DELTA-METHOD INFERENCE FOR A CLASS OF SET-IDENTIFIED SVARS

Bulat Gafarov[1], Matthias Meier[2] and José Luis Montiel Olea[3]

This paper studies Structural Vector Autoregressions that impose equality and/or inequality restrictions to set-identify a single shock (e.g., a monetary shock). We make three contributions to the literature.

(i) We present an algorithm to compute—for each horizon, each variable, a fixed vector of reduced-form parameters, and a given collection of equality and/or inequality restrictions—the largest and smallest value of the coefficients of the structural impulse-response function.

(ii) We provide conditions under which the largest and smallest value of the structural parameters are *directionally differentiable* functions of the reduced-form parameters.

(iii) We propose a computationally convenient *delta-method* confidence interval for the set-identified coefficients of the structural impulse-response function. We present sufficient conditions to guarantee the pointwise consistency in level of the suggested inference approach.

To illustrate our results, we use a monetary Structural Vector Autoregression estimated with monthly U.S. data. We set-identify an *unconventional* monetary policy shock that decreases the two-year government bond rate upon impact, but has no effect over the nominal federal funds rate. We impose two additional sign restrictions on the contemporaneous responses of inflation and output. We use our confidence bands to assess the effects of the announcement of the second part of the so-called Quantitative Easing program in August 2010. (JEL-Classification: C1, C32, E47).

KEYWORDS: Set-Identification, SVAR, Delta Method, Directional Differentiability, Unconventional Monetary Policy.

1. INTRODUCTION

An increasingly popular practice in empirical macroeconomics is to set-identify the parameters of a Structural Vector Autoregression [SVAR]. This approach was pioneered by Faust (1998), Canova and Nicoló (2002) and Uhlig (2005). Most of the follow-up studies have relied on Bayesian methods to construct posterior credible sets for the structural coefficients of the impulse-response function.

There has been recent interest in studying non-Bayesian approaches to summarize uncertainty in set-identified SVARs. Moon, Schorfheide, and Granziera (2013) [MSG13] pro-

pose Projection/Bonferroni frequentist inference based on a moment-inequality-minimum-distance framework. Giacomini and Kitagawa (2014) [GK14] propose robust-Bayesian inference using multiple priors for rotation matrices. Gafarov, Meier, and Montiel Olea (2016) [GMM16] propose frequentist inference based on the projection of a Wald ellipsoid for the SVAR reduced-form parameters. None of these approaches requires the specification of prior beliefs by the researcher.

This paper contributes to the non-Bayesian analysis of set-identified SVARs by proposing a novel delta-method confidence interval for the coefficients of the impulse-response function [IRF]. Broadly speaking, our approach is based on a closed-form characterization of the endpoints of the identified set (given a vector of reduced-form parameters and a collection of binding inequality constraints). Our delta-method confidence interval takes the form of a plug-in estimator for the identified set plus/minus standard errors. In terms of theoretical results, we establish the pointwise consistency in level of our confidence interval. In terms of practical considerations, we argue that the computational cost of our procedure compares very favorably with other non-Bayesian procedures and also with the standard Bayesian algorithm described in Uhlig (2005).

The main limitation of our approach is that the delta-method confidence interval is only defined for SVAR models that impose equality and inequality restrictions on a single structural shock (e.g., a monetary shock). Admittedly, this is problematic, as some popular applications of set-identified SVARs feature restrictions on multiple structural innovations.[1] In spite of this observation, single-shock set-identified models have been applied in several empirical studies: the effects of monetary policy on output [Uhlig (2005)], the impact of monetary policy on the housing market [Vargas-Silva (2008)], the effects of labor market shocks on worker flows [Fujita (2011)], the effects of exchange rates on aggregate prices [An and Wang (2012)], and the effect of optimism shocks on business cycles fluctuations [Beaudry, Nam, and Wang (2014)]. Thus, we think there is room for our results to have an impact on empirical work.

EMPIRICAL APPLICATION—UNCONVENTIONAL MONETARY POLICY SHOCKS: To illustrate the usefulness of our main results, we estimate a monetary Structural Vector Autoregression using monthly U.S. data from July 1979 to December 2007 (a sample that deliberately ends one semester before the financial crisis begins). The goal of our exercise is to use pre-crisis data to learn about the responses of macroeconomic variables to shocks that have effects similar to the 'unconventional' monetary policy interventions implemented after the crisis.

In 'conventional' descriptions of monetary policy, the short-term nominal interest rate

---

[1]SVAR applications for the oil market set-identify both demand and supply shocks using sign restrictions and elasticity bounds [Kilian and Murphy (2012)]. The same is true for recent labor market applications [Baumeister and Hamilton (2015)]. Mountford and Uhlig (2009)—one of the most cited applications of set-identified SVARs—use sign restrictions to identify a government revenue shock as well as a government spending shock, while controlling for a generic business cycle shock and a monetary policy shock.

is assumed to be the central bank's policy instrument. Following any adjustment by the monetary authority, the market participants—households and firms, both domestic and foreign—use available information to form expectations about the future level of longer-term real interest rates relevant for their consumption and investment decisions.

The recent *Great Recession* has forced the Federal Reserve to consider alternative mechanisms to affect market beliefs about the future of real interest rates. Two examples of such unconventional policies are the Federal Open Market Committee's *forward guidance* announcements and the Federal Reserve's *large-scale asset purchases program*. Broadly speaking, through forward guidance "the Federal Open Market Committee provides an indication to households, businesses, and investors about the stance of monetary policy expected to prevail in the future".[2] In a similar fashion, the asset purchase program of the Federal Reserve intends to "put downward pressure on yields of a wide range of longer-term securities, support mortgage markets, and promote a stronger economic recovery".[3]

With this motivation in mind, we set-identify an *unconventional* monetary policy [UMP] shock as an innovation that decreases the two-year government bond rate upon impact, but has no effect over the nominal federal funds rate.[4] We consider two additional sign restrictions on the contemporaneous responses of inflation and output. Namely, we assume that—upon impact—neither inflation nor output can respond negatively to a UMP shock. Since the model is only set-identified, our analysis effectively captures the effects of any historical economic shock that affected the economy in the same way as an UMP shock.

We apply our delta-method approach to construct a confidence interval for the dynamic responses of Industrial Production, inflation, the two-year government bond rate, and the nominal federal funds rate. We use our confidence bands to assess the effects of the announcement of the second part of the so-called Quantitative Easing program (QE2) in August 2010. Pre-crisis data turns out to be extremely useful to learn about the post-crisis response of macroeconomic aggregates to unconventional monetary policy.

The remainder of the paper is organized as follows. Section 2 presents an overview of the main methodological results in this paper. Section 3 introduces our empirical application, which is used as a running example throughout the paper. Section 4 presents our algorithm to evaluate the endpoints of the identified set. Section 5 establishes the differentiability properties of the endpoints. Section 6 presents our delta-method approach and establishes its asymptotic validity. Section 7 presents the confidence intervals for the dynamic responses to the QE2 program. Section 8 concludes. All the proofs are collected in the Appendix.

GENERIC NOTATION: If A is a matrix, A_{ij} denotes the ij-th element of A, vec(A) denotes the vectorization of A, and vech(A) denotes half-vectorization (applicable only if A is sym-

---

[2]Link: http://www.federalreserve.gov/faqs/money_19277.htm

[3]Link: http://www.federalreserve.gov/faqs/what-are-the-federal-reserves-large-scale-asset-purchases.htm

[4]The paper focuses on the two-year rate as this variable changed considerably after the announcement of the second round of the Quantitative Easing program. See Krishnamurthy and Vissing-Jorgensen (2011)

metric). The Kronecker product between matrices A and B is denoted by $A \otimes B$. The vector $e_i^m \in \mathbb{R}^m$ denotes the i-th column of the identity matrix of dimension m. If B is a matrix of dimension $n \times n$, $B_i \equiv Be_i^n$ denotes its i-th column. If the dimension of e_i^n is obvious, we ignore the superscript n.

2. MODEL AND OVERVIEW OF THE MAIN THEORETICAL RESULTS

This section introduces notation, the class of SVAR models we consider, and presents a brief overview of the main methodological results in the paper. It is our hope that this brief summary (which contains references to the main propositions and lemmas in the paper) contributes to the understanding of the theoretical basis behind our delta-method confidence interval.

NOTATION: This paper studies the n-dimensional Structural Vector Autoregression (SVAR) with p lags; i.i.d. structural innovations distributed according to F; and unknown $n \times n$ structural matrix B:

$$(2.1) \qquad Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + B\varepsilon_t, \ \ \mathbb{E}_F[\varepsilon_t] = 0_{n \times 1}, \ \ \mathbb{E}_F[\varepsilon_t \varepsilon_t'] \equiv \mathbb{I}_n,$$

see Lütkepohl (2007), p. 362.

The object of interest is the k-th period ahead structural impulse response function of variable i to a particular shock j (e.g., a monetary shock). In the SVAR model this parameter is given by the (k, i, j)-*coefficient of the structural impulse-response function*:

$$(2.2) \qquad \lambda_{k,i,j}(A, B) \equiv e_i' C_k(A) B_j,$$

where $B_j \equiv Be_j$ and e_i and e_j denote the i-th and j-th column of the identity matrix $\mathbb{I}_n$.[5]

An auxiliary object in the estimation of the structural parameters is the vector of *reduced-form parameters* in the SVAR model:

$$(2.3) \qquad \mu \equiv (\text{vec}(A)', \text{vec}(\Sigma)')' \in \mathcal{M} \subseteq \mathbb{R}^d, \quad A \equiv (A_1, A_2, \ldots, A_p), \quad \Sigma \equiv BB'.$$

The parameter A denotes the autoregressive coefficients in the VAR model, while Σ denotes the covariance matrix of residuals. These parameters can be estimated directly from the data by Ordinary Least-Squares. The (reduced-form) parameter space is $\mathcal{M}$.

SET-IDENTIFYING RESTRICTIONS: Let $\mathcal{R}(\mu) \subseteq \mathbb{R}^n$ be a set of inequality and equality

---

[5]The transformation $C_k(A)$ that appears in equation (2.2) is defined recursively by the formula $C_0 \equiv \mathbb{I}_n$:

$$C_k(A) \equiv \sum_{m=1}^{k} C_{k-m}(A)\, A_m, \quad k \in \mathbb{N},$$

$A_m = 0$ if $m > p$; see Lütkepohl (1990), p. 116.

restrictions imposed on B_j.[6] A common practice in empirical macroeconomics is to use the restrictions in $\mathcal{R}(\mu)$ to *set-identify* the structural parameter in (2.2) as a function of the reduced-form parameters in (2.3). In our paper, the set $\mathcal{R}(\mu)$ takes the form:

$$(2.4) \qquad \mathcal{R}(\mu) \equiv \left\{ B_j \in \mathbb{R}^n \mid Z(\mu)'B_j = \mathbf{0}_{m_z \times 1} \text{ and } S(\mu)'B_j \geq \mathbf{0}_{m_s \times 1} \right\},$$

where $Z(\mu)$ is a matrix of dimension $n \times m_z$ matrix and $S(\mu)$ is a matrix of dimension $n \times m_s$. The matrix $Z(\mu)$ collects the equality restrictions specified by the researcher (we assume there are m_z of them). The matrix $S(\mu)$ collects the inequality restrictions (we assume there are m_s of them). We assume that both $Z(\mu)$ and $S(\mu)$ are differentiable functions of μ.

SCOPE: The simple formulation in (2.4) allows the researcher to incorporate any restriction of the form $R(\mu)'B_j \geq 0$, where $R(\mu)$ is differentiable. Thus, our analysis allows for the following identifying restrictions:

a) Restrictions on the responses of variable i at horizon k to an impulse on the j-th shock:

$$e_i'C_k(A)B_j \geq \text{ or } = 0,$$

as in Uhlig (2005).

b) Long-run restrictions on the response of variable i to an impulse on the j-th shock:

$$e_i'(\mathbb{I}_n - A_1 - \ldots - A_p)^{-1}B_j \geq \text{ or } = 0,$$

as in Blanchard and Quah (1989).

c) Restrictions on the j-th column of $(H')^{-1}$:

$$e_i'(H')^{-1}e_j = e_i'\Sigma^{-1}B_j \geq \text{ or } = 0,$$

as in Rubio-Ramirez, Caldara, and Arias (2015).

d) Elasticity bounds as in Kilian and Murphy (2012); for example, for some $b \in \mathbb{R}$:

$$e_i'B_j/e_{i'}'B_j \geq b \iff (e_i - be_{i'})'B_j \geq 0,$$

provided $e_{i'}'B_j > 0$.

ENDPOINTS OF THE IDENTIFIED SET: The main results in this paper concern the endpoints of the identified set for the structural parameters, given μ. The endpoints of the identified set (which we sometimes refer to as the *maximum and minimum* response) are defined as follows:

---

[6]For example, a contractionary monetary shock increases interest rates and does not affect prices upon impact.

DEFINITION (ENDPOINTS OF THE IDENTIFIED SET): Given a vector of reduced-form parameters μ we define the endpoints of the identified set for $\lambda_{k,i,j}$ as the functions:

$$(2.5) \qquad \overline{v}_{k,i,j}(\mu) \equiv \sup_{B \in \mathbb{R}^{n \times n}} e_i' C_k(A) B e_j, \text{ s.t. } BB' = \Sigma \text{ and } B e_j \in \mathcal{R}(\mu),$$

and

$$(2.6) \qquad \underline{v}_{k,i,j}(\mu) \equiv \inf_{B \in \mathbb{R}^{n \times n}} e_i' C_k(A) B e_j, \text{ s.t. } BB' = \Sigma \text{ and } B e_j \in \mathcal{R}(\mu).$$

The function $\overline{v}_{k,i,j}(\mu)$ corresponds to the largest value of the structural parameter, $\lambda_{k,i,j}$ subject to the restriction that $B_j \in \mathcal{R}(\mu)$ and also that B_j is the j-th column of a square root of Σ. The lower bound is defined analogously.

OVERVIEW OF THE MAIN RESULTS: Our delta-method confidence interval is supported by the three theoretical results described in the abstract. Our results can be summarized as follows:

- SUMMARY OF LEMMA 1 *(Characterization of the maximum and minimum response given a fixed set of active constraints)*: We show that $\overline{v}_{k,i,j}(\mu)$ and $\underline{v}_{k,i,j}(\mu)$ are the *value functions* of a mathematical program whose Karush-Kuhn-Tucker points can be described analytically—up to a set of 'active' inequality constraints.[7] More concretely, we show that the maximum response for $\lambda_{k,i,j}$ is equal to either plus or minus the function:

$$v_{k,i,j}(\mu; r) \equiv \left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2}r} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2},$$

where

$$M_{\Sigma^{1/2}r} \equiv \mathbb{I}_n - \Sigma^{1/2} r (r' \Sigma r)^{-1} r' \Sigma^{1/2},$$

and r' is a matrix collecting the gradient vectors of the constraints in $\mathcal{R}(\mu)$ that are active at a maximum. The minimum response is obtained analogously.
- SUMMARY OF PROPOSITION 1 *(Algorithm to evaluate the maximum and minimum response)*: We use the closed-form expressions of Lemma 1 to present an algorithm that allows a researcher to evaluate the endpoints of the identified set given a vector of reduced-form parameters. The algorithm evaluates different collections of active constraints (different matrices r) and selects the constraints that generate the largest (or smallest) value function—after checking that the inequality constraints not included in r are satisfied.
- SUMMARY OF LEMMA 2: *(Differentiability of the maximum and minimum response for a*

---

[7]The term 'active constraints' or 'active set of is constraints' is the common terminology used in numerical optimization; see p. 308 in Nocedal and Wright (2006).

fixed set of active constraints) We establish the differentiability of the function $v_{k,i,j}$, which depends on a fixed set of active constraints. We allow the constraints in the $n \times l$ matrix r (with $l \leq n-1$) to depend on the reduced-form parameters. We show that the derivative of $v_{k,i,j}$ w.r.t. μ is given by:

$$\dot{v}_{k,i,j}(\mu) \equiv \begin{bmatrix} \frac{\partial v_{k,i,j}(\mu;r)}{\partial \text{vec}(A)} \\[2mm] \frac{\partial v_{k,i,j}(\mu;r)}{\partial \text{vec}(\Sigma)} \end{bmatrix} = \begin{bmatrix} \frac{\partial \text{vec}(C_k(A))}{\partial \text{vec}(A)}(x^* \otimes e_i^n) - \sum_{k=1}^{l} w_k^* \frac{\partial r_k}{\text{vec}(A)} x^* \\[2mm] \lambda^*(\Sigma^{-1} x^* \otimes \Sigma^{-1} x^*) - \sum_{k=1}^{l} w_k^* \frac{\partial r_k}{\text{vec}(\Sigma)} x^* \end{bmatrix},$$

where r_k is the k-th column of r, w_k^* is the k-th component of w^*, and

$$x^* \equiv \Sigma^{1/2}\Big(M_{\Sigma^{1/2}r}\Big)\Sigma^{1/2}C_k(A)'e_i\Big/v_{k,i,j}(\mu;r),$$

$$\lambda^* \equiv \frac{1}{2}v_{k,i,j}(\mu;r), \quad w^* \equiv [r'\Sigma r]^{-1}r'\Sigma C_k(A)'e_i.$$

We argue that λ^* and w^* can be interpreted as the Lagrange multipliers associated to the constraints $BB' = \Sigma$ and to the active constraints in r.

• SUMMARY OF PROPOSITION 2 *(Directional Differentiability of the endpoints)*: We use the formula in Lemma 2 to show that the functions $\overline{v}_{k,i,j}(\cdot)$ and $\underline{v}_{k,i,j}(\cdot)$ are *directionally* differentiable, in a sense we make precise. We relate the expression of the directional derivative with the generalized versions of the envelope theorems in the work of Fiacco and Ishizuka (1990) and Bonnans and Shapiro (2000). We argue that directional differentiability of the value functions (as opposed to full differentiability) arises due to the possibility that different structural models lead to the maximum (or minimum) response. In particular, let $R^*(\mu)$ denote the sets of active restrictions that yield the same maximum response and assume, for simplicity, that $\overline{v}_{k,i,j}(\mu) > 0$. We show that:

$$\sqrt{n}\Big(\overline{v}_{k,i,j}(\mu + h_n/\sqrt{n}) - \overline{v}_{k,i,j}(\mu)\Big) \to \max_{r \in R^*(\mu)} \Big[\dot{v}_{k,i,j}(\mu;r)'h\Big].$$

• SUMMARY OF PROPOSITION 3 *(Delta-Method Confidence Interval)*: We establish the pointwise consistency in level of a delta-method confidence interval, which takes the form:

$$\text{CS}_T(1-\alpha; \lambda_{k,i,j}) \equiv \Big[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\,\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}, \ \ \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\,\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}\Big],$$

where $\widehat{\mu}_T$ is the typical OLS estimator for the VAR reduced-form parameters, $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of a standard normal, and $\widehat{\sigma}_{(k,i,j),T}$ is our formula for the standard errors based on the directional derivatives.

OUTLINE: In the remaining part of the paper, we formalize these propositions and apply them to conduct inference about the responses to an *unconventional monetary shock*.

3. RUNNING EXAMPLE: UNCONVENTIONAL MONETARY SHOCKS

This section introduces our empirical application, which will be used as a running example to illustrate our assumptions and results.

