Structural Vector Autoregressions: Theory of
Identification and Algorithms for Inference

Juan F. Rubio-Ramírez, Daniel F.Waggoner, and Tao Zha

Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference

Juan F. Rubio-Ramírez, Daniel F.Waggoner, and Tao Zha

Abstract: Structural vector autoregressions (SVARs) are widely used for policy analysis and to provide stylized facts for dynamic general equilibrium models. Yet there have been no workable rank conditions to ascertain whether an SVAR is globally identified. When identifying restrictions such as long-run restrictions are imposed on impulse responses, there have been no efficient algorithms for small-sample estimation and inference. To fill these important gaps in the literature, this paper makes four contributions. First, we establish general rank conditions for global identification of both overidentified and exactly identified models. Second, we show that these conditions can be checked as a simple matrix-filling exercise and that they apply to a wide class of identifying restrictions, including linear and certain nonlinear restrictions. Third, we establish a very simple rank condition for exactly identified models that amounts to a straightforward counting exercise. Fourth, we develop a number of efficient algorithms for small-sample estimation and inference.

JEL classification: C32, E50

Key words: linear and nonlinear restrictions, global identification, almost everywhere, rank conditions, orthogonal rotation, transformation, simultaneity

I.   INTRODUCTION

For the last 36 years after the seminal work of Rothenberg (1971), identification of structural vector autoregressions (SVARs) has remained to be an unresolved theoretical issue. Filling this theoretical gap is of vital importance because impulse responses based on SVARs have been widely used for policy analysis and for providing stylized facts for dynamic stochastic general equilibrium (DSGE) models.[1]

In this paper we present a general theory for global identification of SVARs that applies to both linear restrictions and nonlinear restrictions such as those imposed on impulse responses. Such theory is absent in the existing SVAR literature. In particular, we provide rank conditions for global identification of both identified and exactly identified SVARs. These rank conditions are sufficient for identification and are necessary and sufficient for exact identification. We also show that our rank conditions are easy to implement in practice.

In the existing literature, the rank conditions for identification that come close to being practical are those discussed by Giannini (1992) and Hamilton (1994, pages 332-335). Nonetheless, they apply to local identification only, are mainly designed to analyze linear restrictions on the structural parameters, and can only be numerically verified at a particular point in the parameter space. In contrast, our theory extends the work of Fisher (1966, Chapters 3 and 4) and Hausman and Taylor (1983) about global identification, and our rank conditions apply not only to linear restrictions on structural parameters but also to certain *nonlinear restrictions* on the structural parameters, such as restrictions directly imposed on impulse responses. More important, we establish a powerful result such that if our rank conditions for global identification are satisfied at an arbitrary point in the parameter space, they will be satisfied almost everywhere. This result gives a simple and efficient way to determine whether the model is globally identified in a large parameter space before the estimation step.

Our necessary and sufficient conditions for exact identification complement another part of the existing literature. Rothenberg (1971) gives a necessary condition for exact identification, called "an order condition." Rothenberg (1971)'s order condition is easy to implement by simply counting enough restrictions in total.[2] In

---

[1] See, for example, Galí (1999), Smets and Wouters (2003), Fernández-Villaverde and Rubio-Ramírez (2004), Christiano, Eichenbaum, and Evans (2005), and Sims and Zha (2006a).

[2] Rothenberg (1971) gives sufficient conditions for global identification for certain types of restrictions on simultaneous-equation models, but these are not applicable to non-triangular SVARs. Dhrymes (1978), Hsiao (1983), and Dhrymes (1994), among others, give other rank conditions for traditional simultaneous-equation models. None of these conditions is workable for identification of SVARs, where the structural covariance matrix is restricted to be an identity matrix. See Bekker and Pollock (1986) and Leeper, Sims, and Zha (1996) for further discussions.

practice, this simple counting mechanism has been primarily used to check whether a particular SVAR is exactly identified. Except for a triangular system, however, the model may not be exactly identified even if there are $n(n-1)/2$ linear restrictions where n is the number of endogenous variables. Counterexamples are shown in Sections IV.1 and V.4. Clearly, this is an important issue. If the SVAR under study were incorrectly regarded as being globally identified, the empirical results implied by this model would be misleading (Leeper, Sims, and Zha, 1996; Dufour and Hsiao, forthcoming).

The practical distinction between local identification and global identification is important. In Section IV.1, we highlight this importance through an analysis of a simple simultaneous-equation VAR model studied by Hausman and Taylor (1983) and Sims and Zha (1999). We derive theoretical results for this model and use them to illustrate how a structural model can be locally identified but fail to be globally identified. These results provide an interesting case in which the model is locally identified everywhere but at the same time globally unidentified almost everywhere.

The theoretical results developed in this paper also differ from those for identification of traditional simultaneous-equation models. In traditional simultaneous-equation modeling, there are no restrictions imposed on the correlation between structural disturbances. Therefore, rank conditions for this kind of model do not work for SVARs. The restriction of zero correlation between structural shocks as imposed in the SVAR literature makes it a challenging task to derive workable rank conditions, especially for restrictions imposed on impulse responses. On the other hand, the zero-correlation restriction, as a restriction on the second moment of structural disturbances, helps achieve identification of structural equations that are otherwise unidentifiable. We discuss this important issue in Section IV.2.

