}$$

where $\tilde{\boldsymbol{\xi}}(P) \sim \boldsymbol{N}\left(\boldsymbol{0}, \begin{bmatrix} \sigma^2_{lij^*h}(P) & 0 \\ 0 & \sigma^2_{uij^*h}(P) \end{bmatrix}\right)$. ■

B Empirical Appendix

In Figure B.9 and B.10, we compare the estimated bounds with the bias-corrected ones and with those obtained by using Gafarov et al.'s (2018) IRFs as a plug-in, respectively. Figure B.11, which is instrumental to confirm the large identification uncertainty implied by the identification scheme in the main text, shows the dynamics of the IRFs by using the delta-method in Gafarov et al. (2018) and the classical Bayesian approach.

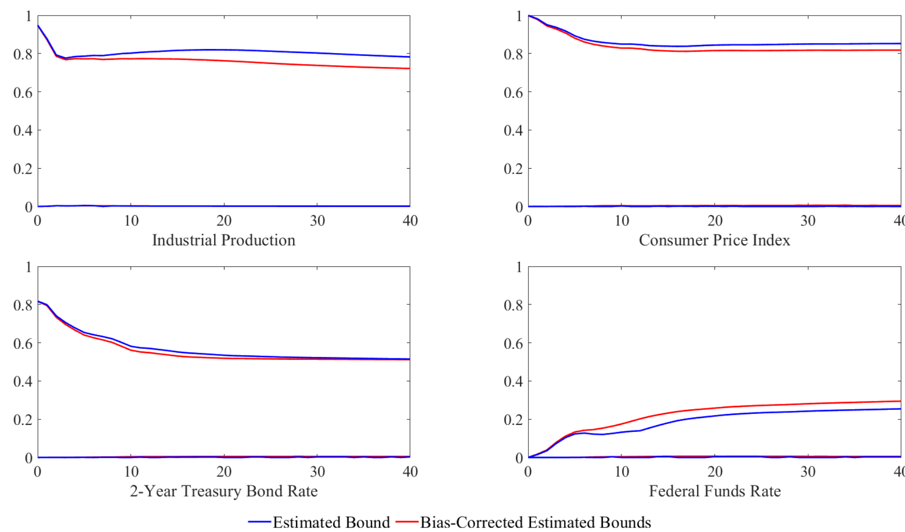

Figure B.9: Estimated Bounds *vs* Bias-Corrected Estimated Bounds.

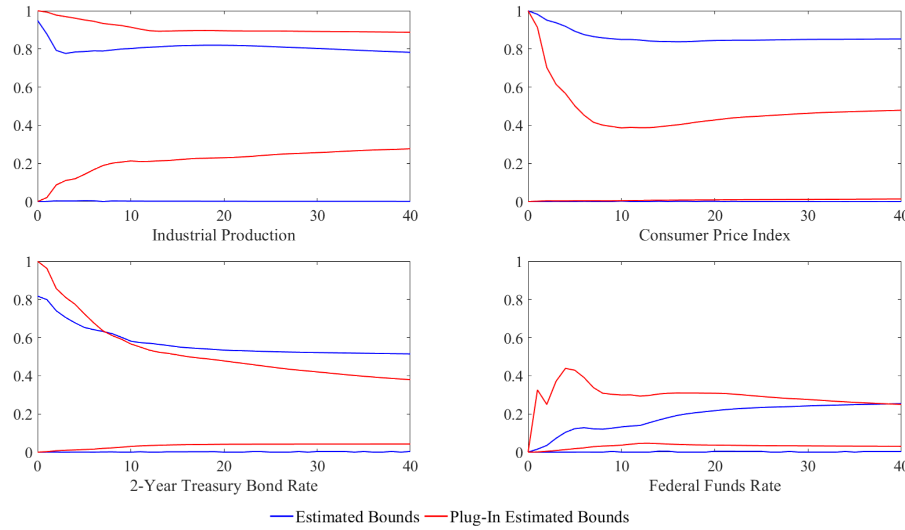

Figure B.10: Estimated Bounds *vs* Plug-In Estimated Bounds
(based on Gafarov et al.'s (2018) IRFs).