MONETARY SVAR: We consider a simple 4-variable model that includes the Consumer Price Index (CPI_t), the Industrial Production Index (IP_t), the 2-year Treasury Bond rate ($2yTB_t$), and the Federal Funds rate (FF_t).[8] We take a logarithmic transformation of CPI_t, IP_t and then work with first differences for all variables. Thus, our vector of macro variables is:

$$Y_t \equiv \Big(\ln CPI_t - \ln CPI_{t-1}, \quad \ln IP_t - \ln IP_{t-1}, \quad 2yTB_t - 2yTB_{t-1}, \quad FF_t - FF_{t-1} \Big)'.$$

We set the number of lags equal to 11 using the Bayesian Information Criterion ($p = 11$). The time span of the monthly series is July 1979 to August 2008 ($T = 342$). To keep our exposition as simple as possible, we ignore potential co-integration issues between short-term and long-term interest rates. Without loss of generality, we assume that column of B corresponding to an UMP shock is the first column; $B_1 \equiv Be_1$. Our equality/inequality restrictions are summarized in the following Table:

TABLE I

SET-IDENTIFICATION OF AN UNCONVENTIONAL MONETARY POLICY SHOCK: RESTRICTIONS

Series	Acronym	UMP	Notation
Consumer Price Index	CPI	$+$	$e_1' B_1 \geq 0$
Industrial Production	IP	$+$	$e_2' B_1 \geq 0$
2-year Treasury Bond rate	2yTB	$-$	$e_3' B_1 \leq 0$
Fed Funds Rate	FF	0	$e_4' B_1 = 0$

DESCRIPTION: Restrictions on contemporaneous responses to a UMP shock. '0' stands for a zero restriction, '$-$' stands for a negative sign restriction and '$+$' for positive sign restriction. These sign restrictions can be justified by the DSGE model calibrated in the work of Bhattarai, Eggertsson, and Gafarov (2014).

The suggested set-identification strategy in this paper is not new. Baumeister and Benati (2013) study an analogous 'spread' monetary shock that leaves the short-term nominal rate unchanged, but affects the spread between the ten-year Treasury-bond yield and the policy rate. They consider a Bayesian SVAR with time varying parameters and stochastic volatility combined with demand and supply structural shocks that satisfy zero/sign restrictions as in Rubio-Ramirez, Waggoner, and Zha (2010). Their main result is that the long-term yield spread exerts a powerful effect on both output growth and inflation. All their inference is Bayesian, while ours is frequentist. In addition, our SVAR model does not consider time-varying parameters, stochastic volatility, and restrictions on other nonmonetary shocks.

---

[8]All these variables are sourced from the data set of Gertler and Karadi (2015). We thank Peter Karadi for making their data set available to us.

4. **THE ENDPOINTS OF THE IDENTIFIED SET**

In this section we formalize Lemma 1 and Proposition 1. We consider the problem of finding the maximum response to an impulse in the j-th structural shock subject to m_z equality ('zero') restrictions and m_s inequality ('sign') restrictions. The focus on the maximum and the minimum is an intermediate step to conduct frequentist inference about the coefficients of the impulse-response function.

This section makes two assumptions on the sign and zero restrictions allowed in the model. First, we require the number of zero restrictions to be less than $n-1$. Second, we assume that every collection of $n-1-m_z$ inequality restrictions and the m_z equality restrictions are linearly independent everywhere in the parameter space.

Just as before, the set $\mathcal{R}(\mu)$ is given by

$$\mathcal{R}(\mu) \equiv \left\{ B_j \in \mathbb{R}^n \mid Z(\mu)'B_j = \mathbf{0}_{m_z \times 1} \text{ and } S(\mu)'B_j \geq \mathbf{0}_{m_s \times 1} \right\},$$

where $Z(\mu)$ is a matrix of dimension $n \times m_z$ matrix and $S(\mu)$ is a matrix of dimension $n \times m_s$.

RUNNING EXAMPLE—$\mathcal{R}(\mu)$: In the UMP example, the set of restrictions $\mathcal{R}(\mu)$ corrresponds to (see Table I):

$$\left\{ B_1 \in \mathbb{R}^4 \mid e_4'B_1 = 0, (e_1, e_2, -e_3)'B_1 \geq \mathbf{0}_{3 \times 1} \right\}.$$

Consequently:

$$Z(\mu) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } S(\mu) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

We note that the equality and inequality restrictions in our example do not depend on the reduced-form parameters (neither Z nor S depend on μ).

MAIN ASSUMPTIONS: The first assumption in this section requires the number of zero restrictions to be strictly smaller than $n-1$. The rationale behind Assumption 1 is as follows: if $m_z > n-1$, then $\mathcal{R}(\mu) = \{\mathbf{0}_{n \times 1}\}$ for every μ for which there are n linearly independent equality restrictions. This is problematic, as the latter implies there is no $B_j \in \mathbb{R}^n$ such that $B_j \in \mathbb{R}^n$ and $B_j = Be_j$ for some $BB' = \Sigma$ (provided B is invertible).[9]

ASSUMPTION 1 $m_z \leq n-1$.

---

[9]If B is invertible, then Σ is invertible and $B'\Sigma^{-1}B = \mathbb{I}_n$, which implies $B_j'\Sigma^{-1}B_j = 1$. Therefore, if $B_j = Be_j$ for some square root of Σ then B_j must be different from $\mathbf{0}_{n \times 1}$.

Our second assumption imposes a 'linear independence' condition on the equality and inequality restrictions on the model (given a particular value of the reduced-form parameter μ). Let $e_1^{m_s}, e_2^{m_s}, \ldots e_{m_s}^{m_s}$ denote the m_s different columns of the identity matrix $\mathbb{I}_{m_s}$. Let $e(k)$ denote an $m_s \times k$ matrix formed by collecting any of the $k \leq n - 1 - m_z$ columns of $\mathbb{I}_{m_s}$. Note that for any matrix S, the matrix $Se(k)$ selects k columns of S.

DEFINITION   We say that $Z(\mu)$ and $S(\mu)$ are linearly independent at μ if for any $k \in \mathbb{N}, k \leq n - 1 - m_z$ and any $e(k)$ the matrix

$$R(\mu; e(k)) \equiv [Z(\mu), S(\mu)e(k)] \in \mathbb{R}^{n \times (m_z + k)},$$

is assumed to have full column rank (rank $m_z + k$).

We use this definition to state our following assumption:

ASSUMPTION 2   The parameter space $\mathcal{M}$ is such that $Z(\mu)$ and $S(\mu)$ are linearly independent at every $\mu \in \mathcal{M}$.

This 'linear independence' property plays an important role in the characterization of the maximum and minimum response in terms of *Karush-Kuhn-Tucker* conditions.

RUNNING EXAMPLE—ASSUMPTION 1 AND 2: In our application $m_z = 1$, so we only need to verify Assumption 2. As we mentioned before, the matrix $e(k)$ is a selector matrix. For example, let $e(2)$ be given by the first and third column of $\mathbb{I}_3$; that is

$$e(2) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

This implies that

$$S(\mu)e(2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 0 \end{pmatrix},$$

and moreover:

$$R(\mu; e(2)) \equiv [Z(\mu), S(\mu)e(2)] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{4 \times (2+1)}.$$

Thus, the matrix $R(\mu; e(2))$ is formed by collecting the gradient of the unique equality restriction and the first and third inequality restrictions in S. Note that regardless of the number of k columns selected from S—and regardless of what $\mathcal{M}$ is—the resulting matrix $R(\mu, e(k))$ will always have full column rank.

Verifying Assumption 2 with more general restrictions requires additional work. For example, suppose that the researcher is interested in including the restriction:

$$e_2' C_1(A) B_1 \geq 0.$$

This restriction says that the UMP shock cannot decrease the growth rate in Industrial Production even one-period after the shock. Since $C_1(A) = A_1$, the vector $e_2' C_1(A)$ is equal to the second row of A_1, which we can denote as $(A_{1,(2,1)}, A_{1,(2,2)}, A_{1,(2,3)}, A_{1,(2,4)})$. The matrix $S(\mu)$ is now given by:

$$S(\mu) = \begin{pmatrix} 1 & 0 & 0 & A_{1,(2,1)} \\ 0 & 1 & 0 & A_{1,(2,2)} \\ 0 & 0 & -1 & A_{1,(2,3)} \\ 0 & 0 & 0 & A_{1,(2,4)} \end{pmatrix}.$$

Hence, we conclude that Assumption 2 will be satisfied as long as $\mathcal{M}$ is such that $A_{1,(2,j)} \neq 0$ for all $j = 1, \ldots 4$, which means that each of the entries in the first lag of Y_{t-1} has predictive power on Y_t after controlling for the rest of the lags.

The third assumption is the following:

ASSUMPTION 3 The matrices $Z(\mu)$ and $S(\mu)$ are differentiable functions of the reduced-form parameter μ.

We are not aware of equality/inequality restrictions in the SVAR literature that does not satisfy this property. In particular, all the examples given in p. 5 of this paper satisfy Assumption 3.

4.1. *Lemma 1: Closed-form Solution for the maximum response given an active set of constraints*

In this section we show that that given a collection $r \in \mathbb{R}^{m \times n}$ of 'active' constraints ($m \leq n-1$) the maximum response is determined in closed-form (and up to sign) by the Karush-Kuhn-Tucker conditions of the program (2.5) and (2.6).

LEMMA 1 *Let r be a matrix of dimension $n \times m$, $m \leq (n-1)$ collecting the gradients of the 'active' (binding) constraints at a solution of the mathematical program (2.5). Suppose that Assumption 1 holds and suppose that $Z(\mu)$ and $S(\mu)$ are linearly independent at μ. Then, if $\overline{v}_{k,i,j}(\mu) \neq 0$:*

a) $\overline{v}_{k,i,j}(\mu)$ is given by either plus or minus the norm of the residual of the projection of $\Sigma^{1/2} C_k(A)' e_i$ into the space spanned by the columns of $\Sigma^{1/2} r$; that is:

$$(4.1) \qquad \overline{v}_{k,i,j}(\mu) = \left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2}r} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2},$$

or

$$(4.2) \qquad \overline{v}_{k,i,j}(\mu) = -\left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2}r} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2},$$

where

$$M_{\Sigma^{1/2}r} \equiv \mathbb{I}_n - \Sigma^{1/2} r (r' \Sigma r)^{-1} r' \Sigma^{1/2}.$$

b) In addition, there is a unique maximizer $x^*(\mu; r)$ such that $r' x^*(\mu; r) = \mathbf{0}_{m \times 1}$ and is given by:

$$x^*(\mu; r) = \Sigma^{1/2} \left(M_{\Sigma^{1/2}r} \right) \Sigma^{1/2} C_k(A)' e_i \Big/ \overline{v}_{k,i,j}(A, \Sigma).$$

Consequently, the sign of $\overline{v}_{k,i,j}(\mu)$ depends on which of the two solutions $x^*(\mu; r)$ (the one with (4.1) in the denominator or the one with (4.2)) satisfies the sign restrictions that are not in r.

PROOF: See Appendix A.1 for the proof, which uses the necessary Karush-Kuhn-Tucker conditions of the optimization problem to characterize the maximizers given a set r of active constraints.[10] Figure 1 presents a graphical representation of the mathematical program of interest. Figure 2 presents an intuitive description of the solution.

---

[10] To guarantee the existence of Karush-Kuhn-Tucker multipliers we use the fact that $Z(\mu)$ and $S(\mu)$ are linear independence at μ. Our assumption implies that the mathematical program defining the endpoints of the identified set satisfies a linear independence constraint qualification (see Fiacco and Ishizuka (1990), p. 224).
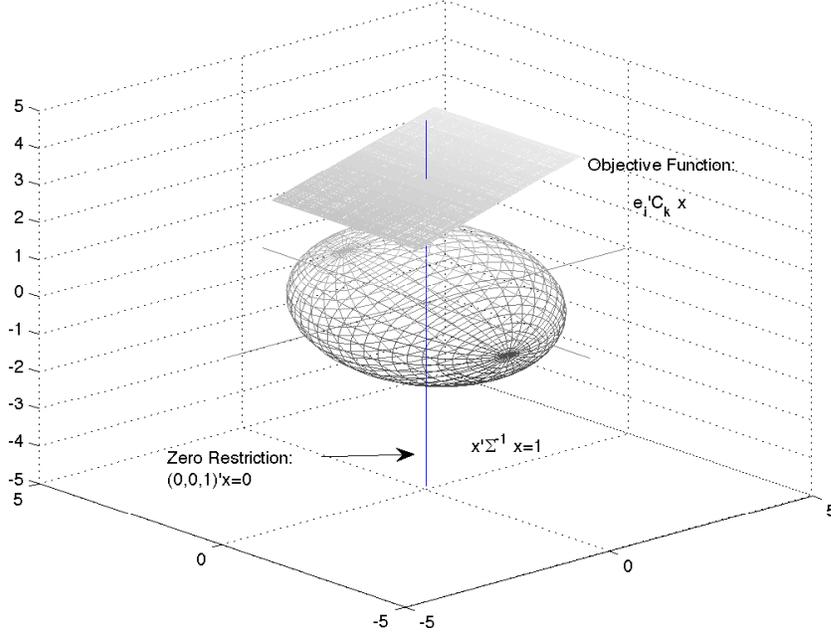
Figure 1: The mathematical program defining $\overline{v}_{k,i,j}(A, \Sigma)$ $(n = 3)$ with one zero restriction.

Figure 1 provides a graphical representation of the mathematical program (2.5), where $BB' = \Sigma$ has been replaced by the 'ellipsoid' constraint $x'\Sigma^{-1}x = 1$, $x \equiv B_j \in \mathbb{R}^3$ (this equivalence will not hold, in general, if there are restrictions on multiple shocks). The objective function corresponds to the hyperplane with normal vector $C_k(A)'e_i \in \mathbb{R}^3$. In this example, there is only one equality restriction with normal vector given by the (BLUE, SOLID) line. This restriction requires the contemporaneous impact of the j-th shock on the third variable to be zero. Note that without the equality restriction the maximizer and minimizer will be given by the point at which the hyperplane is tangent to the ellipsoid.

One way to think about the solution to the problem of interest is explained in Figure 2. Suppose there are only equality constraints. Note that $Z'B_j = \mathbf{0}_{m \times 1}$ implies that re-parameterized choice variable $\widetilde{x} \equiv \Sigma^{-1/2}B_j$ must lie on the orthogonal space of $\Sigma^{1/2}Z$. That is, the selected value of $\widetilde{x}$ should be of the form:

$$\widetilde{x} = M_{\Sigma^{1/2}Z}y, \quad M_{\Sigma^{1/2}Z} \equiv \left(\mathbb{I}_n - \Sigma^{1/2}Z(Z'\Sigma Z)^{-1}Z'\Sigma^{1/2} \right), \ y \in \mathbb{R}^n.$$

The quadratic equality constraint also restricts the choice variable $\widetilde{x}$ to satisfy $\widetilde{x}'\widetilde{x} = 1$. Consequently, the problem can be re-written as

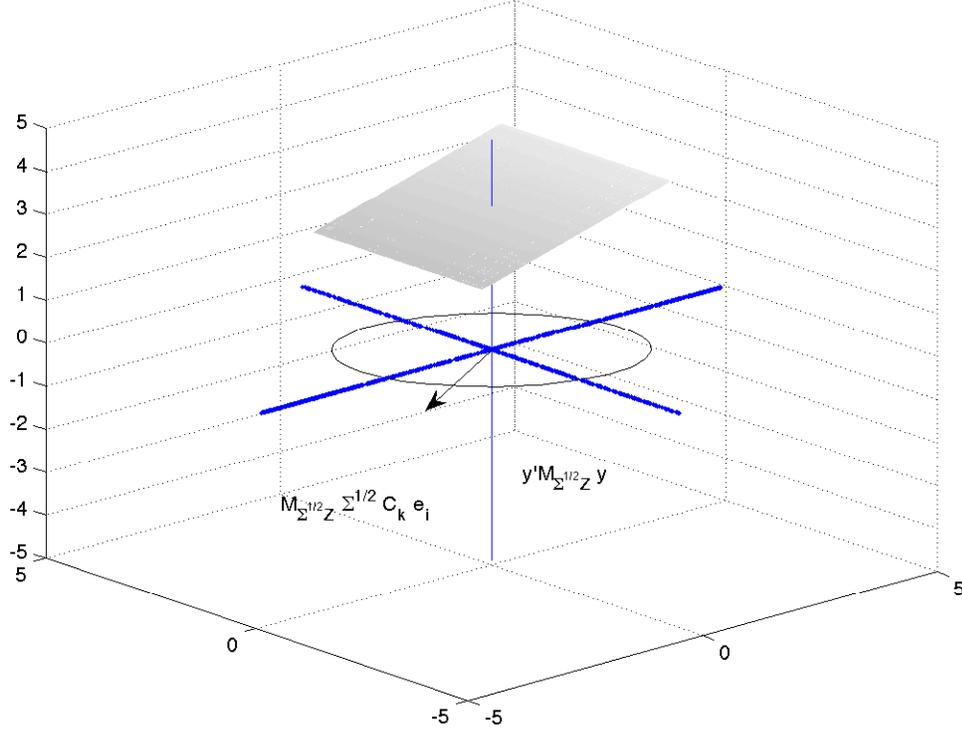$$\max_{y \in \mathbb{R}^n} e_i'C_k\Sigma^{1/2}M_{\Sigma^{1/2}Z}y \quad \text{s.t.} \quad y'M_{\Sigma^{1/2}Z}y = 1.$$

Figure 2: Solving for $\overline{v}_{k,i,j}(A,\Sigma)$ ($n=3$, $\Sigma=\mathbb{I}_3$) with one equality restriction.

Figure 2 provides a graphical representation of the solution to the mathematical program (2.5) when $\Sigma=\mathbb{I}_3$ and there is only one zero restriction. The solution to the program must lie in the orthogonal complement of Z (BLUE, THIN, SOLID). In this picture the orthogonal complement corresponds to the space spanned by the BLUE, THICK, SOLID lines. This implies that the rotated solution, denoted $\widetilde{x} \equiv \Sigma^{-1/2}x$, must be of the form $M_{\Sigma^{1/2}Z}y$ for some $y \in \mathbb{R}^3$. Hence, the only relevant part of $x'\Sigma^{-1}x=1$ becomes the projected version of it: $y'M_{\Sigma^{1/2}Z}y=1$, represented by the BLACK, SOLID ellipsoid. One can find the value of this problem by projecting the gradient of the objective function on the orthogonal complement of $\Sigma^{1/2}z$ (arrow) and selecting a direction in the ellipsoid proportional to it. The value function $\overline{v}_{k,ij}(A,\Sigma)$ will be given by the norm of the arrow.

An application of the Cauchy-Schwartz inequality shows that the positive value in (4.1) gives the maximum response in (2.5).[11]

---

[11]Using the fact that $M_{\Sigma^{1/2}Z}$ is idempotent and using the assumption that

$$\left(e_i'C_k\Sigma^{1/2}M_{\Sigma^{1/2}Z}\Sigma^{1/2}C_k'e_i\right)^{1/2} \neq 0,$$

the problem of interest becomes:

$$\max_{y\in\mathbb{R}^n} \left(e_i'C_k\Sigma^{1/2}M_{\Sigma^{1/2}Z}\Sigma^{1/2}C_k'e_i\right)^{1/2}\left[\frac{e_i'C_k\Sigma^{1/2}M_{\Sigma^{1/2}Z}}{\left(e_i'C_k\Sigma^{1/2}M_{\Sigma^{1/2}Z}\Sigma^{1/2}C_k'e_i\right)^{1/2}}\right]M_{\Sigma^{1/2}Z}y,$$

4.2. *Proposition 1: Algorithm to evaluate the maximum and minimum response*

We have provided a closed-form expression (up to a sign) for the maximum response $\overline{v}_{k,i,j}(\mu)$, given a collection r of active restrictions. We now answer the following question: how does one compute the maximum response $\overline{v}_{k,i,j}(\mu)$ for a given value of μ?