Because our theoretical results are new and different from those in standard textbooks, in Section V we illustrate how to apply our theorems to a number of widely-used SVARs in the literature to determine whether these VARs are globally identified. We show that some slight and reasonable changes in restrictions may result in a model that is not globally identified. These examples are also useful to show how easy it is to apply our theory. In all cases, our theorems enable one to determine whether the model is globally identified as a simple matrix-filling exercise.

Once the global identification issue has been resolved, the next task involves small-sample estimation and inference of the model. For this purpose, both classical and Bayesian methods often require repeated simulations of structural parameters. Such computation is quite expensive, especially when time-varying SVARs are estimated (Uhlig, 1997; Canova and Gambetti, 2004; Cogley and Sargent, 2005; Primiceri, 2005; Sims and Zha, 2006b; Gambetti, Pappa, and Canova, forthcoming). To solve this problem, we use our theoretical results to derive efficient algorithms for exactly identified models and for models identified with sign restrictions. Without these new methods it would be prohibitively expensive to obtain accurate small-sample inferences for many relevant problems. These efficient algorithms make it possible to estimate a variety of models with different identifying restrictions, especially when dealing with time-varying features. Without them it would be practically infeasible to entertain a task of estimating a variety of time-varying SVARs and performing the model comparison in their fit to the data, as did Rubio-Ramírez, Waggoner, and Zha (2005).

The rest of the paper is organized as follows. Section II presents a general theory of global identification. Section III derives necessary and sufficient conditions for exact identification. Section IV discusses two important theoretical issues: local versus global identification and differences between identifying a traditional simultaneous-equation model and an SVAR model. Section V shows how to apply our theory to a variety of SVAR models, whose identifiability has not been established before. Section VI uses our theoretical results to derive efficient algorithms for small-sample estimation and inference. Section VII concludes.

II. A Theory of Global Identification

In this section we present a unified, general theory for global identification of SVARs with both linear and certain nonlinear restrictions. In Section II.1, we present a general class of SVARs. In Section II.2, following Rothenberg (1971), we define global and local identifications for this class of models. In Sections II.3 and II.4, we introduce and discuss a wide class of identifying restrictions. These restrictions encompass those in the literature. In Sections II.5 and II.6, we derive rank conditions that are sufficient for global identification.

Our theory is important for several reasons. First, our rank condition is for global identification, while the recent VAR literature deals with local identification.[3] Second, it is very straightforward to check this rank condition by simply counting the number of independent columns of a matrix. Third, our rank condition works for both linear and certain nonlinear restrictions on the structural parameters of the model, while rank conditions in the recent literature deals with local identification or with linear restrictions on the structural parameters. Fourth, and more important, we show that if the model is globally identified at any point of the parameter space, it is then identified for almost all points.

II.1. **The structural model.** The class of SVARs we study has the general form

$$y_t'A_0 = \sum_{\ell=1}^{p} y_{t-\ell}'A_\ell + z_t'C + \varepsilon_t' \text{ for } 1 \leq t \leq T, \tag{1}$$

where

- p is the lag length,
- T is the sample size,
- y_t is an $n \times 1$ vector of endogenous variables,
- z_t is an $n_z \times 1$ vector of exogenous variables,
- ε_t is an $n \times 1$ vector of exogenous structural shocks,
- A_ℓ is an $n \times n$ matrix of parameters for $0 \leq \ell \leq p$, and
- C is a $n_z \times n$ matrix of parameters.

The distribution of ε_t, conditional on the past information, is Gaussian with mean zero and covariance matrix I_n, the $n \times n$ identity matrix.[4] The exogenous variables z_t are of full rank in the sense that the support of z_t spans $\mathbb{R}^{n_z}$. This assumption precludes any co-linear relationship among the exogenous variables. The initial conditions, $y_0, \cdots, y_{1-p}$, are taken as given.

Let

$$A_+' = \begin{bmatrix} A_1' & \cdots & A_p' & C' \end{bmatrix}$$

---

[3]For example, Giannini (1992) works out a rank condition for local identification when restrictions are placed on the contemporaneous impulse responses.

[4]Because only the first and second moments are used for identification, the important assumption is that the reduced-form shocks $u_t' = \varepsilon_t'A_0^{-1}$ form a family of distributions uniquely parameterized by their means and variances.