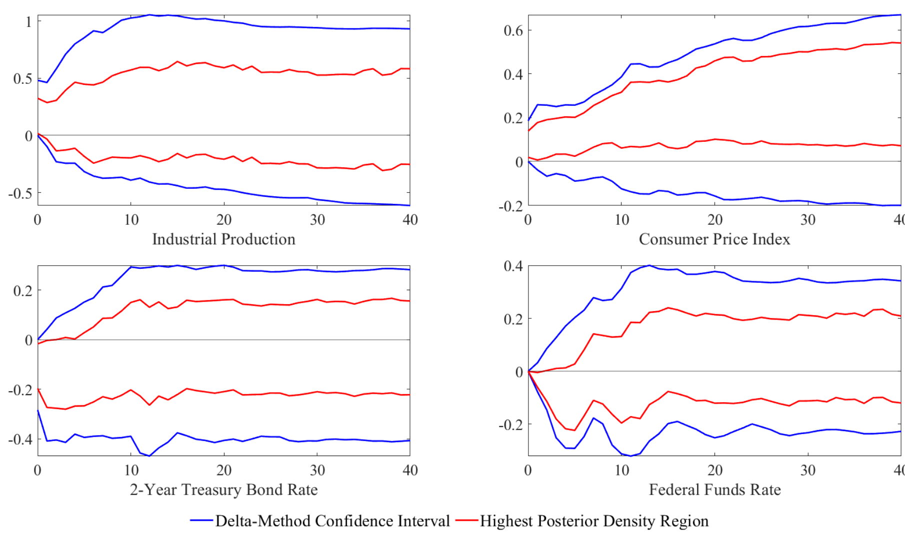

Figure B.11: 68% Delta-Method Confidence Interval *vs* 68% Highest Posterior Density Region
Under Uniform Prior on $\boldsymbol{Q}$.

C Further Simulation Results

Figure C.12 displays the frequentist coverage for the lower bound. As discussed in the main text, in the DGP the lower bound is often close to 0.

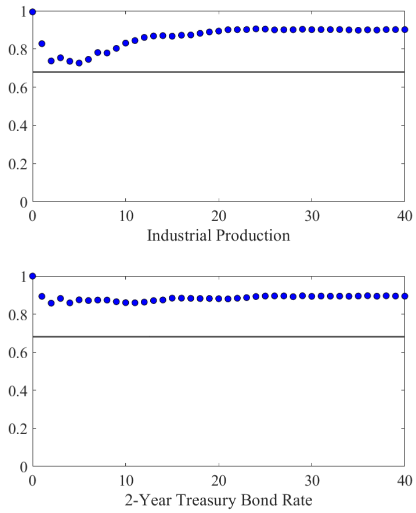

Figure C.12: Monte-Carlo Exercise: Frequentist Coverage (Lower Bound).

Figure C.13 shows how the set-length correction to the critical value is paramount to obtain less conservative intervals for the lower bound. In Figure C.14, we show how the same result for the lower bound also applies to the IRFs. In Figure C.15, we instead assess the coverage of GK21's robust credible interval with respect to the lower bound.

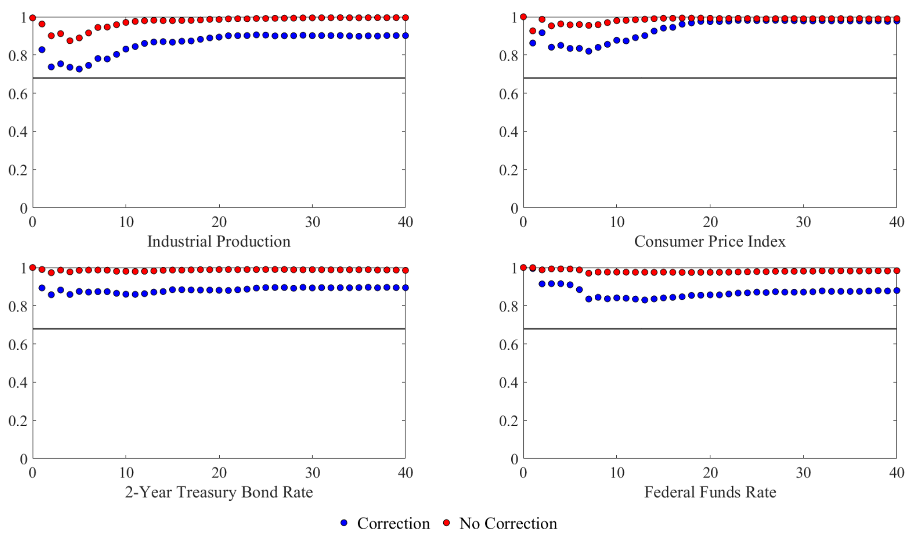

Figure C.13: Monte-Carlo Exercise: Lower Bound Frequentist Coverage With and Without Set-Length Correction to the Critical Value (FEVD).

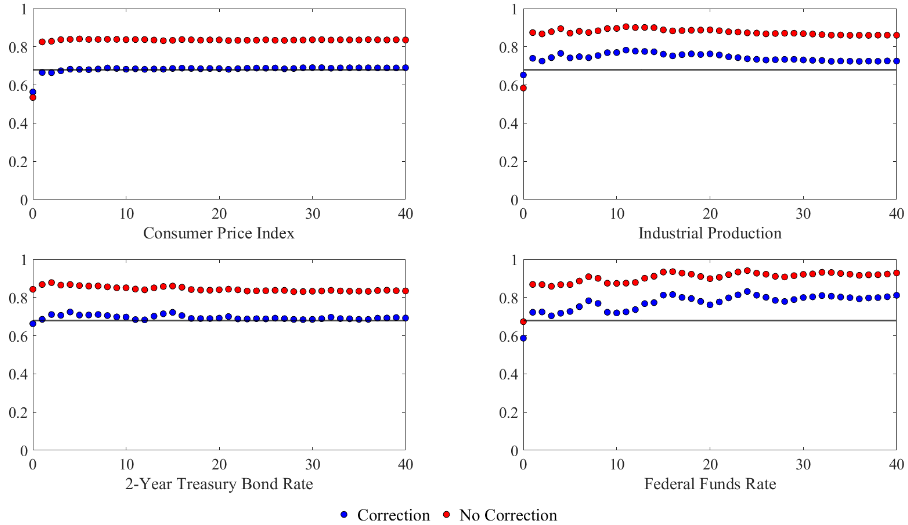

Figure C.14: Monte-Carlo Exercise: Lower Bound Frequentist Coverage With and Without (Gafarov et al.'s (2018)) Set-Length Correction to the Critical Value (IRFs).