We use the result in Lemma 1 to state the solution of the mathematical program (2.5) that includes both equality and inequality restrictions. The main result in this section is that such problem can be solved by 'activating' different combinations of inequality constraints. In other words, the problem in (2.5) can be solved by finding the largest value among the Karush-Kuhn-Tucker points that satisfy a feasibility constraint.

ADDITIONAL NOTATION ILLUSTRATED WITH OUR EXAMPLE—1) COLLECTION OF ACTIVE CONSTRAINTS: Fix (A, Σ) and, in a slight abuse of notation let Z and S denote $Z(\mu)$ and $S(\mu)$. Define first:

$$R_0 \equiv Z,$$

as the $\mathbb{R}^n$ matrix that collects all of the m_z zero restrictions. Hence, in our empirical application:

$$R_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Define also:

$$R_0 \equiv \left\{ R \in \mathbb{R}^{n \times (m_z + 1)} \, \middle| \, R = [R_0, Se_i^{m_s}], \quad i \in \{1, \ldots m_s\} \right\},$$

as the collection of all matrices that *activate* one of the m_s inequality restrictions; analgously, R_1 corresponds to the collection of matrices that impose one of the inequality restrictions as an equality restriction. In our example:

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

Therefore,

---

s.t. $y'M_{\Sigma^{1/2}Z}y = 1$. By the Cauchy-Schwartz inequality this program is bounded above by $(e_i'C_k\Sigma^{1/2}M_{\Sigma^{1/2}Z}\Sigma^{1/2}C_k'e_i)^{1/2}$. This value can be achieved by $x^*(A, \Sigma; Z)$ in Lemma 1.

$$R_1 = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}.$$

More generally, for $l \leq n - m_z - 1$, consider the collection:

$$R_l \equiv \left\{ R \in \mathbb{R}^{n \times (m_z + l)} \;\middle|\; R = [R_0, Se_{m_1}^{m_s}, \ldots Se_{m_l}^{m_s}], \; \{m_i\}_{i=1}^l \text{ is a subsequence of } \{1, \ldots m_s\} \right\}.$$

The matrix $r_l \in R_l$ activates l of the m_s sign-restrictions in the SVAR model. Note that the collection R_l has $m_s!/(l!(m_s - l)!)$ elements and R_{m_s} has a unique element in which all the sign restrictions of the model are active (provided $m_s \leq n - m_z - 1$). In our example, $n - m_z - 1 = 2$. There are 3 different subsequences of two elements from the sequence $\{1, 2, 3\}$: $\{1, 2\}, \{1, 3\}$, and $\{2, 3\}$. Therefore,

$$R_2 = \left\{ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \right\}.$$

Thus, R_1 and R_2 denote the different collection of active constraints formed by choosing one and two of the elements of S, respectively.

ADDITIONAL NOTATION ILLUSTRATED WITH OUR EXAMPLE—2) FEASIBILITY: We define the feasibility of a vector $x \in \mathbb{R}^n$ (with respect to the sign restrictions) as the indicator function

$$(4.3) \qquad \mathbf{1}_{m_s}(x) = \mathbf{1}\left\{ S'x \geq \mathbf{0}_{m_s \times 1} \right\} \in \mathbb{R},$$

where, following convention, $\geq$ is taken component-wise whenever the binary relation is applied to vectors. Hence, $x \in \mathbb{R}^n$ is a feasible point for the mathematical program (2.5) if and only if $\mathbf{1}_{m_s}(x) = 1$ and x satisfies the equality restrictions in Z.[12] In the context of our example:

$$x = \begin{pmatrix} 4 \\ 3 \\ -2 \\ 0 \end{pmatrix} \implies \mathbf{1}_{m_s}(x) = 1, \text{ but } y = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 0 \end{pmatrix} \implies \mathbf{1}_{m_s}(y) = 0.$$

---

[12] We use the term primal feasibility in contrast with *dual* feasibility, which obtains whenever the Lagrange multipliers for the sign restrictions that are not active are all positive (or negative).

The point x is a feasible point for the mathematical program in (2.5) as it satisfies both the equality and inequality restrictions.

PROPOSITION 1 (**Algorithm to evaluate the maximum and minimum response**)
For $\mu \equiv (vec(A)', vec(\Sigma)')'$ consider the mathematical programs

$$\overline{v}_{k,i,j}(\mu) \equiv \max_{B_j} e_i' C_k(A) B_j, \quad \underline{v}_{k,i,j}(\mu) \equiv \min_{B_j} e_i' C_k(A) B_j,$$

with each program subject to

$$BB' = \Sigma,$$

with equality and inequality restrictions:

$$Z' B_j = \mathbf{0}_{m_z \times 1}, \quad S' B_j \geq \mathbf{0}_{m_s \times 1}.$$

Let

$$R \equiv \bigcup_{l=0}^{\min\{n-1-m_z, m_s\}} R_l,$$

denote all possible combinations of up to $n-1$ active constraints and for $r \in R$ define $v_{k,i,j}(\mu; r)$ as the function:

$$v_{k,i,j}(\mu; r) = \left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2} r} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2},$$

and let c be a positive and large constant (penalty term) such that $-c < \underline{v}_{k,i,j}(A, \Sigma) \leq \overline{v}_{k,i,j}(A, \Sigma) < c$ and such that $-c < -v_{k,i,j}(A, \Sigma, r) < v_{k,i,j}(A, \Sigma, r) < c$ for all $r \in R$.
Consider the candidate value functions:

$$v_{k,i,j}^+(\mu; r) = v_{k,i,j}(A, \Sigma; r), \quad v_{k,i,j}^-(\mu; r) = -v_{k,i,j}(A, \Sigma; r).$$

If $v_{k,i,j}(\mu; r) \neq 0$, set:

$$\begin{aligned}
f_{max}^+(\mu; r) &\equiv v_{k,i,j}^+(\mu; r) - 2(1 - \mathbf{1}_{m_s}(x_+^*(\mu; r)))c, \\
f_{max}^-(\mu; r) &\equiv v_{k,i,j}^-(\mu; r) - 2(1 - \mathbf{1}_{m_s}(x_-^*(\mu; r)))c, \\
f_{min}^+(\mu; r) &\equiv v_{k,i,j}^+(\mu; r) + 2(1 - \mathbf{1}_{m_s}(x_+^*(\mu; r)))c, \\
f_{min}^-(\mu; r) &\equiv v_{k,i,j}^-(\mu; r) + 2(1 - \mathbf{1}_{m_s}(x_-^*(\mu; r)))c,
\end{aligned}$$

where

$$x_+^*(\mu; r) = \Sigma^{1/2} \left(M_{\Sigma^{1/2} r} \right) \Sigma^{1/2} C_k(A)' e_i \Big/ v_{k,i,j}(\mu),$$

$$x^*_-(\mu; r) = -\Sigma^{1/2}\left(M_{\Sigma^{1/2}r}\right)\Sigma^{1/2}C_k(A)'e_i\Big/v_{k,i,j}(\mu).$$

If $v_{k,i,j}(\mu; r) = 0$ and there is a point $x^* \neq 0$ satisfying the equality restrictions in r and also the inequality restrictions that are not included in r, set:

$$f^+_{max}(\mu; r) = f^-_{max}(\mu; r) = f^+_{min}(\mu; r) = f^-_{min}(\mu; r) = 0.$$

If $v_{k,i,j}(\mu; r) = 0$ and there is no point $x^* \neq 0$ satisfying the equality restrictions in r and the inequality restrictions that are not in r, set:

$$f^+_{max}(\mu; r) = f^-_{max}(\mu; r) = -c \quad \text{and} \quad f^+_{min}(\mu; r) = f^-_{min}(\mu; r) = c.$$

Then:

$$\overline{v}_{k,i,j}(\mu) = \max_{r \in R}\left(\max\{f^+_{max}(\mu; r), f^-_{max}(\mu; r)\}\right),$$

and

$$\underline{v}_{k,i,j}(\mu) = \min_{r \in R}\left(\min\{f^+_{min}(\mu; r), f^-_{min}(\mu; r)\}\right).$$

That is, the value function $\overline{v}_{k,i,j}(\mu)$ is obtained by computing the Karush-Kuhn-Tucker points in Lemma 1 for each r, penalizing the value $\overline{v}_{k,i,j}(\mu; r)$ if unfeasible, and maximizing over all the possible values of r. The minimum value is obtained analogously.

PROOF: The intuition behind the proof is as follows. Note that any combination of active sign restrictions r for which $x^*_+(A, \Sigma; r)$ (or $x^*_-(A, \Sigma; r)$) is well-defined and feasible must be, by definition, no larger than $\overline{v}_{k,i,j}(A, \Sigma)$. Thus, we only have to show that $\max_{r \in R}\left(\max\{f^+_{max}(A, \Sigma; r), f^-_{max}(A, \Sigma; r)\}\right) \geq \overline{v}_{k,i,j}(A, \Sigma)$. Since Lemma 1 showed that the value of the program (2.5) should be of the form $f^+_{max}(A, \Sigma, r)$ or $f^-_{max}(A, \Sigma, r)$ for some $r \in R$, the result must follow. The proof is formalized in Appendix A.2.

ALGORITHM TO EVALUATE THE ENDPOINTS OF THE IDENTIFIED-SET: The proposition above shows that in order to solve the mathematical problem in (2.5) it is sufficient to apply the following algorithm:

1. *Activate* different combinations of the m_s sign restrictions. Collect the original m_z equality restrictions and the inequality restrictions that were activated in the matrix r. The matrix should have no more than $n - 1$ columns. Note that the total number of matrices r will be given by:

$$\sum_{k=0}^{\min\{m_s, n-1-m_z\}} \binom{m_s}{k}.$$

2. Compute the candidate value functions $\pm v_{k,i,j}(A, \Sigma; r)$ for each of the elements $r \in R$.

3. If $v_{k,i,j}(A, \Sigma; r) \neq 0$, verify if $x_+^*(A, \Sigma; r)$ satisfies the sign restrictions that were not included in r. That is, verify the *feasibility* of the solution $x_+^*(A, \Sigma; r)$. If the primal feasibility condition is satisfied set

$$f_{max}^+(A, \Sigma; r) = f_{min}^+(A, \Sigma; r) = v_{k,i,j}(A, \Sigma; r).$$

If the primal feasibility condition is violated penalize $v_{k,i,j}(A, \Sigma; r)$ to guarantee that it is never a solution by setting:

$$f_{max}^+(A, \Sigma; r) \equiv v_{k,i,j}(A, \Sigma; r) - 2c,$$

and

$$f_{min}^+(A, \Sigma; r) \equiv v_{k,i,j}(A, \Sigma; r) + 2c.$$

Check the feasibility of $x_-^*(A, \Sigma, r)$ and proceed in the same way. If $v_{k,i,j}(A, \Sigma; r) = 0$ do the adjustment described in Proposition 1.

4. Select the maximum value of $\max\{f_{max}^+(A, \Sigma; r), f_{max}^-(A, \Sigma; r)\}$ over $r \in R$; that is, consider the different combinations of active restrictions and select the maximum value $\pm v_{k,i,j}(A, \Sigma, r)$ over them. This gives $\overline{v}_{k,i,j}(A, \Sigma)$. Taking the minimizer over $f_{min}^+(A, \Sigma; r), f_{min}^-(A, \Sigma; r)$ gives the smallest value.

USING THE ALGORITHM IN THE UMP EXAMPLE: We use the algorithm to evaluate the identified set in the running example. We fix μ at its estimated OLS values, denoted $\widehat{\mu}_T$, and we report $\overline{v}_{k,i,j}(\widehat{\mu}_T)$ and $\underline{v}_{k,i,j}(\widehat{\mu}_T)$ for the cumulative IRFs.[13] The scale in Figure 3 corresponds to a one standard deviation structural UMP shock.

We consider first the equality/inequality restrictions in Table I. We note that evaluating the endpoints of the identified for the 4 variables in the VAR, over 40 horizons, takes around .1 seconds. We then include an additional inequality restriction on the response of output to an expansionary UMP shock. Namely, we assume that even one period after the shock, the cumulative effect on IP cannot be negative ($e_2'(C_0 + C_1(A))B_1 \geq 0$). A comparison between the two collections of restrictions suggests that, in this example, the noncontemporaneous constraint has almost no additional identification power.

Of course, one could use the Bayesian algorithm in Uhlig (2005) to approximate the value of the endpoints. Given D draws of the reduced-form parameters (A, Σ) and a unit vector $q \in \mathbb{R}^n$, one could report the maximum and minimum value for $\{\lambda_{k,i,j}^d(A, \Sigma, q)\}_{d=1}^D$ over the different draws. This algorithm is a random grid search approach to solve the programs (2.5) and (2.6). Figure 6 in the appendix presents a comparison between the different approaches. The *grid search* takes around 300 seconds to run and underestimates the identified set.

---

[13]The formula for the maximum (minimum) k-th period ahead *cumulative* IRF replaces $C_k(\widehat{A}_T)$ by $C_0(\widehat{A}_T) + C_1(\widehat{A}_T)+, \ldots, +C_k(\widehat{A}_T)$.

GAFAROV, MEIER, AND MONTIEL-OLEA

Figure 3: Identified Set for the Cumulative Impulse Response Functions to a one standard deviation UMP shock (given $\widehat{\mu}_T$)
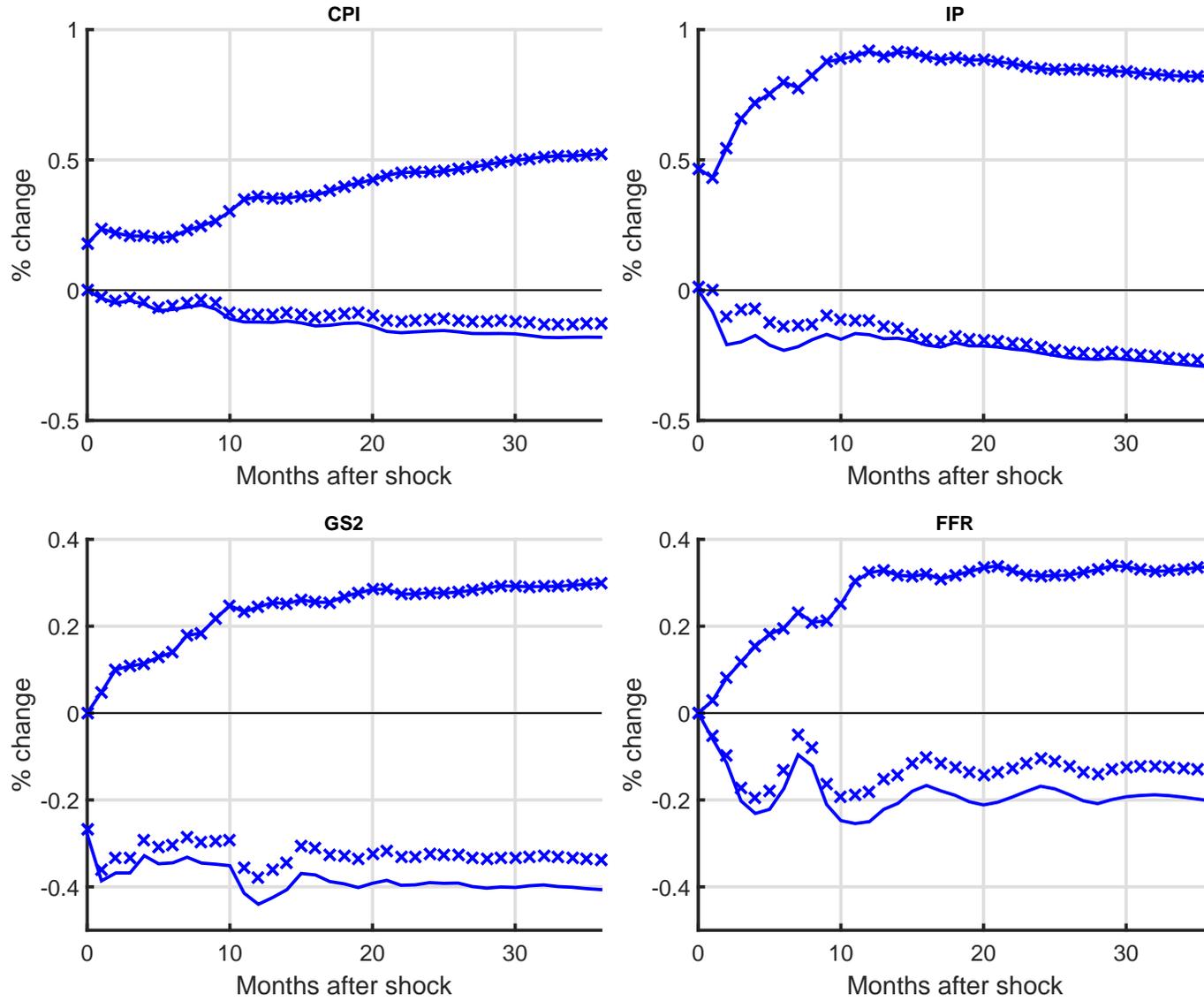
(SOLID, BLUE LINE) Endpoints of the identified set for the cumulative responses given $\widehat{\mu}_T$ and the equality/inequality restrictions in Table I. (BLUE, CROSSES) Endpoints of the identified set with the additional restriction $e_2'(C_0 + C_1(A))B_1 \geq 0$.

5. DIRECTIONAL DIFFERENTIABILITY OF THE ENDPOINTS

Once again, let

$$R(\mu) \equiv \bigcup_{l=0}^{\min\{n-1-m_z, m_s\}} R_l(\mu),$$

denote the set of all possible combinations up to $n-1$ active constraints.[14] Let us denote the typical element in $R(\mu)$ as $r(\mu)$, which we take to be an $n \times l$ matrix with $l \leq n-1$. We will continue working with the auxiliary function:

$$v_{k,i,j}(\mu; r(\mu)) = \left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2} r(\mu)} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2},$$

where we now explicitly acknowledge the possible dependence of r on μ. In Lemma 1 we have shown that if $r(\mu)$ is the active set of constraints at a solution of the program (2.5), then:

$$\overline{v}_{k,i,j}(\mu) = v_{k,i,j}(\mu; r(\mu)) \quad \text{or} \quad \overline{v}_{k,i,j}(\mu) = -v_{k,i,j}(\mu; r(\mu));$$

as long as $\overline{v}_{k,i,j}(\mu) \neq 0$. In order to establish the differentiability of $\overline{v}_{k,i,j}(\mu)$ we prove the following intermediate result.

LEMMA 2 (**Differentiability results for a given active set of constraints**) *If $r(\mu)$ is differentiable with respect to μ and $v_{k,i,j}(\mu; r(\mu)) \neq 0$, then $v_{k,i,j}(\mu; r(\mu))$ is differentiable with respect to μ with derivative $\dot{v}_{k,i,j}(\mu; r(\mu))$ given by:*

$$\begin{bmatrix} \frac{\partial v_{k,i,j}(\mu; r(\mu))}{\partial vec(A)} \\[2mm] \frac{\partial v_{k,i,j}(\mu; r(\mu))}{\partial vec(\Sigma)} \end{bmatrix} = \begin{bmatrix} \frac{\partial vec(C_k(A))}{\partial vec(A)}(x^*(\mu; r(\mu)) \otimes e_i) - \sum_{k=1}^{l} w_k^* \frac{\partial vec(r_k(\mu))}{vec(A)} x^*(\mu, r(\mu)) \\[2mm] \lambda^*(\Sigma^{-1} x^*(\mu, r(\mu)) \otimes \Sigma^{-1} x^*(\mu, r(\mu))) - \sum_{k=1}^{l} w_k^* \frac{\partial vec(r_k(\mu))}{vec(\Sigma)} x^*(\mu, r(\mu)) \end{bmatrix},$$

where $r_k(\mu)$ denotes the k-th column of $r(\mu)$,

$$x^*(A, \Sigma; r(\mu)) = \Sigma^{1/2} \left(M_{\Sigma^{1/2} r(\mu)} \right) \Sigma^{1/2} C_k(A)' e_i \Big/ v_{k,i,j}(A, \Sigma; r(\mu)),$$

$$\lambda^* \equiv \frac{1}{2} v_{k,i,j}(A, \Sigma; r(\mu)), \quad w^* \equiv [r(\mu)' \Sigma r(\mu)]^{-1} r(\mu)' \Sigma C_k(A) e_i,$$

and w_k^* is the k-th component of the vector w^*.