and

$$x_t' = \begin{bmatrix} y_{t-1}' & \cdots & y_{t-p}' & z_t' \end{bmatrix}$$

for $1 \leq t \leq T$. The dimension of A_+ is $m \times n$, where $m = np + n_z$. The model (1) can be written in compact form as

$$y_t' A_0 = x_t' A_+ + \varepsilon_t'. \tag{2}$$

The parameters of the structural model are (A_0, A_+) and we assume that A_0 is invertible. We denote the set of all structural parameters by $\mathbb{P}^S$. The set $\mathbb{P}^S$ is an open dense subset of $\mathbb{R}^{(n+m)n}$. The reduced-form representation implied by the structural model (2) is

$$y_t' = x_t' B + u_t',$$

where $B = A_+ A_0^{-1}$, $u_t' = \varepsilon_t' A_0^{-1}$, and $E[u_t u_t'] = \Sigma = (A_0 A_0')^{-1}$. The parameters of the reduced-form model are (B, Σ), where Σ is a symmetric and positive definite matrix. We denote the set of all reduced-form parameters by $\mathbb{P}^R$. The set $\mathbb{P}^R$ is an $nm + n(n+1)/2$ dimensional sub-manifold of $\mathbb{R}^{(n+m)n}$ but can be mapped, using the Cholesky decomposition of Σ, to an open subset of $\mathbb{R}^{nm+n(n+1)/2}$. For future reference, define $g : \mathbb{P}^S \to \mathbb{P}^R$ by $g(A_0, A_+) = \left(A_+ A_0^{-1}, (A_0 A_0')^{-1} \right)$.

II.2. **Identification.** We begin by defining when structural parameters are observational equivalent.

Definition 1. Two parameter points, (A_0, A_+) and $(\tilde{A}_0, \tilde{A}_+)$, are observationally equivalent if and only if $g(A_0, A_+) = g(\tilde{A}_0, \tilde{A}_+)$.

Definition 1 implies that two sets of structural parameters are observationally equivalent if they have the same reduced-form representation. We have chosen to define observational equivalence using the relationship between the structural parameters and their reduced-form representation as in Fisher (1966). An alternative approach would be to define observational equivalence using the relationship between the structural parameters and the distribution of the endogenous variables y_t for $1 \leq t \leq T$, as in Rothenberg (1971). These two definitions are equivalent because of our distributional assumption about the exogenous variables z_t and ε_t.

The following theorem gives an equivalent formulation of Definition 1 that is more convenient in the analysis of SVAR models.

Theorem 1. Two parameter points, (A_0, A_+) and $(\tilde{A}_0, \tilde{A}_+)$, are observationally equivalent if and only if there exists an $n \times n$ orthogonal matrix P such that $A_0 = \tilde{A}_0 P$ and $A_+ = \tilde{A}_+ P$.

Proof. If $A_0 = \tilde{A}_0 P$ and $A_+ = \tilde{A}_+ P$ for some orthogonal matrix P, then

$$
\begin{aligned}
g(A_0, A_+) &= \left(A_+ A_0^{-1}, \left(A_0 A_0' \right)^{-1} \right) = \left(\tilde{A}_+ P P^{-1} \tilde{A}_0^{-1}, \left(\tilde{A}_0 P P' \tilde{A}_0' \right)^{-1} \right) \\
&= \left(\tilde{A}_+ \tilde{A}_0^{-1}, \left(\tilde{A}_0 \tilde{A}_0' \right)^{-1} \right) = g(\tilde{A}_0, \tilde{A}_+).
\end{aligned}
$$

Hence the structural parameters are observationally equivalent.

On the other hand, if the structural parameters are observationally equivalent, then

$$
A_+ A_0^{-1} = \tilde{A}_+ \tilde{A}_0^{-1} \text{ and } \left(A_0 A_0' \right)^{-1} = \left(\tilde{A}_0 \tilde{A}_0' \right)^{-1}.
$$

The second equality implies that

$$
\left(\tilde{A}_0^{-1} A_0 \right)' \left(\tilde{A}_0^{-1} A_0 \right) = I,
$$

and therefore $P = \tilde{A}_0^{-1} A_0$ is orthogonal and $A_0 = \tilde{A}_0 P$. This result, together with the fact that $A_+ A_0^{-1} = \tilde{A}_+ \tilde{A}_0^{-1}$, implies that $A_+ = \tilde{A}_+ P$. □

Theorem 1 implies that two sets of structural parameters are observationally equivalent if we can find an orthogonal matrix P that rotates one into the other. This result will be useful in developing both our theory and efficient algorithms. As one might gather from this result, the set of all $n \times n$ orthogonal matrices plays a central role in our analysis. Following the usual convention, we denote the set of all $n \times n$ orthogonal matrices by $O(n)$.

As it is well known, an unrestricted SVAR is neither globally nor locally identified and restrictions are needed for identification. To have as general definitions of global and local identification as possible, we first define a set of restrictions in an abstract way. The class of identifying restrictions considered in this paper will be defined in a more concrete manner later in Section II.3.

Let $R \subset \mathbb{P}^S$ denote the set of all restricted structural parameters. We now define what we mean by global and local identifications of the restricted model.

Definition 2. The parameter point $(A_0, A_+) \in R$ is globally identified if and only if there is no other parameter point $(\tilde{A}_0, \tilde{A}_+) \in R$ that is observationally equivalent to (A_0, A_+).

According to Theorem 1, $(A_0, A_+) \in R$ is globally identified if and only if

$$(A_0 P, A_+ P) \notin R$$

for every orthogonal matrix $P \neq I_n$. Although this paper focuses on global identification, we give the following definition of local identification for a comparison with the existing literature. Let $B_\varepsilon (A_0, A_+)$ denote the open ε-ball centered at (A_0, A_+).