In order to evaluate the coverage of GK21's robust credible interval, we proceed as follows:

1. Generate a random sample by keeping the VAR coefficients fixed at $\boldsymbol{B}_{OLS}$, and drawing the reduced-form errors $\boldsymbol{u}_t$ from $\boldsymbol{u}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\text{OLS}})$.
2. Compute the robust credible interval based on the sample obtained in step 1.[20]
3. Repeat step 1 and 2 M times.[21]
4. Compute coverage as the percentage of the M robust credible intervals that contain the estimated mean upper and lower bounds.

As shown below, and similarly to what we discussed in Section 4, the coverage for the lower bound is close to 1 for all variables and across all horizons.

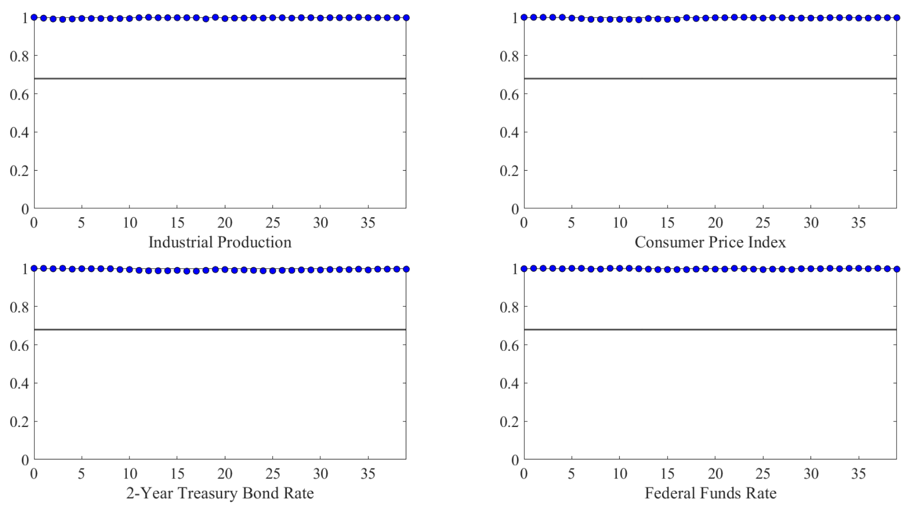

Figure C.15: Monte Carlo Exercise: Coverage of GK's Robust Credible Interval (Lower Bound).

Finally, Figure C.16 displays the Bayesian coverage of our confidence inteval for the lower bound.

[20]Bayesian estimation is performed by using a Normal-Inverse-Wishart prior with 1000 draws for the reduced-form coefficients and 1000 draws for $\boldsymbol{Q}$.

[21]To limit the computational burden, M is set equal to 500.

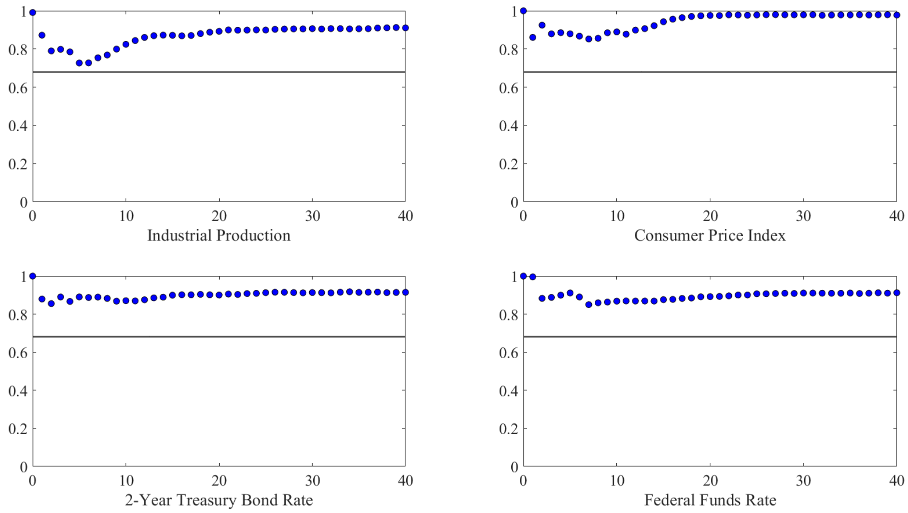

Figure C.16: Monte-Carlo Exercise: Robust Bayesian Credibility (Lower Bound).