PROOF: See Appendix A.3.

---

[14]The dependence of the set R on the parameter μ was omitted in the previous section for notational simplicity.

The *envelope theorem* sheds light on the derivative formula provided in Lemma 2. Note first that

$$v_{k,i,j}(\mu, r(\mu)) = \max_{x \in \mathbb{R}^n} e_i' C_k(A) x \quad \text{s.t.} \quad x' \Sigma^{-1} x = 1 \text{ and } r'(\mu) x = \mathbf{0}_{l \times 1}.$$

The auxiliary *Lagrangian function* of this problem is given by:

$$\mathcal{L}(x; \mu, r(\mu)) = (x' \otimes e_i') \text{vec}(C_k(A)) - \lambda \Big((x' \otimes x') \text{vec}(\Sigma^{-1}) - 1 \Big) - w'(r(\mu)'x),$$

where λ is the Lagrange multiplier corresponding to the quadratic equality restriction and $w \in \mathbb{R}^l$ is the vector of Lagrange multipliers corresponding to the l equality restrictions. The envelope theorem suggests that $\dot{v}_{k,i,j}(\mu; r(\mu))$ is given by the formula in Lemma 2. We confirm this intutition in the prove of Lemma 2; provided $v_{k,i,j}(\mu; r(\mu)) \neq 0$.

We now establish the differentiability of $\overline{v}_{k,i,j}(\mu)$. Without loss of generality, assume that $\overline{v}_{k,i,j}(\mu) > 0$ (and also that $\underline{v}_{k,i,j}(\mu) < 0$). For a fixed vector of reduced-form parameters define the sets:

$$R^*(\mu) \equiv \Big\{ r(\mu) \in R(\mu) \mid \overline{v}_{k,i,j}(\mu) = v_{k,i,j}(\mu, r(\mu)) \Big\},$$

and

$$R_*(\mu) \equiv \Big\{ r(\mu) \in R(\mu) \mid \underline{v}_{k,i,j}(\mu) = -v_{k,i,j}(\mu, r(\mu)) \Big\}.$$

The set $R^*(\mu)$ collects the different active constraints that could lead to the maximum value. The set $R^*(\mu)$ is a singleton if and only if the program (2.5) has a unique solution. The set $R_*(\mu)$ is defined analogously.

PROPOSITION 2 (Directional differentiability of the endpoints of the identified set) *Suppose, w.l.o.g., that $\overline{v}_{k,i,j}(\mu) > 0$ and $\underline{v}_{k,i,j}(\mu) < 0$. Then, for any sequence $h_n \in \mathbb{R}^d$ such that $h_n \to h$ and any sequence $t_n \to \infty$:*

$$t_n \Big(\overline{v}_{k,i,j}(\mu + h_n/t_n) - \overline{v}_{k,i,j}(\mu) \Big) \to \max_{r(\mu) \in R^*(\mu)} \Big[\dot{v}_{k,i,j}(\mu; r(\mu))' h \Big],$$

and

$$t_n \Big(\underline{v}_{k,i,j}(\mu + h_n/t_n) - \underline{v}_{k,i,j}(\mu) \Big) \to \min_{r(\mu) \in R_*(\mu)} \Big[-\dot{v}_{k,i,j}(\mu; r(\mu))' h \Big].$$

Thus, $\overline{v}_{k,i,j}$ and $\underline{v}_{k,i,j}$ are directionally differentiable functions of the reduced-form parameters with directional derivative:

$$\max_{r(\mu) \in R^*(\mu)} \Big[\dot{v}_{k,i,j}(\mu; r(\mu))' h \Big],$$

and

$$\min_{r(\mu)\in R_*(\mu)}\left[-\dot{v}_{k,i,j}(\mu;r(\mu))'h\right],$$

respectively.

Whenever $R^*(\mu)$ is a singleton—i.e., $R^*(\mu) = \{r(\mu)\}$—the value function $\overline{v}_{k,i,j}(\mu)$ is fully differentiable with derivative $\dot{v}_{k,i,j}(\mu, r(\mu))$. Likewise if $R_*(\mu)$ is a singleton, the value function $\underline{v}_{k,i,j}(\mu)$ is differentiable with derivative $-\dot{v}_{k,i,j}(\mu, r(\mu))$.

PROOF: See Appendix A.4.

Theorem 4.2, p. 223 in Fiacco and Ishizuka (1990) and Theorem 4.24, p. 280 in the book of Bonnans and Shapiro (2000) present a generalized version of the envelope theorem. They show that—under suitable constraint qualifications—the directional derivative (in direction h and evaluated at parameter μ) of the largest and smallest value in a mathematical program with equality and inequality constraints is given by:

$$\sup_{r\in R^*(\mu)}\left[\nabla_\mu \mathcal{L}(x^*(\mu;r);\mu)h\right],$$

and

$$\inf_{r\in R_*(\mu)}\left[\nabla_\mu \mathcal{L}(x^*(\mu;r);\mu)h\right],$$

provided there is a unique set of Lagrange Multipliers supporting the optimal solutions $x^*(\mu;r)$. Proposition 2 uses the results in Lemma 1 and Lemma 2 to verify this formula.

DELTA-METHOD VS. BOOTSTRAP: We also note that directionally differentiable functions have been a topic of recent research. Fang and Santos (2014) show that the standard bootstrap is not consistent when applied to parameters of the form $v(\mu)$, where v is a directionally differentiable function. Kitagawa, Payne, and Montiel Olea (2016) show that Bayesian credible sets based on the quantiles of the posterior distribution of $v(\mu)$ will be asymptotically equivalent to the frequentist bootstrap (which is not consistent in this case).

These results imply that typical frequentist and Bayesian inference for directionally differentiable functions is problematic. The next section shows that the special form of the directional derivative in the class of SVARs models studied in this paper allows the researcher to conduct delta-method inference, with a slight adjustment on the standard errors.

6. DELTA-METHOD INFERENCE

This section proposes a delta-method confidence interval of the form

$$\text{CS}_T(1-\alpha) \equiv \left[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\,\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}, \; \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\,\widehat{\sigma}_{(k,i,j),T}/\sqrt{T} \right],$$

where

$$\widehat{\mu}_T \equiv (\text{vec}(\widehat{A}_T)', \text{vec}(\widehat{\Sigma}_T)'),$$

is the OLS estimator for μ defined as:

$$\widehat{A}_T \;\; \equiv \;\; \left(\frac{1}{T}\sum_{t=1}^{T} Y_t X_t'\right)\left(\frac{1}{T}\sum_{t=1}^{T} X_t X_t'\right)^{-1}, \quad \widehat{\Sigma}_T \equiv \frac{1}{T}\sum_{t=1}^{T}\widehat{\eta}_t\widehat{\eta}_t',$$

with

$$X_t \equiv (Y_{t-1}', \dots, Y_{t-p}')', \quad \widehat{\eta}_t \equiv Y_t - \widehat{A}_T X_t.$$

We work under the assumption that $\sqrt{T}(\widehat{\mu}_T - \mu)$ is asymptotically normal with some co-variance matrix Ω. A common formula to estimate the asymptotic variance of $\widehat{\mu}_T$ is:

$$\widehat{\Omega}_T \equiv \left(\frac{1}{T}\sum_{t=1}^{T}\text{vec}\left([\widehat{\eta}_t X_t', \widehat{\eta}_t\widehat{\eta}_t' - \widehat{\Sigma}_T]\right)\text{vec}\left([\widehat{\eta}_t X_t', \widehat{\eta}_t\widehat{\eta}_t' - \widehat{\Sigma}_T]\right)'.$$

We use the results in Proposition 2 and the asymptotically normality of $\widehat{\mu}_T$ to suggest the following formula for $\widehat{\sigma}_{(k,i,j),T}$:

$$(6.1) \qquad \widehat{\sigma}_{(k,i,j),T} \equiv \max_{r \in R(\widehat{\mu}_T)} \left[\dot{v}_{k,i,j}(\widehat{\mu}_T; r)'\widehat{\Omega}_T \dot{v}_{k,i,j}(\widehat{\mu}_T; r) \right],$$

where $R(\widehat{\mu}_T)$ is the set of all possible active constraints evaluated at $\widehat{\mu}_T$.

MAIN RESULT IN THIS SECTION: Let P denote the data generating process and let $\mathcal{I}_{k,i,j}^{\mathcal{R}}(\mu(P))$ denote the identified set for the structural parameter $\lambda_{k,i,j}$ given the equality/inequality restrictions in $\mathcal{R}(\mu)$. This section shows that under our proposed specification of $\widehat{\sigma}_{(k,i,j),T}$:

$$\liminf_{T\to\infty} \inf_{\lambda \in \mathcal{I}^R(\mu(P))} P\left(\lambda \in \text{CS}_T(1-\alpha)\right) \geq 1-\alpha.$$

Consequently, the delta-method confidence interval presented in this paper is *pointwise consistent in level.* We now describe the main large-sample assumptions concerning P.

DATA GENERATING PROCESS: The SVAR parameters $(A_1, \ldots, A_p, B, F)$ define a probability measure, denoted P, over the data observed by the econometrician. Our main assumption concerning P is as follows:

ASSUMPTION 4 (**Asymptotic Normality of** $\widehat{\mu}_T$) The data generating process P is such that for $\mu(P) \in \mathbb{R}^d$:

$$\sqrt{T}(\widehat{\mu}_T - \mu(P)) \xrightarrow{d} \zeta(P) \sim \mathcal{N}_d\Big(\mathbf{0}, \, \Omega(P)\Big),$$

and

$$\widehat{\Omega}_T \xrightarrow{p} \Omega(P).$$

Thus, our only restriction on $(A_1, A_2, \ldots, A_p, B, F)$ is that, whatever these parameters are, the OLS estimator $\widehat{\mu}_T$ is asymptotically normal with a covariance matrix that can be estimated consistently.

DELTA-METHOD FOR DIRECTIONALLY DIFFERENTIABLE FUNCTIONS: Dümbgen (1993), Shapiro (1991), and Fang and Santos (2014) have shown if v is a directionally differentiable function with directional derivative $\dot{v}_\mu(h)$ (in direction h evaluated at μ) then:

$$\sqrt{T}(v(\widehat{\mu}) - v(\mu)) \xrightarrow{d} \dot{v}_\mu(\zeta),$$

whenever Assumption 4 holds.[15] Proposition 2 in the previous section established that the directional derivative of $\overline{v}_{k,i,j}$—in direction h evaluated at μ—is given by:

$$\max_{r \in R^*(\mu)} \Big[\dot{v}_{k,i,j}(\mu; r))'h \Big].$$

where $R^*(\mu)$ collects the active constraints that generate $\overline{v}_{k,i,j}(\mu)$. Thus, Proposition 2 and Assumption 4 imply that:

$$\sqrt{T}(\overline{v}_{k,i,j}(\widehat{\mu}) - \overline{v}_{k,i,j}(\mu)) \xrightarrow{d} \max_{r \in R^*(\mu)} \Big[\dot{v}_{k,i,j}(\mu, r)'\zeta \Big],$$

where

$$\dot{v}_{k,i,j}(\mu, r(\mu))'\zeta \sim \mathcal{N}_1\Big(0, \dot{v}_{k,i,j}(\mu, r(\mu))'\Omega\dot{v}_{k,i,j}(\mu, r(\mu))\Big).$$

INFERENCE FOR DIRECTIONALLY DIFFERENTIABLE FUNCTIONS: How can we use the delta-method result above to construct a confidence set for $\lambda_{k,i,j}$? Our suggestion—which

---

[15]The map $v : \mathbb{R}^r \to \mathbb{R}$ is said to be Hadamard directionally differentiable at $\mu \in \mu \subseteq \mathbb{R}^r$, tangentially to $\mathbb{R}^r$, if there is a continuous (not necessarily linear) map $\dot{v}(\cdot, \mu) : \mathbb{R}^r \to \mathbb{R}$ such that:

$$\lim_{T \to \infty} \left| \frac{v(\mu + t_T h_T) - v(\mu)}{t_T} - \dot{v}(h; \mu) \right| = 0,$$

for all sequences $\{h_T\} \subseteq \mathbb{R}^r$ and $\{t_T\} \subseteq \mathbb{R}_+$ such that $t_T \to 0^+$, $h_T \to h \in \mathbb{R}^r$ and $\mu + t_T h_T \in \mu$ for all T. The function $v(\cdot)$ is Fully Differentiable at μ if and only if the mapping $\dot{v}(\cdot; \mu)$ is linear. See Fang and Santos (2014) for a recent elegant exposition on directionally differentiable functions. See also Shapiro (1990).

exploits the specific form of the directional derivative in the SVAR context—is to consider:

$$\widehat{\sigma}_{(k,i,j),T} \equiv \max_{r \in R(\widehat{\mu}_T)} \left[\dot{v}_{k,i,j}(\widehat{\mu}_T; r)' \widehat{\Omega}_T \dot{v}_{k,i,j}(\widehat{\mu}_T; r) \right],$$

where $R(\widehat{\mu}_T)$ is the set of all the different collections of active constraints evaluated at $\widehat{\mu}_T$. The resulting standard error will then be used to enlarge the plug-in estimator of the endpoints of the identified set. The suggested confidence interval is shown to be pointwise consistent in level.[16] This is formalized in the following proposition.

PROPOSITION 3 (**Pointwise Consistency in Level of the Delta-Method Confidence Interval**) *Let $\widehat{\sigma}_{(k,i,j),T}$ be defined as in (6.1). Suppose there are at most $n-1$ equality restrictions and that data generating process P is such that $Z(\mu(P))$ and $S(\mu(P))$ are linearly independent at $\mu(P)$ as defined in Section 4, and also differentiable. Suppose in addition that:*

$$\min_{r \in R(\mu)} \left[\dot{v}_{k,i,j}(\mu(P), r)' \Omega(P) \dot{v}_{k,i,j}(\mu(P), r) \right] > 0.$$

Then, under Assumption 4:

$$\liminf_{T \to \infty} \inf_{\lambda \in \mathcal{I}^{\mathcal{R}}_{k,i,j}(\mu(P))} P\Big(\lambda \in \Big[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}, \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}\Big]\Big)$$

is at least $1 - \alpha$.

PROOF: See Appendix A.5.

The intuition behind our proof is as follows. Note first that if λ belongs to the identified set, (i.e., $\lambda \in \mathcal{I}^{\mathcal{R}}_{k,i,j}(\mu(P))$), then such parameter must lie between the maximum and the minimum response; that is $\lambda \in [\underline{v}_{k,i,j}(\mu(P)), \overline{v}_{k,i,j}(\mu(P))]$. Consequently, one can show that:

$$P\Big(\lambda \in \Big[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\,\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}, \ \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\,\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}\Big]\Big),$$

is larger than or equal to

$$P\Big(\underline{v}_{k,i,j}(\mu(P)), \overline{v}_{k,i,j}(\mu(P)) \in \Big[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}, \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\widehat{\sigma}_{(k,i,j),T}/\sqrt{T}\Big]\Big).$$

Thus, a sufficient condition for the validity of the delta-method confidence interval is that it covers the identified set with probability at least $1 - \alpha$. Note that the probability of covering

---

[16]The question of how to build a *uniformly consistent in level*, delta-method confidence set for a set-identified parameter is beyond the scope of this paper. For the readers interested in uniform inference for set-identified parameters in SVARs our suggestion is to apply the projection approach developed in Gafarov et al. (2016). Compared to GMM16, the delta-method approach described in this paper is faster to implement.

the identified set can be written as one minus the sum of the following two terms:

$$P\Big(\sqrt{T}(\underline{v}_{k,i,j}(\widehat{\mu}_T) - \underline{v}_{k,i,j}(\mu(P))) > z_{1-\alpha/2}\widehat{\sigma}_{(k,i,j),T}\Big),$$

and

$$P\Big(\sqrt{T}(\overline{v}_{k,i,j}(\widehat{\mu}_T) - \overline{v}_{k,i,j}(\mu(P)) < -z_{1-\alpha/2}\,\widehat{\sigma}_{(k,i,j),T} >)\Big).$$

Using our large sample assumptions and the delta-method for directional differentiable functions, these probabilities are approximately equal to:

$$P\Big(\min_{r \in R_*(\mu(P))}[-\dot{v}_{k,i,j}(\mu(P),r)'Z_d] > z_{1-\alpha/2}\sigma_{(k,i,j)}\Big),$$

and

$$P\Big(\max_{r \in R^*(\mu(P))}[\dot{v}_{k,i,j}(\mu(P),r)'Z_d] < -z_{1-\alpha/2}\sigma_{(k,i,j)}\Big),$$

where:

$$\sigma_{k,i,j} \equiv \max_{r \in R(\mu)}\Big[\dot{v}_{k,i,j}(\mu(P),r)'\Omega(P)\dot{v}_{k,i,j}(\mu(P),r)\Big], \quad Z_d \sim \mathcal{N}(0,\Omega).$$

Take any $r_* \in R_*(\mu(P))$ for which $\sigma_{(k,i,j)}(r_*) \equiv \dot{v}_{k,i,j}(\mu(P),r_*)'\Omega\dot{v}_{k,i,j}(\mu(P),r_*) > 0$. (we have assumed that such r_* exists). It follows that:

$$P\Big(\min_{r \in R_*(\mu(P))}[-\dot{v}_{k,i,j}(\mu(P),r)'Z] > z_{1-\alpha/2}\sigma_{(k,i,j)}\Big) \leq P\Big(\sigma_{k,i,j}(r_*)N(0,1) > z_{1-\alpha/2}\sigma_{(k,i,j)}\Big),$$

and the last term is bounded above by $\alpha/2$. Analogously, we can select an $r^* \in R^*(\mu(P))$ for which $\sigma_{(k,i,j)}(r^*) \equiv \dot{v}_{k,i,j}(\mu(P),r^*)'\Omega\dot{v}_{k,i,j}(\mu(P),r^*) > 0$ and we can show that:

$$P\Big(\max_{r \in R^*(\mu(P))}[\dot{v}_{k,i,j}(\mu(P),r)'Z] < -z_{1-\alpha/2}\sigma_{(k,i,j)}\Big) \leq P\Big(\sigma_{k,i,j}(r^*)N(0,1) < -z_{1-\alpha/2}\sigma_{(k,i,j)}\Big),$$

which is also bounded above by $\alpha/2$. These inequalities suffice to establish the pointwise validity of our delta-method approach.

MONTE-CARLO EVIDENCE: We conduct a simple Monte-Carlo exercise to study the coverage probability of our delta-method confidence set. We set $(1 - \alpha) = .68$, which implies that $z_{1-\alpha/2} = .9945$. We generate $10,000$ draws from the multivariate normal model $N_d(\widehat{\mu}_T, \widehat{\Omega}_T/T)$ and for each draw (denoted μ^*) we compute the confidence interval:

$$\Big[\underline{v}_{k,i,j}(\mu^*) - .9945\,\sigma^*_{(k,i,j),T}/\sqrt{T}, \ \overline{v}_{k,i,j}(\mu^*) + .9945\,\sigma^*_{(k,i,j),T}/\sqrt{T}\Big].$$

We check whether $[\underline{v}_{k,i,j}(\widehat{\mu}_T), \overline{v}_{k,i,j}(\widehat{\mu}_T)]$ is contained in the confidence interval or not. The estimated probability provides a lower bound on the coverage of the identified parameter. The results are reported in Figure 4.