Definition 3. The parameter point $(A_0, A_+) \in R$ is locally identified if and only if there exists an $\varepsilon > 0$ such that no other parameter point $(\tilde{A}_0, \tilde{A}_+) \in R \cap B_\varepsilon (A_0, A_+)$ is observationally equivalent to (A_0, A_+).

By Theorem 1 and Definition 3, $(A_0, A_+) \in R$ is locally identified if and only if $(A_0 P, A_+ P) \notin R$ for every orthogonal matrix $P \neq I_n$ sufficiently close to the identity matrix. Although we have not been specific about the metric used in Definition 3, any metric that delivers the usual topology gives an equivalent result. Thus, any standard metric would be appropriate.

At this point it is important to discuss the notion of admissible parameters as in Rothenberg (1971) and Dhrymes (1978). In some situations, one is not interested in the set of all structural or reduced-form parameter points, but only in a subset that satisfies a priori constraints. A primary example is that Σ must be symmetric and positive definite so that the reduced-form model is always identified (Hsiao, 2001). Another example appears when long-run impulse responses are used to identify the model. In that case, we are interested only in the set of reduced-form parameters for which long-run impulse responses exist.[5] In general, structural or reduced-form parameters are said to be admissible if they satisfy a priori constraints. We denote the set of all admissible reduced-form parameters by $\hat{U}$ and define the set of all admissible structural parameters to be $U = g^{-1}(\hat{U})$. Here, the notation g^{-1} does not denote the inverse function of g, which does not exist, but instead refers to the preimage under g. In this paper, we follow Rothenberg (1971) and assume that $\hat{U}$ is an open subset of $\mathbb{P}^R$ and U is an open subset of $\mathbb{P}^S$. Because $U = g^{-1}(\hat{U})$, if $(A_0, A_+) \in U$, then $(A_0 P, A_+ P) \in U$ for every $P \in O(n)$.

---

[5]Although a long-run impulse response depends on structural parameters, its existence depends only on its reduced form representation. As we shall see in Example 3 in Section II.4, long-run impulse responses exist if and only if the matrix $I_n - \sum_{\ell=1}^{p} B_\ell$ is invertible, where $B_\ell = A_\ell A_0^{-1}$.

II.3. **Identifying restrictions.** In the last subsection we have defined a set of restrictions in an abstract way with the objective of having definitions of global and local identification as general as possible. In this subsection we will be more specific and explicitly define a set of restrictions to be studied in the paper. This set will include a wide class of linear and nonlinear restrictions on structural parameters that is used in the literature. Specifically, we study two important sets of restrictions. The first set concerns the commonly used linear restrictions on the structural parameters (A_0, A_+). This class of restrictions includes the triangular identification as described by Christiano, Eichenbaum, and Evans (1996) and the non-triangular identification as described by Sims (1986), King, Plosser, Stock, and Watson (1991), Gordon and Leeper (1994), Bernanke and Mihov (1998), Zha (1999), and Sims and Zha (2006b).

The second set of restrictions concern nonlinear restrictions on the structural parameters. This class includes restrictions directly imposed on impulse responses, such as short-run and long-run restrictions studied by Blanchard and Quah (1993) and Galí (1992).[6] The restrictions on impulse responses are nonlinear restrictions on the structural parameter space. To determine whether a set of nonlinear restrictions on the structural parameter space identifies the model globally is clearly a difficult task. What is new in this paper is to find a way to transform nonlinear restrictions on the original parameter space to linear restrictions on the transformed parameter space represented by a set of $k \times n$ matrices. Working on the linear restrictions on the transformed parameter space is, in general, a much easier task. The transformation is represented by $f(\cdot)$, as described in the following condition.

Condition 1. Let $U \subset \mathbb{P}^S$ be an open set of admissible structural parameter point and $f(\cdot)$ be a mapping from U to a dense set of $k \times n$ matrices with $f(U)$, where $k \geq 1$. Condition 1 is satisfied if and only if

   (1) for any $P \in O(n)$ and $(A_0, A_+) \in U$, $f(A_0 P, A_+ P) = f(A_0, A_+) P$;

   (2) the function $f(\cdot)$ is continuously differentiable for all $(A_0, A_+) \in U$, and the derivative of $f(\cdot)$ evaluated at (A_0, A_+) is of rank kn.

If the dimension of the transformed space $k > m + n$, then (2) in Condition 1 cannot hold. Therefore, it must be that $1 \leq k \leq m + n$. In practice, the dimension

---

[6]Sign restrictions on impulse responses, as in Faust (1998), Canova and De Nicoló (2002), and Uhlig (2005), are of a different nature and will be analyzed in Section VI.

k is usually a multiple of n. For instance, if the identifying restrictions concern only the contemporaneous matrix A_0, then $k = n$. If identifying restrictions involve the short-run and long-run restrictions of the Galí type (1992), then $k = 2n$.

At this point it is worth making a relevant observation. While our examples and applications concentrate on linear restrictions and on restrictions concerning impulse responses, our global identification theory is valid for linear restrictions on any transformation $f(\cdot)$ of structural parameters that satisfies Condition 1. This advance highlights one of the most salient features of our paper: we develop a rank condition for global identification of an SVAR with linear and nonlinear restrictions on the structural parameters, while most of the recent SVAR literature focuses on local identification with linear restrictions on the structural parameters.