Figure 4: Monte-Carlo Coverage Probability based on the model $\mu^* \sim \mathcal{N}(\widehat{\mu}_T, \widehat{\Omega}_T/T)$, $T = 342$.
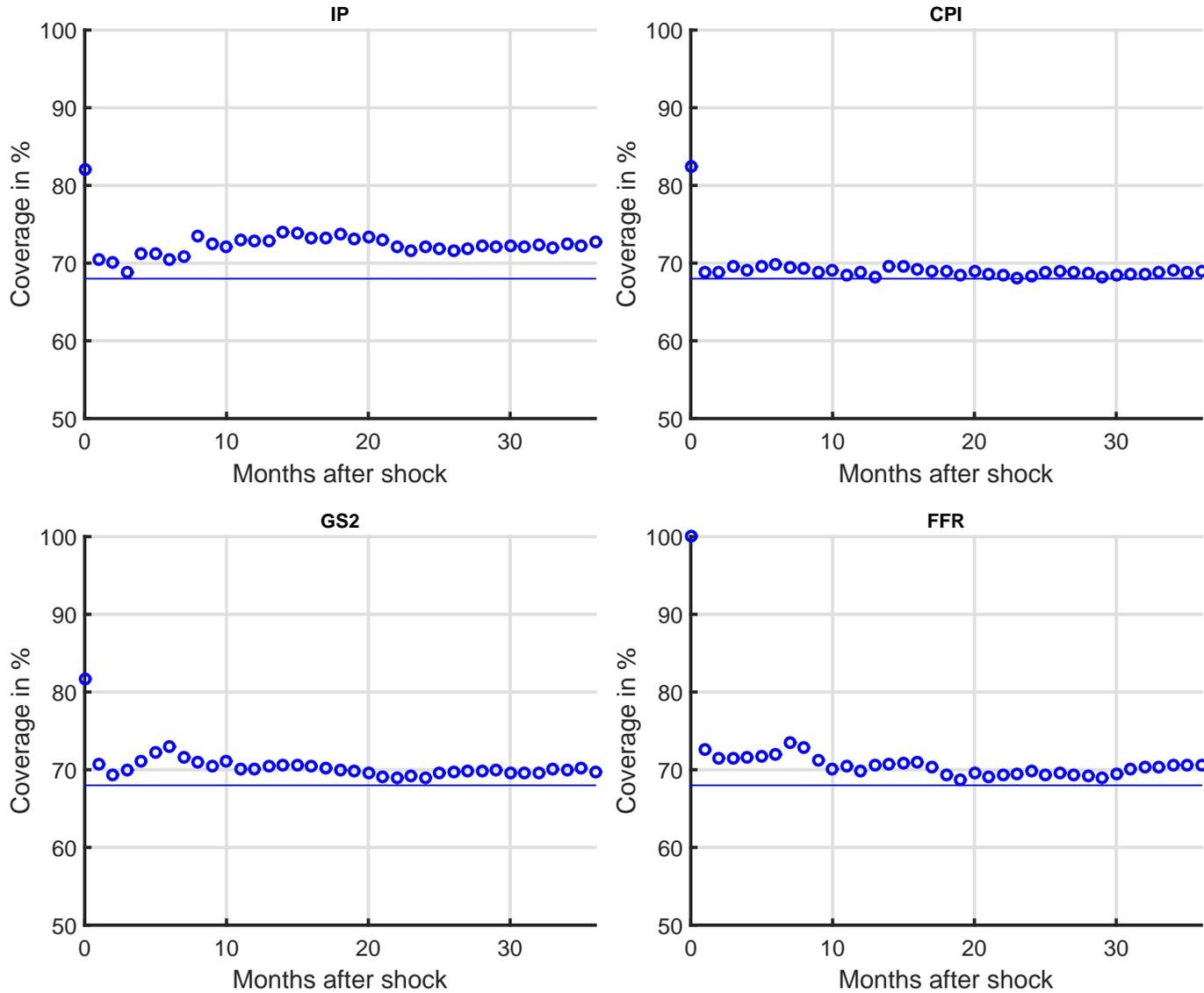
(BLUE, CIRCLES) Monte-Carlo estimate of the probability $P\left([\underline{v}_{k,i,j}(\widehat{\mu}_T), \overline{v}_{k,i,j}(\widehat{\mu}_T)] \in [\underline{v}_{k,i,j}(\mu^*) - .9945\,\sigma^*_{(k,i,j),T}/\sqrt{T}, \ \overline{v}_{k,i,j}(\mu^*) + .9945\,\sigma^*_{(k,i,j),T}/\sqrt{T}]\right)$

for the model $\mu^* \sim \mathcal{N}(\widehat{\mu}_T, \widehat{\Omega}_T)$, with $T = 342$. The values $\widehat{\mu}_T$ and $\widehat{\Omega}_T$ correspond, respectively, to the estimators of the reduced-form parameter and its asymptotic covariance matrix in the UMP application. (BLUE, SOLID LINE) Nominal confidence level for the delta-method confidence interval (68%).

7. UNCONVENTIONAL MONETARY SHOCKS

As we mentioned before, the identification strategy in this paper was motivated by two mechanisms used by the Federal Reserve to affect market beliefs during the Great Recession: forward guidance announcements and the large-scale asset purchase program. We will focus on one particular episode of the Great Recession illustrating the role of forward guidance. In August 2010 the Federal Open Market Committee announced: "*The Committee will keep constant the Federal Reserve's holdings of securities at their current level by reinvesting principal payments from agency debt and agency mortgage-backed securities in longer-term Treasury securities.*" This announcement was an important prelude for the second part of the Quantitative Easing program (QE2) (see p. 244 in Krishnamurthy and Vissing-Jorgensen (2011) for a detailed discussion). In addition, this announcement generated a drop in the intraday yield for two- and ten- year treasury bond. In fact, from the end of July 2010 to the end of August 2010 the 2 year Treasury bond rate fell by 10 basis points.

Figure 5 uses our delta-method approach to construct confidence bands for the evolution of the levels of the four variables in the monetary SVAR. We fix all the variables at their level on July 2010 and we trace their evolution (over a 12-month window) according to the confidence set for their cumulative responses. The motivation for this exercise is as follows. Suppose that—back in August 2010—an econometrician is asked to provide confidence bands for the evolution of IP, CPI, 2YTB, and FF after the August 2010 announcement of the Federal Open Market Committee (FOMC). The econometrician observes the realization of the macroeconomic variables from July 1979 until August 2010, but decides to deliberately ignore the two years of data after the crisis (to avoid introducing structural changes, stochastic volatility, or any other feature that will complicate the estimation of the VAR).

The econometrician uses the data until December 2007—one semester before the financial crisis—to conduct delta-method inference on the cumulative responses to a one standard deviation unconventional monetary policy shock. The econometrician then uses these cumulative responses to get a rough idea of the evolution of the variables (in levels) following the announcement of the Federal Reserve in August 2010. The econometrician assumes there is a linear trend for CPI/IP, and ignores sampling uncertainty coming from the trend estimation in reporting the bands.
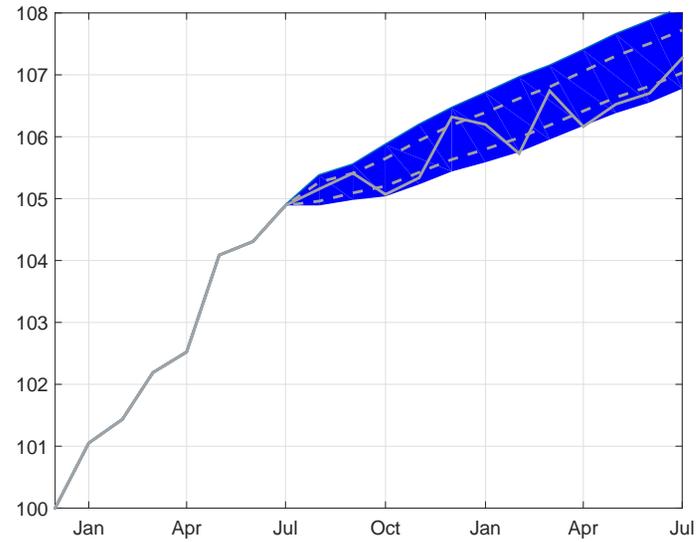
An ex-post evaluation of this exercise (over a window of 12 months) is reported in Figure 5.[17] We note that the observed dynamics for CPI, IP, GS2, and FFR from August 2010 to July 2011 fall within the bounds motivated by our delta-method confidence interval. We also note that our delta-method confidence interval misses the observed value at most three out of 12 months, which means that our 68% confidence-set covers each of these variables at least 75% of the time. We also report the 68% Bayesian credible sets.

---

[17]The reason to focus in a 12-month window is to cover the period between the QE2 announcement and the announcement of the so-called "Operation Twist" in September 2011. See http://www.federalreserve.gov/newsevents/press/monetary/20110921a.htm.
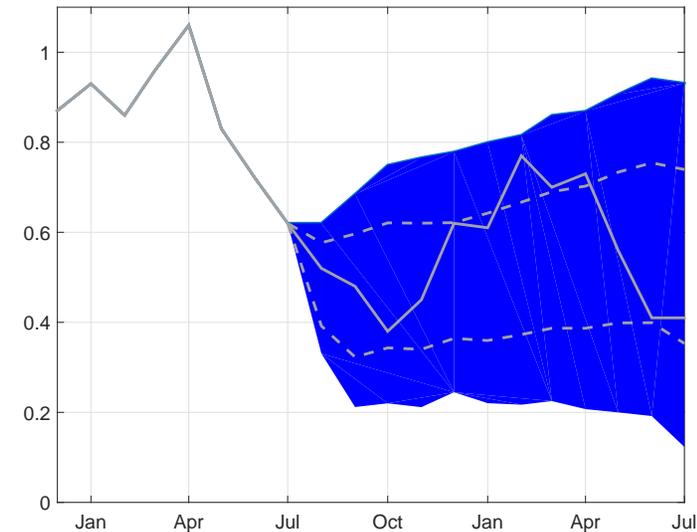
Figure 5: Delta-Method Confidence Interval for CPI, IP, 2yTB, FF after the August 2010 announcement
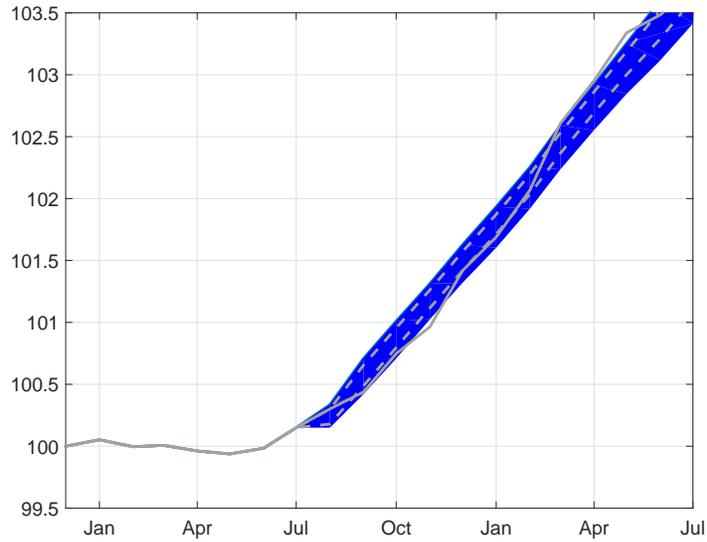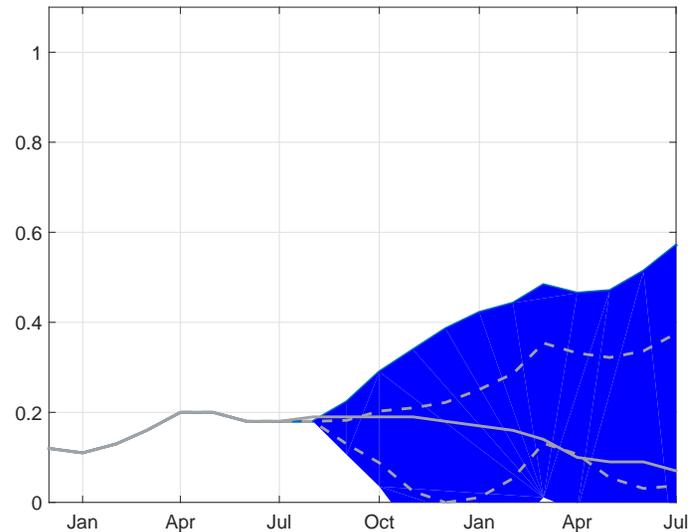

(a) Consumer Price Index

(b) Industrial Production Index

(c) 2-year Treasury Bond Rate

(d) Federal Funds Rate

(SHADED, BLUE AREA) Evolution of the Levels CPI, IP, 2yTB, and FF based on our 68% delta method confidence bands for the coefficients of Cumulative Impulse-Response Functions. (GRAY, SOLID LINE) Observed Levels of CPI, IP, 2yTB, and FF from December 2009 to July 2011. Both the CPI index and the IP index were normalized to have a starting value of 100. (GRAY, DASHED LINE) Evolution of the Levels CPI, IP, 2yTB, and FF based on the 68% credible set constructed using the priors in Uhlig (2005).

COMPUTATIONAL COST: We close this section with some comments regarding the computational cost of our delta-method procedure. Most of the work to compute the endpoints of the identified set and its derivatives is analytical. Consequently, practitioners can expect the computational burden of our procedure to be low. We note that the implementation of our delta-method confidence interval in the running example takes only around .15 seconds (using a standard Laptop @2.4GHz IntelCore i7). With the same equipment, the standard Bayesian implementation required around 327 seconds for 10,000 draws (which means that we could have constructed 2,000 delta-method confidence intervals while we generated the Bayesian credible set).

COMPARISON WITH THE PROJECTION APPROACH: Figure 7 presents a comparison between the delta-method approach and the *projection* confidence interval recently proposed by Gafarov et al. (2016) [GMM16]. The projection confidence interval has three theoretical properties that we were not able to verify for the delta-method approach. First, the projection confidence interval is *uniformly* consistent in level. Second, the projection confidence interval yields valid inference for the whole impulse-response function and not only its scalar coefficients. Third, the projection confidence interval has—in large samples—Bayesian credibility of at least the nominal level (for a large class of priors).

In order to exploit our formulas for the endpoints of the identified set, we followed a different algorithm to the one suggested in GMM16. We used a random grid over the Wald ellipsoid for the reduced-form parameters and reported the range of IRFs over this grid. The implementation of the projection confidence set (based on a random grid of 10,000 points) took around 1300 seconds (4 times slower than the Bayesian credible set and 8,000 times slower than the delta-method). We note that the projection confidence interval (which is wider than the delta-method bands) contains the realized value of IP, CPI, 2YTB, and FF for every horizon under consideration.

COMPARISON WITH CALIBRATED PROJECTION: The projection confidence interval covers the structural parameters more often than necessary. GMM16 show that when the endpoints of the identified set are differentiable, one can project a Wald ellipsoid with a radius given by $z_{1-\alpha}^2$ to eliminate projection bias. The calibrated confidence set will be approximately given by:

$$\left[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha} \, \underline{\widehat{\sigma}}_{(k,i,j),T}/\sqrt{T}, \ \ \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha} \, \overline{\widehat{\sigma}}_{(k,i,j),T}/\sqrt{T} \right],$$

where $\underline{\widehat{\sigma}}_{(k,i,j),T}, \overline{\widehat{\sigma}}_{(k,i,j),T}$ are estimators of the asymptotic variance of the plug-in estimators for the endpoints of the identified set. Figure 8 reports the calibrated projection.

Under the differentiability assumption, both $\underline{\widehat{\sigma}}_{(k,i,j),T}, \overline{\widehat{\sigma}}_{(k,i,j),T}$ are smaller (for every data realization) than $\widehat{\sigma}_{(k,i,j),T}$. This suggests that—whenever the endpoints of the identified set are fully differentiable—the calibrated projection should deliver smaller confidence intervals than our delta-method approach, as our formula for the standard error takes into account the

possibility that the endpoints are only directionally differentiable. In terms of computation time, the calibrated projection takes approximately 109,365 seconds (solving the nonlinear program described in GMM16).

COMPARISON WITH THE ROBUST APPROACH: Finally, Figure 6 in the Appendix reports the robust-Bayesian credible set in Giacomini and Kitagawa (2014)[2014]. The implementation of the robust-Bayes credible set (based on 10,000 posterior draws and using our algorithm to evaluate the endpoints) took around 9106 seconds.[18]

8. **CONCLUSION**

This paper focused on set-identified SVARs that impose equality and inequality restrictions to set-identify only one structural shock. For this class of models, the endpoints of the identified set have special properties that allow an intuitive and computationally simple approach to conduct frequentist inference. Specifically, the paper made three contributions:

(i) We presented an algorithm to compute—for each horizon, each variable, a fixed vector of reduced-form parameters, and a given collection of equality and/or inequality restrictions—the largest and smallest value of the coefficients of the structural IRF (see Proposition 1). Our algorithm does not require random sampling from the space of rotation matrices or unit vectors. Instead, we treated the bounds of the identified set as the *maximum and minimum value* of a mathematical program whose solutions we can characterize analytically.

(ii) We provided sufficient conditions under which the largest and smallest value of the structural parameters are directionally differentiable functions of the reduced-form parameters (see Proposition 2). This result seems to be of interest in its own right, and for example, could be used to explore the frequentist properties of the robust-Bayesian procedure in Giacomini and Kitagawa (2014).

(iii) We proposed a computationally convenient delta-method confidence interval for the set-identified coefficients of the structural IRF. We presented sufficient conditions to guarantee the pointwise consistency in level of the suggested inference approach. The delta-method in this paper exploited the structure of the directional derivative.

REFERENCES

AN, L. AND J. WANG (2012): "Exchange rate pass-through: Evidence based on vector autoregression with sign restrictions," *Open Economies Review*, 23, 359–380.

BAUMEISTER, C. AND L. BENATI (2013): "Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound," *International Journal of Central Banking*, 9, 165–212.

---

[18]Out of which 1266 seconds were used just to compute the identified set for each posterior draw, and the remaining time to translate the posterior bounds into the GK robust bounds

BAUMEISTER, C. AND J. HAMILTON (2015): "Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information," *Econometrica*, 5, 1963–1999.

BEAUDRY, P., D. NAM, AND J. WANG (2014): "Do mood swings drive business cycles and is it rational?" *NBER Workind Paper (w17651)*.

BHATTARAI, S., G. B. EGGERTSSON, AND B. GAFAROV (2014): "Time Consistency and the Duration of Government Debt: A Signalling Theory of Quantitative Easing," *Working paper, Penn State University*.

BLANCHARD, O. J. AND D. QUAH (1989): "The dynamic effects of aggregate demand and supply disturbances," *The American Economic Review*, 79, 655–673.

BONNANS, J. F. AND A. SHAPIRO (2000): *Perturbation analysis of optimization problems*, Springer.

CANOVA, F. AND G. D. NICOLÓ (2002): "Monetary disturbances matter for business fluctuations in the G-7," *Journal of Monetary Economics*, 49, 1131–1159.

DÜMBGEN, L. (1993): "On nondifferentiable functions and the bootstrap," *Probability Theory and Related Fields*, 95, 125–140.