In the previous subsection, we have mentioned the central role orthogonal matrices play in identification of SVAR models. Condition 1 also requires $f(\cdot)$ to respect right multiplication by orthogonal matrices. As one will see, some of the identification problems deal with whether or not certain subsets of possibly nonlinear sub-manifolds of the structural parameter space are of measure zero. The second requirement in Condition 1 allows us to transform this problem to whether or not certain subsets of linear subspaces of the transformed space consisting of $k \times n$ matrices are of measure zero (see Appendix A for a detailed analysis on this issue).

Linear restrictions on the transformed parameters $f(A_0, A_+)$ can be represented by $k \times k$ matrices Q_j for $1 \leq j \leq n$. Each matrix Q_j has rank q_j. The structural parameters (A_0, A_+) satisfy the linear and nonlinear restrictions if and only if

$$Q_j f(A_0, A_+) e_j = 0, \text{ for } 1 \leq j \leq n, \tag{3}$$

where e_j is the jth column of the $n \times n$ identity matrix I_n. The number of restrictions on the jth equation is q_j. Because the ordering of the columns of $f(\cdot)$ is completely arbitrary, we assume without loss of generality that

$$q_1 \geq q_2 \geq \cdots \geq q_n. \tag{4}$$

When one applies the theory developed in this paper, it is important that this convention be followed.

The restrictions given by (3) alone are insufficient to obtain either global or local identification. To see this point, suppose that D is any $n \times n$ diagonal matrix with plus or minus ones along the diagonal. Such matrices are orthogonal. Since

$De_j = \pm e_j$, if (A_0, A_+) satisfy (3), $(A_0 D, A_+ D)$ will also satisfy (3), and thus the system cannot be identified. Consequently, one must employ a normalization rule to determine the sign of each equation, as in standard textbooks (e.g., Dhrymes (1978, p.284) and Greene (1993, p.590)). While the theory developed in this paper works for any choice of normalization, it is worth noting that a poor choice can distort inference concerning impulse responses (Waggoner and Zha, 2003; Hamilton, Waggoner, and Zha, 2007).

We now give a general definition of normalization as follows.

Definition 4. A normalization rule can be characterized by a set $N \subset \mathbb{P}^S$ such that

(1) For any structural parameter point $(A_0, A_+) \in \mathbb{P}^S$, there exists an $n \times n$ diagonal matrix D with plus or minus ones along the diagonal such that $(A_0 D, A_+ D) \in N$.

(2) For any $n \times n$ diagonal matrix $D \neq I_n$ with plus or minus ones along the diagonal, $N \cap ND = \varnothing$, where ND is the set of all elements of N multiplied by the matrix D.

The set N is the collection of normalized structural parameters. The first condition implies that for all structural parameters, the sign of each equation can be chosen so that the normalization rule is satisfied. The second condition implies that this choice is unique. Throughout this paper we assume that all the SVAR models are normalized via some normalization rule N.

We now fully specify the set of restrictions represented by R using the function f, the sets U and N, and the matrices Q_j's:

$$R = \left\{ (A_0, A_+) \in U \cap N \mid Q_j f(A_0, A_+) e_j = 0 \text{ for } 1 \leq j \leq n \right\}, \tag{5}$$

where Q_j is a $k \times k$ matrix of rank q_j with $q_1 \geq \cdots \geq q_n$ and $f(\cdot)$ satisfies Condition 1. From this point on, when we make reference to the set of restrictions represented by R, we refer to the set (5).

II.4. **Examples of transformation.** As discussed in Section II.3, the transformation $f(\cdot)$ allows us to transform difficult nonlinear problems into easier linear problems. In this subsection we show that linear restrictions on (A_0, A_+), as well as short-run

and long-run restrictions on impulse responses, can be represented as linear restrictions on the columns of the transformed parameter matrix $f(A_0, A_+)$. In particular, we use three well-established examples to show how to find the transformation $f(\cdot)$.

Example 1. *Linear restrictions on* (A_0, A_+). For linear restrictions imposed on contemporaneous and lagged coefficients in individual structural equations, the transformation $f(\cdot)$ is simply the identity mapping,

$$f(A_0, A_+) = \begin{bmatrix} A_0 \\ A_+ \end{bmatrix}.$$

If the restrictions concern the contemporaneous coefficient matrix A_0 only, then the transformation $f(\cdot)$ is a projection onto the contemporaneous matrix, and hence $f(A_0, A_+) = A_0$. In either case, the requirement (1) of Condition 1 holds trivially. The transformation is continuously differentiable. Since f is a linear projection for this example, the derivative of f has the required rank.[7] Thus, the requirement (2) of Condition 1 holds as well. Finally, because U is the set of all structural parameters for which A_0 is invertiable, $f(U)$ is the set of all $k \times n$ matrices for which the upper $n \times n$ block is invertiable, which is a dense set.