FANG, Z. AND A. SANTOS (2014): "Inference on Directionally Differentiable Functions," *Working paper, University of California at San Diego*.

FAUST, J. (1998): "The Robustness of Identified VAR Conclusions about Money," in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 49, 207–244.

FIACCO, A. V. AND Y. ISHIZUKA (1990): "Sensitivity and stability analysis for nonlinear programming," *Annals of Operations Research*, 27, 215–235.

FUJITA, S. (2011): "Dynamics of worker flows and vacancies: evidence from the sign restriction approach," *Journal of Applied Econometrics*, 26, 89–121.

GAFAROV, B., M. MEIER, AND J. L. MONTIEL OLEA (2016): "Projection Inference for Set-Identified SVARs," *Working paper, New York University*.

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

GIACOMINI, R. AND T. KITAGAWA (2014): "Robust-Bayes Inference for Non-Identified SVARs," *Working Paper, University College London*.

KILIAN, L. AND D. P. MURPHY (2012): "Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models," *Journal of the European Economic Association*, 10, 1166–1188.

KITAGAWA, T., J. PAYNE, AND J. L. MONTIEL OLEA (2016): "Calibrated Inference for nondifferentiable functions," *Working paper, New York University*.

KRISHNAMURTHY, A. AND A. VISSING-JORGENSEN (2011): "The effects of quantitative easing on interest rates: channels and implications for policy," *Brookings Papers on Economic Activity*.

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.

——— (2007): *New Introduction to Multiple Time Series Analysis*, Springer.

Moon, H. R., F. Schorfheide, and E. Granziera (2013): "Inference for VARs Identified with Sign Restrictions," *University of Southern California*.

Mountford, A. and H. Uhlig (2009): "What are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics*, 24, 960–992.

Nocedal, J. and S. Wright (2006): *Numerical optimization*, Springer Science & Business Media, second edition ed.

Rubio-Ramirez, J., D. Caldara, and J. Arias (2015): "The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure," Tech. rep., Society for Economic Dynamics.

Rubio-Ramirez, J. F., D. F. Waggoner, and T. Zha (2010): "Structural vector autoregressions: Theory of identification and algorithms for inference," *The Review of Economic Studies*, 77, 665–696.

Shapiro, A. (1990): "On concepts of directional differentiability," *Journal of optimization theory and applications*, 66, 477–487.

——— (1991): "Asymptotic analysis of stochastic programs," *Annals of Operations Research*, 30, 169–186.

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

Vargas-Silva, C. (2008): "Monetary policy and the US housing market: A VAR analysis imposing sign restrictions," *Journal of Macroeconomics*, 30, 977–990.

Appendix

APPENDIX A: MAIN RESULTS

A.1. *Lemma 1*

Let $S(\mu)$ denote the $n \times m_s$ matrix of m_s 'sign' restrictions and let $Z(\mu)$ denote the $m_z \times n$ matrix of 'zero' restrictions. For notational simplicity, we deliberately ignore the dependence of the equality/inequality restrictions on μ. The problem in equation (2.5) is equivalent to:

(A.1) $\qquad \overline{v}_{k,i,j}(A, \Sigma) \equiv \max_{x \in \mathbb{R}^n} e_i' C_k(A) x \quad \text{subject to} \quad x' \Sigma^{-1} x = 1, \quad S'x = \mathbf{0}_{m_s \times 1}, \quad Z'x = \mathbf{0}_{m_z \times 1}.$

The auxiliary Lagrangian function is given by:

$$\mathcal{L}(x, w_1, w_2, w_3; A, \Sigma) = e_i' C_k x - w_1(x' \Sigma^{-1} x - 1) - w_2'(S'x) - w_3'(Z'x).$$

SInce $Z(\mu)$ and $S(\mu)$ are linearly independent at μ we can characterize the maximum response using the Karush-Kuhn-Tucker conditions for the mathematical program in (2.5). The Karush-Kuhn-Tucker necessary conditions for this problem are as follows:

$$
\begin{aligned}
\text{Stationarity} \quad &: \quad C_k'(A)e_i - 2w_1 \Sigma^{-1} x - S w_2 - Z w_3 = \mathbf{0}_{n \times 1}, \\
\text{Primal Feasibility} \quad &: \quad x' \Sigma^{-1} x = 1, \\
& \qquad S'x \geq 0_{m_s \times 1}, \\
& \qquad Z'x = \mathbf{0}_{m_z \times 1}, \\
\text{Complementary Slackness} \quad &: \quad \omega_{2i}(e_i' S x) = 0 \quad \forall \quad i = 1 \ldots m_s,
\end{aligned}
$$

plus the additional dual feasibility constraint requiring the Lagrange multipliers, ω_{2i}, to be smaller than or equal to zero.

Let $x^*(A, \Sigma, r)$ be one (out of possibly many) maximizer of the program of interest and suppose that the $m \times n$ matrix $(m \leq n - 1)$ r collects all the restrictions that are active (binding). Because of Assumption 1 and the fact that Z and S are linearly independent at μ, the matrix r is of full rank m and m must be smaller than or equal $n - 1$. Using stationarity, primal feasibility, and complementary slackness at x^* we get:

$$
\begin{aligned}
0 = x^{*\prime}[C_k(A)'e_i - 2w_1^* \Sigma^{-1} x^* - S w_2 - Z w_3] \quad &= \quad x^{*\prime} C_k(A)'e_i - 2w_1^* x^{*\prime} \Sigma^{-1} x^* - x^{*\prime} S w_2 - x^{*\prime} Z w_3 \\
&= \quad x^{*\prime} C_k(A)'e_i - 2w_1^* - x^{*\prime} S w_2 - x^{*\prime} Z w_3 \\
& \qquad \text{(where we have used } x^{*\prime} \Sigma^{-1} x^* = 1) \\
&= \quad x^{*\prime} C_k(A)'e_i - 2w_1^* \\
& \qquad \text{(where we have used } x^{*\prime} Z = \mathbf{0}_{m_z \times 1} \text{ and complementary slackness)} \\
&= \quad \overline{v}_{k,i,j}(A, \Sigma) - 2w_1^*.
\end{aligned}
$$

where $\overline{v}_{k,i,j}(A, \Sigma)$ denotes the value of the maximum response when the constraints in r are active. Thus, the Lagrange multiplier w_1^* is unique and given by:

$$w_1^* = \frac{1}{2} \overline{v}_{k,i,j}(A, \Sigma).$$

Note also that $w_1^* \neq 0$ if and only if $\overline{v}_{k,ij}(A, \Sigma) \neq 0$. We now show that if $\overline{v}_{k,i,j}(A, \Sigma) \neq 0$ there are unique w_2^* and w_3^* that satisfy the Karush-Kuhn Tucker conditions. Note that left multiplying the stationarity condition by Σ we have:

$$2w_1^* x^{*\prime} = \left(C_k(A)'e_i - rw\right)' \Sigma.$$

where w collects the nonzero components of ω_2 and all the components of ω_3.

$$(A.2) \qquad \left(C_k(A)'e_i - rw\right)' \Sigma \left(C_k(A)'e_i - rw\right) \;=\; 4w_1^2 x' \Sigma^{-1} x$$

$$= \; 4w_1^2$$

$$\text{(where we have used } \; x^{*\prime}\Sigma^{-1}x^* = 1\text{)}$$

$$= \; 4\left(\frac{1}{2}\overline{v}_{k,i,j}(A,\Sigma)\right)^2$$

$$= \; \overline{v}_{k,i,j}(A,\Sigma)^2.$$

Consequently the value function given active constraints z is given by either:

$$\overline{v}_{k,i,j}(A,\Sigma) = \left[\left(C_k(A)'e_i - rw\right)' \Sigma \left(C_k(A)'e_i - rw\right)\right]^{1/2},$$

or

$$\overline{v}_{k,i,j}(A,\Sigma) = -\left[\left(C_k(A)'e_i - rw\right)' \Sigma \left(C_k(A)'e_i - rw\right)\right]^{1/2}.$$

We will use the first order conditions to find the vector of Lagrange multipliers w and show that they are unique given $\overline{v}_{k,i,j}(A,\Sigma) \neq 0$. Note that

$$0 = 2w_1^* r' x^* \;=\; \overline{v}_{k,i,j}(A,\Sigma)\left[r'\Sigma(C_k(A)'e_i - rw)\right]$$

$$= \; \overline{v}_{k,i,j}(A,\Sigma)\left[r'\Sigma C_k(A)'e_i - r'\Sigma rw\right].$$

Under the assumptions of the Lemma, r is of rank m. If $\overline{v}_{k,i,j}(A,\Sigma) \neq 0$, the equation above holds if and only if

$$w^* = (r'\Sigma r)^{-1} r'\Sigma C_k(A)'e_i.$$

Consequently, the Lagrange multipliers for the active restrictions are unique. To conclude the proof, we get an explicit expression of the value function in terms of (A,Σ). To do so, note that:

$$\Sigma^{1/2}\left(C_k(A)'e_i - rw\right) \;=\; \Sigma^{1/2}C_k(A)'e_i - \Sigma^{1/2}rw$$

$$= \; \Sigma^{1/2}C_k(A)'e_i - \Sigma^{1/2}r(r'\Sigma r)^{-1}r'\Sigma C_k(A)'e_i$$

$$= \; \left(\mathbb{I}_n - \Sigma^{1/2}r(r'\Sigma r)^{-1}r'\Sigma^{1/2}\right)\Sigma^{1/2}C_k(A)'e_i$$

$$= \; \left(\mathbb{I}_n - P_{\Sigma^{1/2}r}\right)\Sigma^{1/2}C_k(A)'e_i$$

$$= \; M_{\Sigma^{1/2}r}\Sigma^{1/2}C_k(A)'e_i.$$

Therefore, the equation above and (A.2) imply that if $\overline{v}_{k,i,j}(A,\Sigma) \neq 0$ then either:

$$\overline{v}_{k,i,j}(A,\Sigma) = \left[e_i'C_k(A)\Sigma^{1/2}M_{\Sigma^{1/2}r}\Sigma^{1/2}C_k(A)'e_i\right]^{1/2}$$

or

$$\overline{v}_{k,i,j}(A,\Sigma) = -\left[e_i'C_k(A)\Sigma^{1/2}M_{\Sigma^{1/2}r}\Sigma^{1/2}C_k(A)'e_i\right]^{1/2}.$$

Furthermore, since any solution for which z is the set of binding constraints satisfies $2w_1^* x^{*\prime} = (C_k(A)' e_i - rw)'\Sigma$, then the solution vectors for which the constraints in r are binding are:

$$x^* = \Sigma^{1/2}\left(M_{\Sigma^{1/2}r}\right)\Sigma^{1/2}C_k(A)'e_i \Big/ \left[e_i'C_k(A)\Sigma^{1/2}M_{\Sigma^{1/2}r}\Sigma^{1/2}C_k(A)'e_i\right]^{1/2},$$

or

$$x^* = -\Sigma^{1/2}\left(M_{\Sigma^{1/2}r}\right)\Sigma^{1/2}C_k(A)'e_i \Big/ \left[e_i'C_k(A)\Sigma^{1/2}M_{\Sigma^{1/2}r}\Sigma^{1/2}C_k(A)'e_i\right]^{1/2}.$$

In any case the lagrange multipliers for the active constraints are given (as shown above) by,

$$w^* = (r'\Sigma r)^{-1}r'\Sigma C_k(A)'e_i.$$

REMARK TO THE PROOF OF LEMMA 1: If $\overline{v}_{k,i,j}(A, \Sigma) = 0$, then neither the maximizer x^* nor the Lagrange multipliers w^* are unique. Note that any $x \in \mathbb{R}^n$ orthogonal to $C_k(A)'e_i$ satisfying the ellipsoid constraint $x'\Sigma^{-1}x = 1$ and the sign/zero restrictions is a solution. In addition, any vector of Lagrange multipliers satisfying the equation $C_k'e_i - rw = \mathbf{0}_{n\times1}$, satisfies the F.O.C.

A.2. *Proof of Proposition 1*

Let $x^* \in \mathbb{R}^n$ be a solution to the program (2.5) and let r^* be the set of constraints that are active at x^*.

STEP 1: We show first that

$$\overline{v}_{k,i,j}(A, \Sigma) \geq \max_{r \in R} \left(\max\{f^+_{max}(A, \Sigma; r), f^-_{max}(A, \Sigma; r)\} \right).$$

We do so by considering two different cases.

Case 1.1: Take any $r \in R$, and assume first that $v_{k,i,j}(A, \Sigma, r) \neq 0$. If $\mathbf{1}_{m_s}(x^*_+(A, \Sigma, r)) = 0$, then

$$f^+_{max}(A, \Sigma, r) = v_{k,i,j}(A, \Sigma, r) - 2c \leq c - 2c = -c < \overline{v}_{k,i,j}(A, \Sigma).$$

If, however, $r \in \mathbb{R}$ is such that $\mathbf{1}_{m_s}(x^*_+(A, \Sigma, r)) = 1$, then $x^*_+(A, \Sigma, r)$ satisfies all the equality and inequality restrictions in (2.5) and, by construction, also satisfies the ellipsoid constraint

$$x^*_+(A, \Sigma, r)' \Sigma^{-1} x^*_+(A, \Sigma, r) = 1.$$

Consequently, $\overline{v}_{k,i,j}(A, \Sigma) \geq f^+_{max}(A, \Sigma, r)$ for all $r \in R$. An analogou argument shows that $\overline{v}_{k,i,j}(A, \Sigma) \geq f^-_{max}(A, \Sigma, r)$. This implies that:

$$\overline{v}_{k,i,j}(A, \Sigma) \geq \max\{f^+_{max}(A, \Sigma, r), f^-_{max}(A, \Sigma, r)\},$$

for all $r \in \mathbb{R}$ such that $v_{k,i,j}(A, \Sigma, r) \neq 0$.

Case 1.2: Consider now any r such that $v_{k,i,j}(A, \Sigma, r) = 0$. If there is no feasible point x^* that gives such value, then $f^+_{max}(A, \Sigma, r) = f^-_{max}(A, \Sigma, r) = -c \leq \overline{v}_{k,i,j}(A, \Sigma)$. If there is such a feasible point $x^* \neq 0$ then $f^+_{max}(A, \Sigma, r) = f^-_{max}(A, \Sigma, r) = 0$ Since x^* is in the choice set of the program 2.5, then $f^+_{max}(A, \Sigma, r) = f^-_{max}(A, \Sigma, r) = 0 \leq \overline{v}_{k,i,j}(A, \Sigma)$.

Therefore, Case 1.1 and 1.2 imply that:

$$\overline{v}_{k,i,j}(A, \Sigma) \geq \max\{f^+_{max}(A, \Sigma, r), f^-_{max}(A, \Sigma, r)\} \quad \text{for all } r \in R.$$

STEP 2: We now show that

$$\overline{v}_{k,i,j}(A, \Sigma) \leq \max_{r \in R} \max\{f^+_{max}(A, \Sigma, r), f^-_{max}(A, \Sigma, r)\}.$$

Again, we consider two cases.

Case 2.1: Assume first that $\overline{v}_{k,i,j}(A, \Sigma) \neq 0$. Without loss of generality, let $\overline{v}_{k,i,j}(A, \Sigma) > 0$. Let $r^* \in R$ denote the set of active sign restrictions at the solution x^*. By Lemma 1, we know that

$$\overline{v}_{k,i,j}(A, \Sigma) = \left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2} r^*} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2},$$

and

$$x^*(A, \Sigma; r^*) = \Sigma^{1/2} \left(M_{\Sigma^{1/2} r^*} \right) \Sigma^{1/2} C_k(A)' e_i \Big/ \left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2} r^*} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2}.$$

Since this point satisfies the sign restrictions not in r^*, then

$$\left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2} r^*} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2} = f^+_{max}(A, \Sigma, r^*).$$

Consequently,

$$\overline{v}_{k,i,j}(A, \Sigma) \leq \max_{r \in R} \max\{f^+_{max}(A, \Sigma, r), f^-_{max}(A, \Sigma, r)\}.$$

Case 2.2: If $\overline{v}_{k,i,j}(A, \Sigma) = 0$, there is an $x^* \neq 0$ in the choice set. Hence, the Karush-Kuhn-Tucker conditions implies that $C_k(A)'e_i$ is a linear combination of the active constraints that generate the value of zero (which means there is an r^* such that $f^+_{max}(A, \Sigma; r^*) = f^-_{max}(A, \Sigma; r^*) = 0$). Therefore, $\overline{v}_{k,i,j}(A, \Sigma) = f(A, \Sigma, r^*) \leq \max_{r \in R} \max\{f^+_{max}(A, \Sigma, r), f^-_{max}(A, \Sigma, r)\}$.

Step 1 and Step 2 shows that the value function $\overline{v}_{k,i,j}(A, \Sigma)$ is obtained by computing the Karush-Kuhn-Tucker points in Lemma 1 for each r, penalizing the value $\overline{v}_{k,i,j}(A, \Sigma; r)$ if unfeasible, and maximizing over all the possible values of r. The proof for the lower bound is analogous.