Example 2. *Short-run restrictions on impulse responses.* The impulse response of the ith variable at horizon h to the jth shock corresponds to the element in row i and column j of

$$L_h = \left(A_0^{-1} J' F^h J \right)', \tag{6}$$

where

$$F = \begin{bmatrix} A_1 A_0^{-1} & I_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{p-1} A_0^{-1} & 0 & \cdots & I_n \\ A_p A_0^{-1} & 0 & \cdots & 0 \end{bmatrix} \quad \text{and } J = \begin{bmatrix} I_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

When $h = 0$, this means that $L_0 = \left(A_0^{-1} \right)'$. If $f(\cdot)$ is the transformation that maps (A_0, A_+) to

$$\begin{bmatrix} L_0' & \cdots & L_p' \end{bmatrix}',$$

---

[7]The derivative of a linear function at any point is a linear function itself. Thus the derivative has the required rank if and only if a projection is an onto function.

it is straightforward to verify that $f(\cdot)$ satisfies the first requirement of Condition 1 and that the transformation is continuously differentiable. As in the first example, $f(U)$ consists of a dense set of all $(n + m) \times n$ matrices for which the first $n \times n$ block is invertible.

To show that the rank of the derivative of f is $(n + m)n$, it suffices to show that f has a differentiable inverse. Let us consider $p = 2$, which the reader can easily generalize. In this case, the transformation $f(\cdot)$ is

$$f(A_0, A_1, A_2) = \left[\begin{array}{ccc} A_0^{-1} & A_0^{-1}A_1A_0^{-1} & A_0^{-1}\left(A_2 + A_1A_0^{-1}A_1\right)A_0^{-1} \end{array} \right]',$$

and its inverse is

$$f^{-1}(X_0, X_1, X_2) = \left[\begin{array}{ccc} X_0^{-1} & X_0^{-1}X_1X_0^{-1} & X_0^{-1}\left(X_2 - X_1X_0^{-1}X_1\right)X_0^{-1} \end{array} \right]',$$

which is differentiable.

Example 3. *Long-run restrictions on impulse responses.* To see how the identification studied by Galí (1992) or Blanchard and Quah (1993) can be represented in our framework, we need a representation of the long-run impulse response function. When the ith variable of the structural model is in first difference, the long-run impulse response of the ith variable to the jth shock is the element in row i and column j of

$$L_\infty = \left(A_0' - \sum_{\ell=1}^{p} A_\ell' \right)^{-1}.$$

Galí and Blanchard and Quah focus on impulse responses at either a short-run horizon (L_0) or the infinite horizon (L_∞) or both. For these cases, the transformation takes one of the following forms:

$$f(A_0, A_+) = L_0, \ f(A_0, A_+) = L_\infty, \ f(A_0, A_+) = \left[\begin{array}{c} L_0 \\ L_\infty \end{array} \right].$$

Note that long-run impulse responses are defined only if the matrix $I_n - \sum_{\ell=1}^{p} B_\ell$, where $B_\ell = A_\ell A_0^{-1}$ is invertible. It can be easily verified that the first requirement of Condition 1 is satisfied and that $f(\cdot)$ is continuously differentiable. Note that $f(U)$ consists of a set of all $k \times n$ matrices for which the first $n \times n$ block is invertible and the second $n \times n$ block, if present, is also invertible. As in Examples 1 and 2, this set is a dense subset of all $k \times n$ matrices.

To see that the derivative of f has the required rank, we let f be a composition of three functions. The first two functions are given by

$$
\begin{bmatrix} A_0 \\ A_+ \end{bmatrix} \rightarrow \begin{bmatrix} A_0' \\ A_0' - \sum_{\ell=1}^{p} A_\ell' \end{bmatrix}, \quad \begin{bmatrix} X \\ Y \end{bmatrix} \rightarrow \begin{bmatrix} X^{-1} \\ Y^{-1} \end{bmatrix},
$$

and the third function is either a projection onto the first $n \times n$ block, a projection onto the second $n \times n$ block, or the identity function. The first and third projections are linear and the second projection is differentiable with a differentiable inverse. Consequently, the composition has the required rank.

Example 4. If one wishes to impose restrictions on (A_0, A_+) jointly with restrictions on impulse responses, one can combine the three transformations described above as long as the rank of the derivative is kn. We would like to re-emphasize that our analysis, not exclusively confined to these three examples of transformations, is valid for linear restrictions on any transformation $f(\cdot)$ of structural parameters that satisfies Condition 1.

II.5. **A rank condition for global identification.** In this subsection we develop a sufficient condition for global identification. This rank condition is very general, and in Section V we show how to apply this condition to a number of widely-used models in the literature.

The following matrix is the key to the establishment of our rank condition. For $1 \leq j \leq n$ and any $k \times n$ matrix X, we define $M_j(X)$ by

$$
\underset{(k+j)\times n}{M_j(X)} = \begin{bmatrix} \underset{k\times k}{Q_j} \underset{k\times n}{X} \\ \begin{bmatrix} \underset{j\times j}{I} & \underset{j\times(n-j)}{\mathbf{0}} \end{bmatrix} \end{bmatrix}.
$$

We now state and prove the following key theorem.