A.3. *Lemma 2*

Note that if $r_l(\mu)$ is differentiable and $v_{k,i,j}(\mu, r_l(\mu)) \neq 0$, then the function:

$$v_{k,i,j}(\mu; r_l(\mu)) = \left(e_i' C_k(A) \Sigma^{1/2} M_{\Sigma^{1/2} r_l(\mu)} \Sigma^{1/2} C_k(A)' e_i \right)^{1/2},$$

is differentiable as well. Moreover, the function

$$x^*(\mu; r_l(\mu)) \equiv \Sigma^{1/2} \left(M_{\Sigma^{1/2} r_l(\mu)} \right) \Sigma^{1/2} C_k(A)' e_i \Big/ v_{k,i,j}(\mu; r_l(\mu))$$

is also differentiable. Therefore,

$$
\begin{aligned}
\frac{\partial v_{k,i,j}(\mu, r_l(\mu))}{\partial \mu} &= \frac{\partial [e_i' C_k(A) x^*(\mu, r_l(\mu))]}{\partial \mu} \\
&\qquad \text{(since } v_{k,i,j}(\mu, r_l(\mu)) = e_i' C_k(A) x^*(\mu, r_l(\mu))) \\
&= \frac{\partial x^*(\mu, r_l(\mu))}{\partial \mu} C_k'(A) e_i + \frac{\partial (x^*(\mu; r_l(\mu))' \otimes e_i') \mathrm{vec}(C_k(A))}{\partial \mu}, \\
&\qquad \text{(where we have re-written } e_i' C_k(A) x^* \text{ as } (x^{*\prime} \otimes e_i') \mathrm{vec}(C_k(A))) \\
&= \frac{\partial x^*(\mu, r_l(\mu))}{\partial \mu} C_k'(A) e_i + \frac{\partial \mathrm{vec}(C_k(A))}{\partial \mu} (x^*(\mu; r_l(\mu)) \otimes e_i) \\
&\qquad \text{(where we have applied the chain rule for matrix derivatives).}
\end{aligned}
$$

We now use the envelope theorem to compute this derivative. Note that we have shown the existence of unique multipliers $\lambda^* \in \mathbb{R}$ and $w^* \in \mathbb{R}^l$ such that:

$$C_k(A)' e_i = \lambda^* 2 \Sigma^{-1} x^*(\mu; r_l(\mu)) + r_l(\mu) w^*.$$

Therefore:

$$
\begin{aligned}
\frac{\partial v_{k,i,j}(\mu, r_l(\mu))}{\partial \mathrm{vec}(A)} &= \frac{\partial x^*(\mu, r_l(\mu))}{\partial \mathrm{vec}(A)} \left[\lambda^* 2 \Sigma^{-1} x^*(\mu; r_l(\mu)) + r_l(\mu) w^* \right] \\
&\quad + \frac{\partial \mathrm{vec}(C_k(A))}{\partial \mathrm{vec}(A)} (x^*(\mu; r_l(\mu)) \otimes e_i).
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial v_{k,i,j}(\mu, r_l(\mu))}{\partial \mathrm{vec}(\Sigma)} &= \frac{\partial x^*(\mu, r_l(\mu))}{\partial \mathrm{vec}(\Sigma)} \left[\lambda^* 2 \Sigma^{-1} x^*(\mu; r_l(\mu)) + r_l(\mu) w^* \right] \\
&\quad + \frac{\partial \mathrm{vec}(C_k(A))}{\partial \mathrm{vec}(\Sigma)} (x^*(\mu; r_l(\mu)) \otimes e_i).
\end{aligned}
$$

Note also that:

$$
0 = \frac{\partial x^*(\mu, r_l(\mu))' \Sigma^{-1} x^*(\mu, r_l(\mu))}{\partial \mathrm{vec}(A)} = 2 \frac{\partial x^*(\mu, r_l(\mu))}{\partial \mathrm{vec}(A)} \Sigma^{-1} x^*(\mu, r_l(\mu))
$$

and

$$
\begin{aligned}
0 &= \frac{\partial r_l(\mu)' x^*(\mu, r_l(\mu))}{\partial \mathrm{vec}(A)} \\
&= \frac{\partial x^*(\mu, r_l(\mu))}{\partial \mathrm{vec}(A)} r_l(\mu) + \left(\frac{\partial r_{1,l}(\mu)}{\mathrm{vec}(A)} x^*(\mu, r_l(\mu)), \quad \cdots \quad , \frac{\partial r_{l,l}(\mu)}{\mathrm{vec}(A)} x^*(\mu, r_l(\mu)) \right),
\end{aligned}
$$

where $r_{k,l}(\mu)$ denotes the k-th column of $r_l(\mu)$. Consequently:

$$
\frac{\partial v_{k,i,j}(\mu, r_l(\mu))}{\partial \mathrm{vec}(A)} = \frac{\partial \mathrm{vec}(C_k(A))}{\partial \mathrm{vec}(A)} (x^*(\mu; r_l(\mu)) \otimes e_i) - \sum_{k=1}^{l} w_k^* \frac{\partial \mathrm{vec}(r_{k,l}(\mu))}{\mathrm{vec}(A)} x^*(\mu, r_l(\mu)),
$$

where w_k^* is the k-th entry of the vector of lagrange multipliers w^*. This gives the partial derivative of $v_{k,i,j}(\mu, r_l(\mu))$ with respect to $\text{vec}(A)$. We note that this derivative can also be written as:

$$\frac{\partial v_{k,i,j}(\mu, r_l(\mu))}{\partial \text{vec}(A)} \quad = \quad \frac{\partial \text{vec}(C_k(A))}{\partial \text{vec}(A)}(x^*(\mu; r_l(\mu)) \otimes e_i) - \frac{\partial \text{vec}(r_l(\mu)')}{\text{vec}(A)}(x^*(\mu, r_l(\mu)) \otimes \mathbb{I}_l)w^*,$$

which is the expression given in the overview. Finally, to get the derivative with respect to $\text{vec}(\Sigma)$ we note that:

$$\begin{aligned}
0 \quad &= \quad \frac{\partial x^*(\mu, r_l(\mu))'\Sigma^{-1}x^*(\mu, r_l(\mu))}{\partial \text{vec}(\Sigma)} = 2\frac{\partial x^*(\mu, r_l(\mu))}{\partial \text{vec}(\Sigma)}\Sigma^{-1}x^*(\mu, r_l(\mu)) \\
&\quad - \quad (\Sigma^{-1}x^*(\mu, r_l(\mu)) \otimes \Sigma^{-1}x^*(\mu, r_l(\mu))),
\end{aligned}$$

and

$$\begin{aligned}
0 \quad &= \quad \frac{\partial r_l(\mu)'x^*(\mu, r_l(\mu))}{\partial \text{vec}(\Sigma)} \\
&= \quad \frac{\partial x^*(\mu, r_l(\mu))}{\partial \text{vec}(\Sigma)}r_l(\mu) + \left(\frac{\partial r_{1,l}(\mu)}{\text{vec}(\Sigma)}x^*(\mu, r_l(\mu)), \quad \cdots \quad , \frac{\partial r_{l,l}(\mu)}{\text{vec}(\Sigma)}x^*(\mu, r_l(\mu)) \right).
\end{aligned}$$

Consequently,

$$\frac{\partial v_{k,i,j}(\mu, r_l(\mu))}{\partial \text{vec}(\Sigma)} \quad = \quad \lambda^*(\Sigma^{-1}x^*(\mu, r_l(\mu)) \otimes \Sigma^{-1}x^*(\mu, r_l(\mu))) - \sum_{k=1}^{l} w_k^* \frac{\partial \text{vec}(r_{k,l}(\mu))}{\text{vec}(\Sigma)}x^*(\mu, r_l(\mu)).$$

A.4. *Proof of Proposition 2*

Let $r_1(\mu), r_2(\mu), \ldots r_M(\mu)$ denote the elements of $R(\mu)$. Each of these elements activate a different collection of sign restrictions. Without loss of generality, assume that $R^*(\mu)$ contains the first L elements $r_1(\mu), \ldots, r_L(\mu)$ of $R(\mu)$. Consider any sequence (A_T, Σ_T) such that

$$(\mathrm{vec}A_T{}', \mathrm{vech}\Sigma_T')' = (\mathrm{vec}A', \mathrm{vech}\Sigma')' + h_T/r_T,$$

where $h_T \to h$, $r_T \to \infty$ and such that for large enough T, $(\mathrm{vec}A_T, \mathrm{vech}\Sigma_T)$ belongs to the parameter space μ. Suppose that $\overline{v}_{k,i,j}(\mu) > 0$. Note that by Proposition 1:

$$\overline{v}_{k,i,j}(A_T, \Sigma_T) = \begin{cases} f(A_T, \Sigma_T, r_1(\mu_T)) & \text{if} & f(A_T, \Sigma_T, r_1(\mu_T)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r), \\ \vdots & \vdots & \vdots \\ f(A_T, \Sigma_T, r_L(\mu_T)) & \text{if} & f(A_T, \Sigma_T, r_L(\mu_T)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r), \\ \vdots & \vdots & \vdots \\ f(A_T, \Sigma_T, r_M(\mu_T)) & \text{if} & f(A_T, \Sigma_T, r_M(\mu_T)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r). \end{cases}$$

Since each $r_l(\mu)$ is assumed to be continuous and only the first L elements of $R(\mu)$ belong to $R^*(\mu)$ then (for T large enough):

$$\overline{v}_{k,i,j}(A_T, \Sigma_T) = \begin{cases} \overline{v}_{k,i,j}(A_T, \Sigma_T, r_1(\mu_T)) & \text{if} & v(A_T, \Sigma_T, r_1(\mu_T)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r), \\ \vdots & \vdots & \vdots \\ \overline{v}_{k,i,j}(A_T, \Sigma_T, r_L(\mu_T)) & \text{if} & v(A_T, \Sigma_T, r_L(\mu)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r), \\ \overline{v}_{k,i,j}(A_T, \Sigma_T, r_{L+1}(\mu_T)) & \text{if} & v(A_T, \Sigma_T, r_{L+1}(\mu)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r), \\ \quad -2(1 - \mathbf{1}_{m_s}(x_+^*(\mu, r_{L+1}(\mu)))) & & \\ \vdots & \vdots & \vdots \\ \overline{v}_{k,i,j}(A_T, \Sigma_T, r_M(\mu_T)) & \text{if} & f(A_T, \Sigma_T, r_M(\mu)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r). \\ \quad -2(1 - \mathbf{1}_{m_s}(x_+^*(\mu, r_M(\mu)))) & & \end{cases}$$

Note that $\overline{v}_{k,i,j}(\mu) = \overline{v}_{k,i,j}(\mu; r_l(\mu))$ for all $r_l(\mu) \in R^*(\mu)$ and $\overline{v}_{k,i,j}(\mu) > \overline{v}_{k,i,j}(\mu; r_l(\mu))$ for all $r_l(\mu) \notin R^*(\mu)$. Therefore, the equation above implies that $r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T) - \overline{v}_{k,i,j}(A, \Sigma))$ equals

$$\begin{cases} r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T, r_1(\mu_T)) - \overline{v}_{k,i,j}(\mu, r_1(\mu_T))) & \text{if} & v(A_T, \Sigma_T, r_1(\mu_T)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r), \\ \vdots & \vdots & \vdots \\ r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T, r_L(\mu_T)) - \overline{v}_{k,i,j}(\mu, r_L(\mu))) & \text{if} & v(A_T, \Sigma_T, r_L(\mu)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r), \\ \vdots & \vdots & \vdots \\ r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T, r_M(\mu_T)) - \overline{v}_{k,i,j}(\mu, r_M(\mu))) & \text{if} & f(A_T, \Sigma_T, r_M(\mu)) = \max_{r \in R(\mu_T)} f(A_T, \Sigma_T, r). \\ \quad + r_T(\overline{v}_{k,i,j}(\mu, r_M(\mu)) - \overline{v}_{k,i,j}(\mu)) & & \\ \quad - r_T 2(1 - \mathbf{1}_{m_s}(x^*(\mu, r_M(\mu)))) & & \end{cases}$$

For T large enough, this implies that $r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T) - \overline{v}_{k,i,j}(A, \Sigma))$ equals

$$\begin{cases} r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T, r_1(\mu_T)) - \overline{v}_{k,i,j}(\mu, r_1(\mu))) & \text{if} & v(A_T, \Sigma_T, r_1(\mu_T)) = \max_{r \in R^*(\mu_T)} \overline{v}_{k,i,j}(A_T, \Sigma_T, r), \\ \vdots & \vdots & \vdots \\ r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T, r_L(\mu_T)) - \overline{v}_{k,i,j}(\mu, r_L(\mu))) & \text{if} & v(A_T, \Sigma_T, r_L(\mu)) = \max_{r \in R^*(\mu_T)} \overline{v}_{k,i,j}(A_T, \Sigma_T, r). \end{cases}$$

Since each of $r_T(\overline{v}_{k,i,j}(A_T, \Sigma_T, r_1(\mu_T)) - \overline{v}_{k,i,j}(\mu, r_l(\mu)))$ in the previous expression is, by Lemma 2, approximately equal to

$$\dot{v}_{k,i,j}(\mu; r_l(\mu))'h.$$

Then

$$r_T\left(\overline{v}_{k,i,j}(A_T, \Sigma_T) - \overline{v}_{k,i,j}(A, \Sigma)\right) \to \max_{r \in R^*(\mu)}\left(\dot{v}_{k,i,j}(\mu; r)'h\right).$$

A.5. *Proof of Proposition 3*

Let P denote the data generating process. For notational simplicity we write μ instead of $\mu(P)$ and Ω instead of $\Omega(P)$ whenever convenient. Note first that

(A.3) $\qquad P\left(\lambda \in \left[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T},\ \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T}\right]\right)$

equals

$$P\left(\sqrt{T}(\underline{v}_{k,i,j}(\widehat{\mu}_T) - \underline{v}_{k,i,j}(\mu)) \le z_{1-\alpha/2}\,\widehat{\sigma}_T + \sqrt{T}(\lambda - \underline{v}_{k,i,j}(\mu))\ \text{and}\right.$$

$$\left.\sqrt{T}(\lambda - \overline{v}_{k,i,j}(\mu)) - z_{1-\alpha/2}\widehat{\sigma}_T \le \sqrt{T}(\overline{v}_{k,i,j}(\widehat{\mu}_T) - \overline{v}_{k,i,j}(\mu))\right).$$

Since

$$\sqrt{T}(\lambda - \underline{v}_{k,i,j}(\mu)) \ge 0 \ \text{and} \ \sqrt{T}(\lambda - \overline{v}_{k,i,j}(\mu)) \le 0,$$

(A.3) is bounded from below by

$$P\left(\sqrt{T}(\underline{v}_{k,i,j}(\widehat{\mu}_T) - \underline{v}_{k,i,j}(\mu)) \le z_{1-\alpha/2}\,\widehat{\sigma}_T \ \text{and} \ -z_{1-\alpha/2}\widehat{\sigma}_T \le \sqrt{T}(\overline{v}_{k,i,j}(\widehat{\mu}_T) - \overline{v}_{k,i,j}(\mu))\right),$$

which is itself bounded from below by:

$$P\left(\sqrt{T}(\underline{v}_{k,i,j}(\widehat{\mu}_T) - \underline{v}_{k,i,j}(\mu)) \le z_{1-\alpha/2}\widehat{\sigma}_T \ \text{and} \ -z_{1-\alpha/2}\widehat{\sigma}_T \le \sqrt{T}(\overline{v}_{k,i,j}(\widehat{\mu}_T) - \overline{v}_{k,i,j}(\mu)), \ \text{and} \ \|\sqrt{T}(\widehat{\mu}_T - \mu)\| \le M_\epsilon\right),$$

where M_ϵ is such that

$$P\left(\|\zeta(P)\| > M_\epsilon\right) \le \epsilon.$$

Since, by Proposition 2, both $\underline{v}_{k,i,j}(\cdot)$ and $\overline{v}_{k,i,j}(\mu)$ are directionally differentiable with directional derivatives:

$$\min_{r \in R_*(\mu)}\left[\dot{v}_{k,i,j}(\mu, r)'h\right],$$

and

$$\max_{r \in R^*(\mu)}\left[\dot{v}_{k,i,j}(\mu, r)'h\right].$$

The directional differentiability implies that for any compact set K there is T large enough such that for any $h \in K$:

$$-\epsilon \le \sqrt{T}(\underline{v}_{k,i,j}(\mu + h/\sqrt{T}) - \underline{v}_{k,i,j}(\mu)) - \min_{r \in R_*(\mu)}\left[\dot{v}_{k,i,j}(\mu, r)'h\right] \le \epsilon,$$

and

$$-\epsilon \le \sqrt{T}(\overline{v}_{k,i,j}(\mu + h/\sqrt{T}) - \overline{v}_{k,i,j}(\mu)) - \max_{r \in R^*(\mu)}\left[\dot{v}_{k,i,j}(\mu, r)'h\right] \le \epsilon.$$

Therefore, for T large enough:

$$\inf_{\lambda \in \mathcal{I}_{k,i,j}^{\mathcal{R}}(\mu(P))} P\left(\lambda \in \left[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T},\ \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T}\right]\right)$$

is bounded from below by:

$$P\Big(\min_{r \in R_*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\sqrt{T}(\widehat{\mu}_T - \mu)\Big] \le z_{1-\alpha/2}\,\widehat{\sigma}_T \text{ and}$$

$$-z_{1-\alpha/2}\widehat{\sigma}_T \le \max_{r \in R^*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\sqrt{T}(\widehat{\mu} - \mu)\Big], \text{ and } ||\sqrt{T}(\widehat{\mu}_T - \mu)|| \le M_\epsilon\Big),$$

which, by Assumption 4, converges in distribution to:

$$P\Big(\min_{r \in R_*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\zeta(P)\Big] \le z_{1-\alpha/2}\,\sigma \text{ and}$$

$$-z_{1-\alpha/2}\sigma \le \max_{r \in R_*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\zeta(P)\Big], \text{ and } ||\zeta(P)|| \le M_\epsilon\Big),$$

where σ is the probability limit of $\widehat{\sigma}_T$:

$$\sigma \equiv \max_{r \in R(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\Omega\dot{v}_{k,i,j}(\mu,r)\Big],$$

and where we have used the fact that $\sigma > 0$.[19] Consequently,

$$\liminf_{T \to \infty} \inf_{\lambda \in \mathcal{I}^{\mathcal{R}}_{k,i,j}(\mu(P))} P\Big(\lambda \in \Big[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T},\ \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T}\Big]\Big)$$

is larger than or equal:

$$1 - P\Big(\min_{r \in R_*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\zeta(P)\Big] > z_{1-\alpha/2}\,\sigma\Big) - P\Big(\max_{r \in R_*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\zeta(P)\Big] < -z_{1-\alpha/2}\,\sigma\Big)$$

$$-P\Big(||\zeta(P)|| > M_\epsilon\Big).$$

Take some $r \in R_*(\mu)$ for which $\sigma(r) \equiv \dot{v}_{k,i,j}(\mu,r)'\Omega\dot{v}_{k,i,j}(\mu,r) > 0$ (one such r must exist for we have asumed that $\min_{r \in R(\mu)}[\dot{v}_{k,i,j}(\mu,r)'\Omega\dot{v}_{k,i,j}(\mu,r)] > 0$). Therefore:

$$P\Big(\min_{r \in R_*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\zeta(P)\Big] > z_{1-\alpha/2}\sigma\Big) \le P\Big(\dot{\overline{v}}_{k,i,j}(\mu,r)'\zeta(P) > z_{1-\alpha/2}\sigma\Big) \le \mathbb{P}\Big(N(0,1) > z_{1-\alpha/2}\frac{\sigma}{\sigma(r)}\Big),$$

which is at most $\alpha/2$ since $\sigma \ge \sigma(r)$.

Now, take some $r \in R^*(\mu)$ for which $\sigma(r) \equiv \dot{v}_{k,i,j}(\mu,r)'\Omega\dot{v}_{k,i,j}(\mu,r) > 0$. Note that

$$P\Big(\max_{r \in R^*(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\zeta(P)\Big] < -z_{1-\alpha/2}\sigma\Big) \le P\Big(\dot{v}_{k,i,j}(\mu,r)'\zeta(P) < -z_{1-\alpha/2}\sigma\Big) \le \mathbb{P}\Big(N(0,1) < -z_{1-\alpha/2}\frac{\sigma}{\sigma(r)}\Big),$$

which is less than or equal to $\le \alpha/2$ as $\sigma > \sigma(r)$. We conclude that

$$\liminf_{T \to \infty} \inf_{\lambda \in \mathcal{I}^{\mathcal{R}}_{k,i,j}(\mu(P))} P\Big(\lambda \in \Big[\underline{v}_{k,i,j}(\widehat{\mu}_T) - z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T},\ \overline{v}_{k,i,j}(\widehat{\mu}_T) + z_{1-\alpha/2}\,\widehat{\sigma}_T/\sqrt{T}\Big]\Big) \ge 1 - \alpha - \epsilon.$$
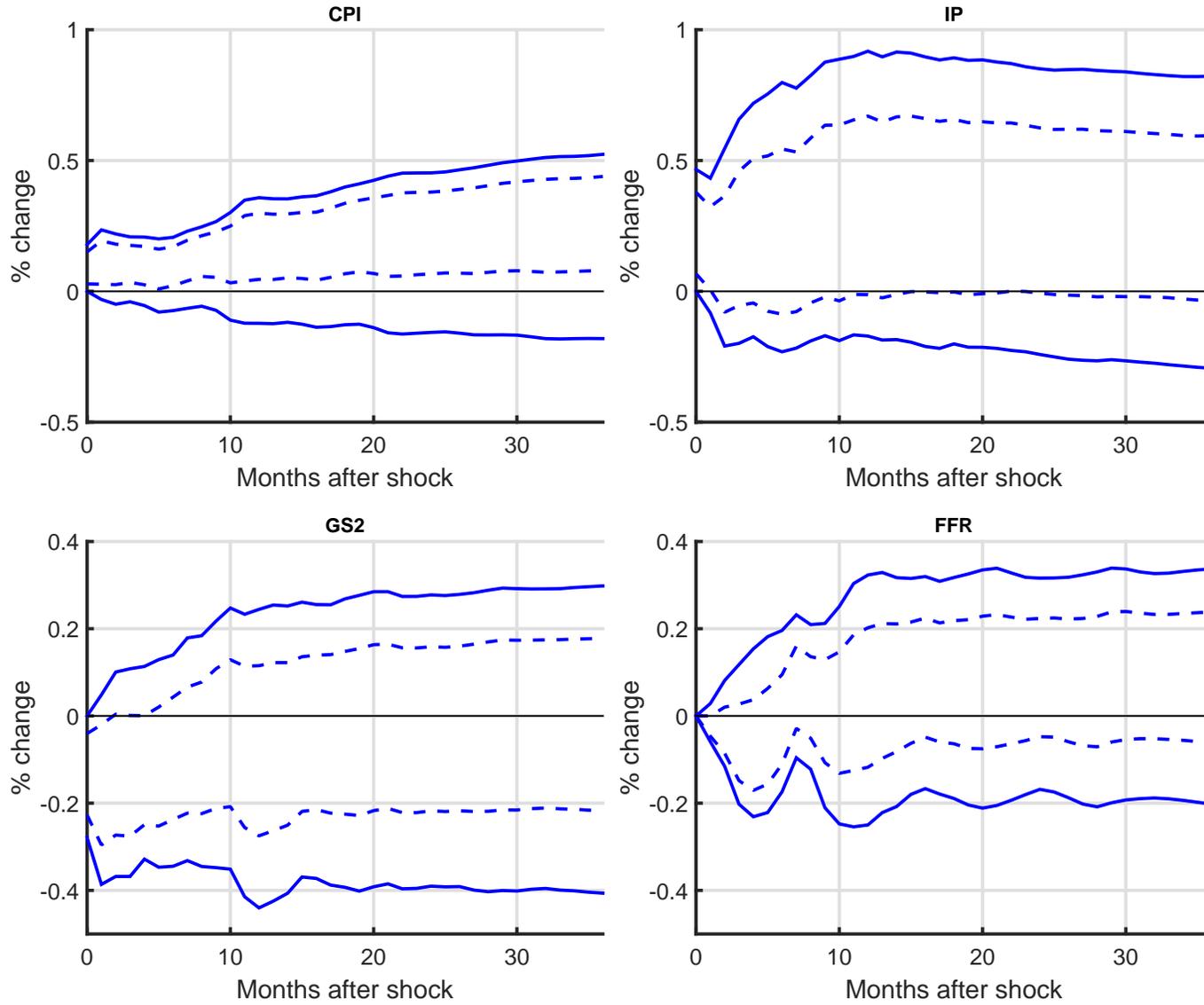
Since $\epsilon > 0$ is arbitrary, the desired result follows.

---

[19]This follows from the fact that:

$$\sigma \equiv \max_{r \in R(\mu)} \Big[\dot{v}_{k,i,j}(\mu,r)'\Omega\dot{v}_{k,i,j}(\mu,r)\Big] \ge \min_{r \in R(\mu)} \Big[\dot{v}_{k,i,j}(\mu(P),r)'\Omega(P)\dot{v}_{k,i,j}(\mu(P),r)\Big] > 0,$$

which we have assumed to be strictly bigger than zero.

Figure 6: Identified Set for the Cumulative Impulse Response Functions to an UMP shock (given $\widehat{\mu}_T$)
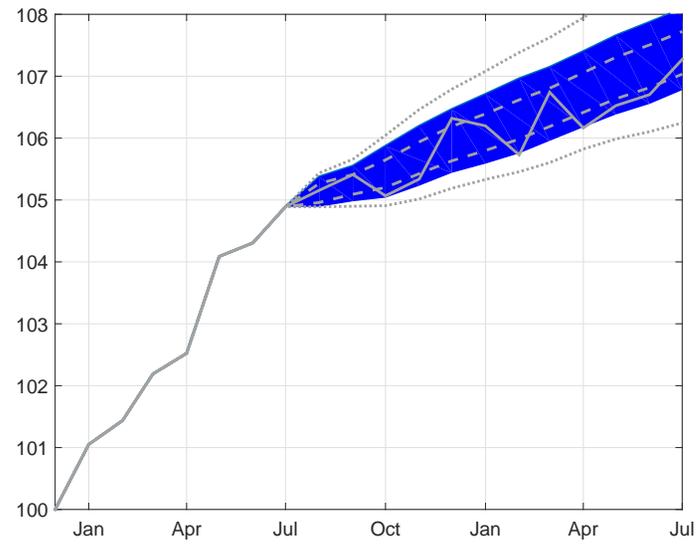
(SOLID, BLUE LINE) Endpoints of the identified set for the cumulative responses given $\widehat{\mu}_T$ and the equality/inequality restrictions in Table I. (BLUE, DASHED) Grid search approach to evaluate the identified set (10,000 uniform draws from a unit vector $q \in \mathbb{R}^4$).
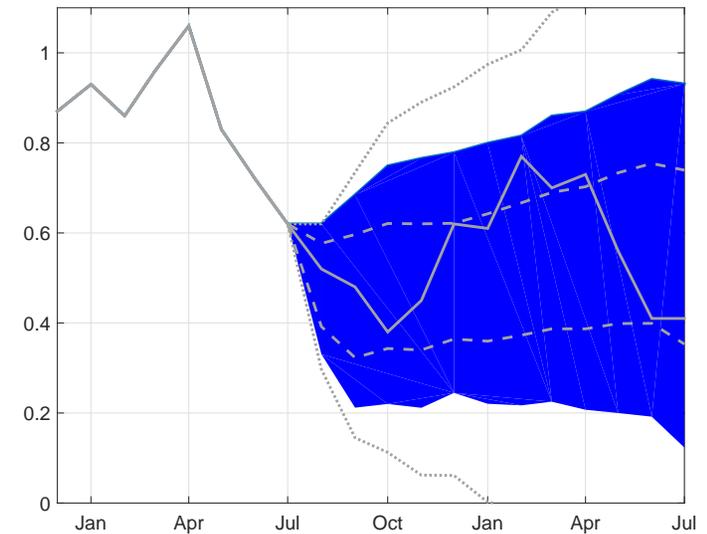
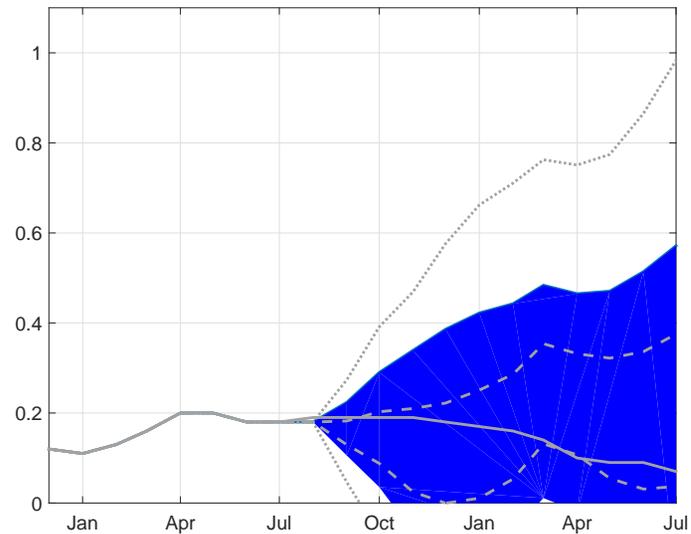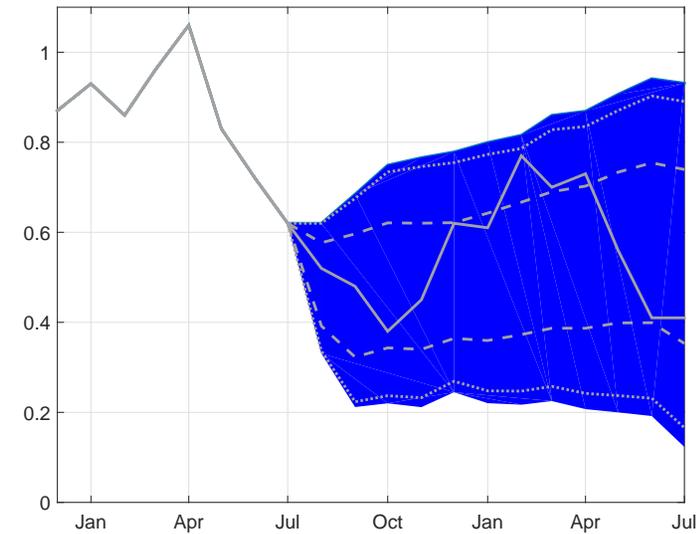Figure 7: Projection Confidence Interval for CPI, IP, 2yTB, FF after the August 2010 announcement
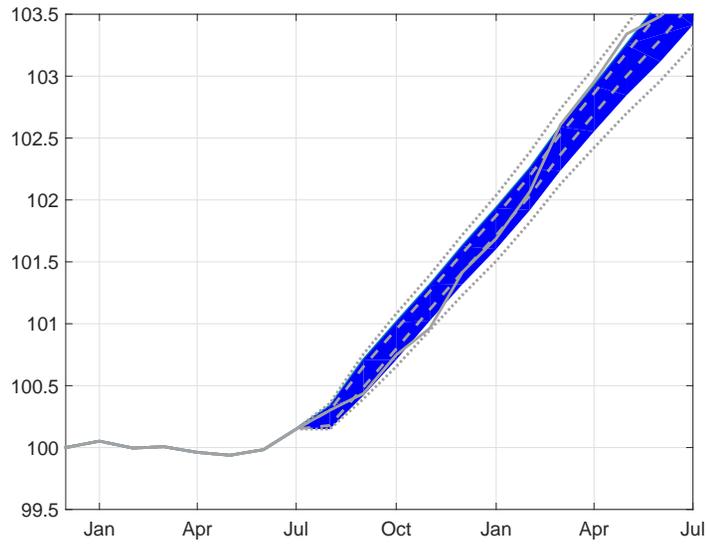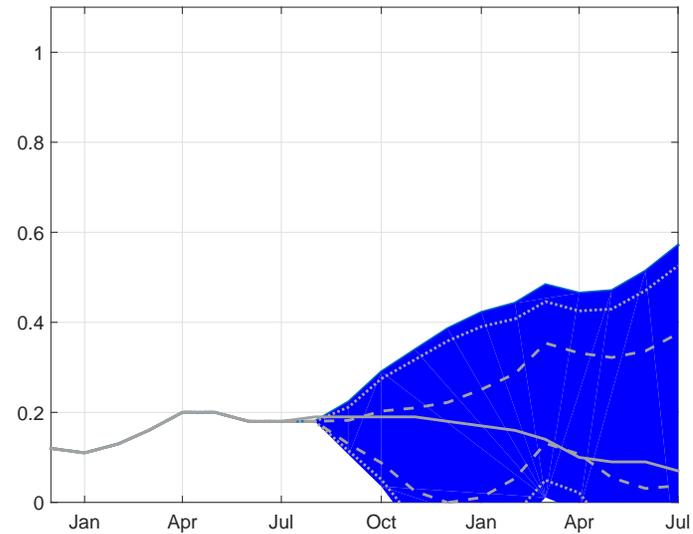

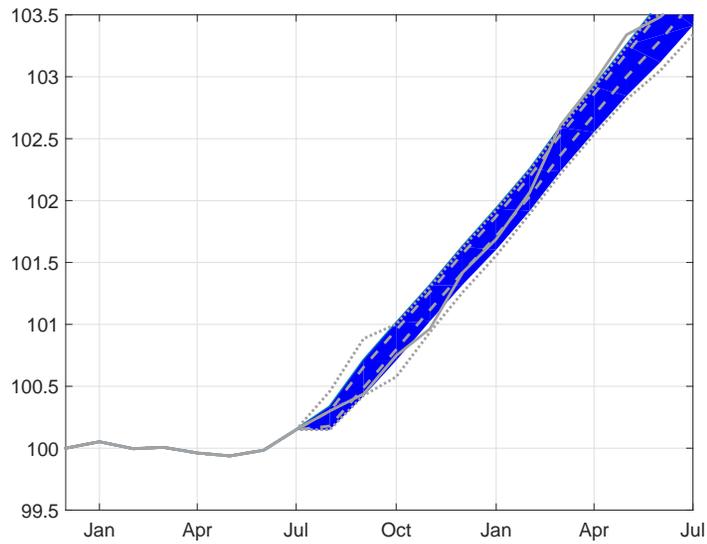(a) Consumer Price Index

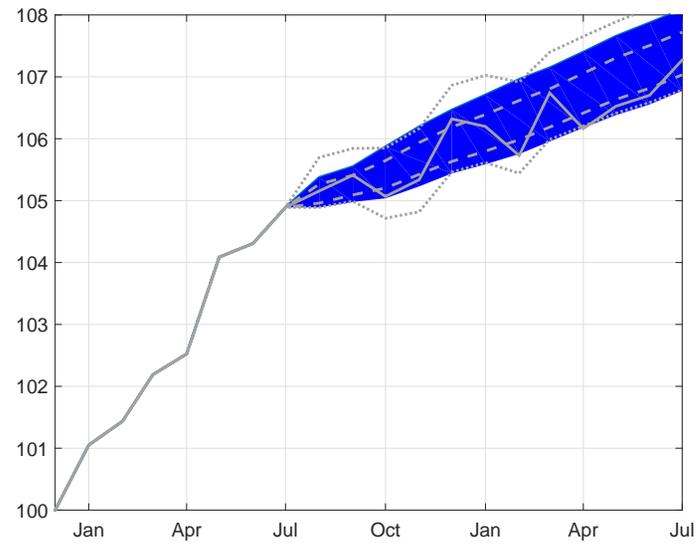(b) Industrial Production Index

(c) 2-year Treasury Bond Rate

(d) Federal Funds Rate

(SHADED, BLUE AREA) Evolution of the Levels CPI, IP, 2yTB, and FF based on our 68% delta method confidence bands for the coefficients of Cumulative Impulse-Response Functions. (GRAY, SOLID LINE) Observed Levels of CPI, IP, 2yTB, and FF from December 2009 to July 2011. Both the CPI index and the IP index were normalized to have a starting value of 100. (GRAY, DASHED LINE) 68% credible set constructed using the priors in Uhlig (2005). (GRAY, DOTTED LINE) Gafarov et al. (2016)'s 68% confidence interval based on the projection approach.

Figure 8: Calibrated Projection Confidence Interval for CPI, IP, 2yTB, FF after the August 2010 announcement
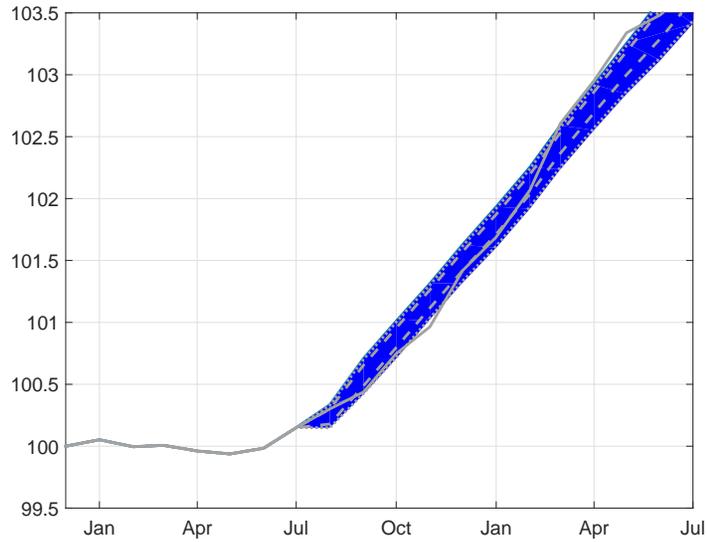
(a) Consumer Price Index

(b) Industrial Production Index

(c) 2-year Treasury Bond Rate

(d) Federal Funds Rate

(SHADED, BLUE AREA) Evolution of the Levels CPI, IP, 2yTB, and FF based on our 68% delta method confidence bands for the coefficients of Cumulative Impulse-Response Functions. (GRAY, SOLID LINE) Observed Levels of CPI, IP, 2yTB, and FF from December 2009 to July 2011. Both the CPI index and the IP index were normalized to have a starting value of 100. (GRAY, DASHED LINE) 68% credible set constructed using the priors in Uhlig (2005). (GRAY, DOTTED LINE) Gafarov et al. (2016)'s 68% *differentiable* projection.
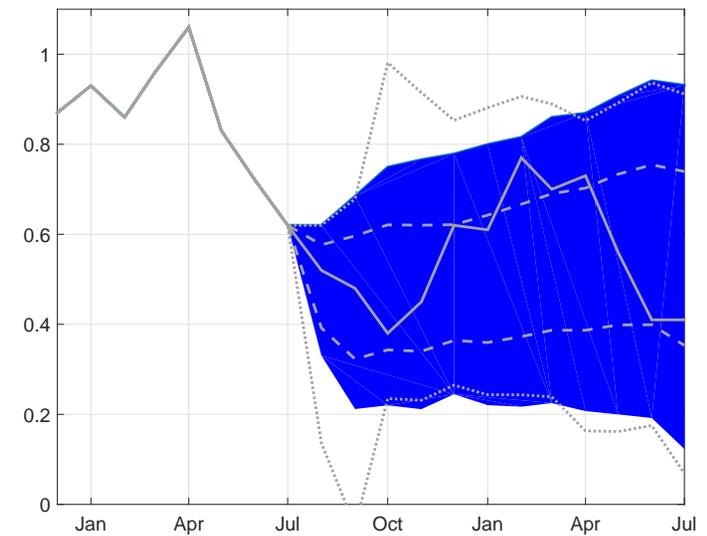
Figure 9: Robust Credible Set for CPI, IP, 2yTB, FF after the August 2010 announcement
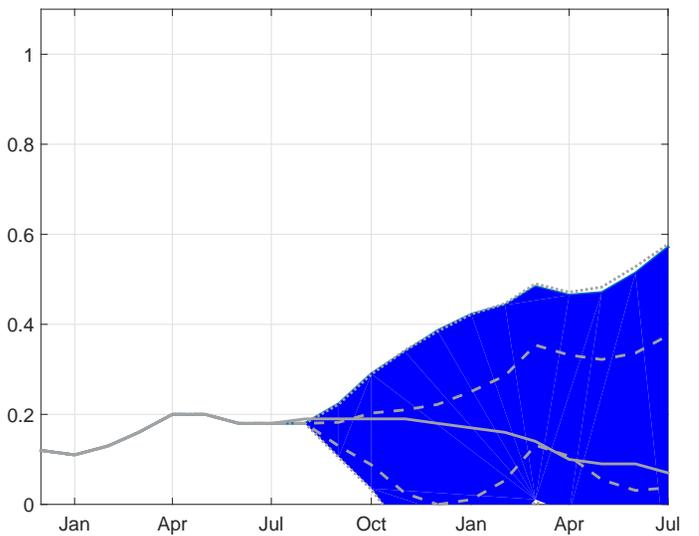

(a) Consumer Price Index

(b) Industrial Production Index

(c) 2-year Treasury Bond Rate

(d) Federal Funds Rate

(SHADED, BLUE AREA) Evolution of the Levels CPI, IP, 2yTB, and FF based on our 68% delta method confidence bands for the coefficients of Cumulative Impulse-Response Functions. (GRAY, SOLID LINE) Observed Levels of CPI, IP, 2yTB, and FF from December 2009 to July 2011. Both the CPI index and the IP index were normalized to have a starting value of 100. (GRAY, DASHED LINE) 68% credible set constructed using the priors in Uhlig (2005). (GRAY, DOTTED LINE) Giacomini and Kitagawa (2014)'s 68% robust-Bayes credible set constructed using the priors for the reduced-form parameters in Uhlig (2005)