Theorem 2. Consider an SVAR with restrictions represented by R. If $(A_0, A_+) \in R$ and $M_j(f(A_0, A_+))$ is of rank n for $1 \leq j \leq n$, then the SVAR is globally identified at the parameter point (A_0, A_+).

Proof. To prove the theorem, it suffices to show that if the SVAR is not identified at (A_0, A_+), then there exists a j such that $M_j(f(A_0, A_+))$ is of rank strictly less than n. By Theorem 1 and Definition 2, if the SVAR is not identified at (A_0, A_+), then there exists a $P = (p_{i,j}) \in O(n)$ such that $P \neq I_n$ and $(A_0 P, A_+ P) \in R$. Since $P \neq I_n$, let

j be the index of the first column of P that has a non-zero off-diagonal element. We shall show that the rank of $M_j \left(f \left(A_0, A_+ \right) \right)$ is strictly less than n.

Let $q_j = P e_j - p_{j,j} e_j$, where e_j is the jth column of I_n. Since $q_j \neq 0$, it suffices to show $M_j \left(f \left(A_0, A_+ \right) \right) q_j = 0$ to complete the proof. Because both (A_0, A_+) and $(A_0 P, A_+ P)$ are in R, $Q_j f \left(A_0, A_+ \right) q_j = 0$. Thus the upper block of $M_j \left(f \left(A_0, A_+ \right) \right) q_j$ is zero. Because P is orthogonal and j is the index of the first column of P that has a non-zero off-diagonal element, the first $j - 1$ elements of q_j are zero and the jth element of q_j is zero by construction. This result implies that the lower block of $M_j \left(f \left(A_0, A_+ \right) \right) q_j$ is also zero. Thus $M_j \left(f \left(A_0, A_+ \right) \right) q_j = 0$ as required. $\qquad\square$

Since the ordering of columns of $f \left(\cdot \right)$ is arbitrary and the condition in Theorem 2 may be satisfied under one ordering but not under another, one might wish to experiment with all possible orderings when applying the rank condition. Choosing a correct ordering, however, can eliminate unnecessary search and make it more efficient to check the rank condition. We find that our convention of ordering the columns of $f \left(\cdot \right)$ so that $q_1 \geq \cdots \geq q_n$ is, in general, sufficient to ascertain whether there exists a set of structural parameters (A_0, A_+) such that $M_j \left(f \left(A_0, A_+ \right) \right)$ is of rank n for $1 \leq j \leq n$.

In contrast to the well-established rank conditions of Giannini (1992) and Hamilton (1994) for local identification, Theorem 2 establishes a rank condition for global identification. Thus, even though an SVAR is locally identified according to Giannini (1992) and Hamilton (1994, pages 332-335), it may not be identified globally (as will be shown in Section IV.1).

In the case of linear restrictions on structural parameters, our rank condition is closely related to Fisher (1966, chapter 4) and Hausman and Taylor (1983). These earlier works focused on identification of one equation at a time and did not explicitly derive a workable condition for global identification of the whole system. Our condition not only encompasses and unifies various sufficient conditions provided by Fisher (1966, chapter 4) and Hausman and Taylor (1983), but also is much easier to implement than the previous conditions.

It is important to note that our theory of global identification applies to a large class of nonlinear restrictions on the structural parameters, while most of the earlier work provides sufficient conditions only for linear restrictions (Fisher, 1966; Hausman and Taylor, 1983; Hamilton, 1994).

II.6. **Global identification almost everywhere.** In the existing literature, the rank conditions for local identification are (numerically) checked at a particular parameter point. Such a point is typically chosen at the estimate of the model parameters. Often it is important to know whether the model is identified at different points in the parameter space prior to the estimation step. In this subsection we develop two theorems to answer this question. We begin with the following key definition.

Define the set K by

$$K = \left\{ (A_0, A_+) \in R \mid \text{rank} \left(M_j \left(f \left(A_0, A_+ \right) \right) \right) = n \text{ for } 1 \leq j \leq n \right\} \tag{7}$$

According to Theorem 2, the model is globally identified on the set K. The next theorem states that this set is open.

Theorem 3. The set K is open.

Proof. The function from the set of all $(k + j) \times n$ matrices to $\mathbb{R}$, which maps a matrix to the volume of the parallelepiped spanned by its columns, is continuous, and a $(k + j) \times n$ matrix is of rank n if and only if the volume of the parallelepiped spanned by it columns is non-zero. This result implies that the set of all $(k + j) \times n$ matrices of rank n is open, and since $f(\cdot)$ is continuous, the set of all $(A_0, A_+) \in R$ such that $M_j \left(f \left(A_0, A_+ \right) \right)$ is of rank n will also be open. $\square$

Theorem 3 is important for the following reasons. If the structural parameter point $(A_0, A_+) \in R$ satisfies the rank condition, then there exists a neighborhood around (A_0, A_+) such that all the structural parameters within that neighborhood satisfy the rank condition. The implication of this result is that if the model is globally identified at the estimated value of the structural parameters, there is no need to check if it is globally identified at nearby points. The next theorem, building on Theorem 3, gives an even stronger result: if the model is globally identified at any point in the structural parameter space, the model is, in fact, globally identified almost everywhere.

Theorem 4. Either K is empty or the complement of K in R is of measure zero in R.

Proof. The proof is provided in Appendix B. $\square$

This theorem is powerful because it gives a practical and efficient way of checking whether the model is globally identified almost everywhere prior to the estimation step. It follows from this theorem that one can randomly choose an element of R

and then check the rank condition. If the rank condition is satisfied, we know that the model is globally identified almost everywhere. In particular, we do not need to perform an brute-force search to determine if the model is globally identified at different points in the parameter space. This result makes our rank condition both powerful and extremely easy to apply.

III. EXACT IDENTIFICATION

In the last section we study globally identified models that includes overidentified cases. Since much of the SVAR literature involves exactly identified models, we show in this section how the sufficient condition for global identification described in Section II.5 becomes a necessary and sufficient condition for exact identification.

One familiar class of exactly identified SVARs is that linear restrictions on A_0 are triangular as defined below.

Definition 5. Let the transformation $f(\cdot)$ be given by $f(A_0, A_+) = A_0$ so that the SVAR is identified via linear restrictions on the contemporaneous matrix A_0. The restrictions on A_0 are said to be triangular if and only if there exists an invertible matrix P_1 such that the matrix $P_1 f(A_0, A_+)$ is triangular.

The intuitive interpretation of Definition 5 is that identifying restrictions A_0 are triangular if A_0 can be transformed into a triangular matrix. The recursive identification of Christiano, Eichenbaum, and Evans (1996) is a classic example of triangular restrictions on A_0. Because there are exactly identified SVARs in the SVAR literature that have non-triangular restrictions, we now give a precise definition of exact identification. Our definition differs slightly from Hamilton (1994, page 250)'s definition and, as we shall see below, the difference is crucial to understanding SVARs with non-triangular restrictions on A_0.

Definition 6. Consider an SVAR with restrictions represented by R. The SVAR is said to be exactly identified if, for almost any admissible reduced-form parameter point (B, Σ), there exists a unique structural parameter point $(A_0, A_+) \in R$ such that $g(A_0, A_+) = (B, \Sigma)$.

According to Definition 6, an SVAR is said to be exactly identified if, for almost any point in the admissible reduced-form parameter space, there exists a unique

set of structural parameters that implies these reduced-form parameters. In contrast, Hamilton (1994, page 250) defines exact identification by insisting that for any (not just for *almost any*) point in the admissible reduced-form parameter space, there exists a unique set of structural parameters that implies these reduced-form parameters. As stated in the following theorem, it turns out that Hamilton's definition precludes all SVAR with non-triangular linear restrictions on A_0 from being exactly identified.

Theorem 5. Let the transformation $f(\cdot)$ be given by $f(A_0, A_+) = A_0$ so that the SVAR is identified via linear restrictions on the contemporaneous matrix A_0. If, for *every* reduced-form parameter point (B, Σ), there exists a unique structural parameter point $(A_0, A_+) \in R$ such that $g(A_0, A_+) = (B, \Sigma)$, then the restrictions on A_0 must be triangular.

Proof. The proof is provided in Appendix D. $\square$

Theorem 5 is important because it implies that Definition 6, not Hamilton (1994)'s original definition, is needed to allow for the possibility of non-triangular models to be exactly identified, such as the simultaneous-equation model studied by Hamilton (1994, pages 332-335) and other non-triangular exactly identified models in the SVAR literature.

While Definition 6 deals with the reduced form parameters, it is often useful to work with an equivalent formulation of exact identification in the form of structural parameters. The following theorem gives this formulation.

Theorem 6. Consider an SVAR with restrictions represented by R. The SVAR is exactly identified if and only if, for almost every structural parameter point $(A_0, A_+) \in U$, there exists a unique matrix $P \in O(n)$ such that $(A_0 P, A_+ P) \in R$.

Proof. Let $\hat{G}$ be

$$\{(B, \Sigma) \in \hat{U} \mid \text{there is not a unique } (A_0, A_+) \in R \text{ s.t. } g(A_0, A_+) = (B, \Sigma)\} \quad (8)$$

and let G be

$$\{(A_0, A_+) \in U \mid \text{there is not a unique } P \in O(n) \text{ s.t. } (A_0 P, A_+ P) \in R\}. \quad (9)$$

Note that $G = g^{-1}(\hat{G})$. Definition 6 states that an SVAR model is exactly identified if and only if $\hat{G}$ is of measure zero, and Theorem 6 states that an SVAR is exactly

identified if and only if G is of measure zero. Consider the function from $\mathbb{P}^R \times O(n)$ to $\mathbb{P}^S$ to which maps $(B, \Sigma) \times P$ to $\left(TP, BP'T^{-1}\right)$, where T is the unique lower triangular matrix with positive diagonal such that $TT' = \Sigma^{-1}$. The matrix T can be obtained from the Cholesky decomposition of Σ^{-1}. This function and its inverse are continuously differentiable; $\hat{G} \times O(n)$ maps to G. Thus it follows that G is of measure zero if and only $\hat{G}$ is of measure zero. □

Theorem 6 states that if an SVAR is exactly identified, there exists a unique orthogonal matrix P such that (A_0P, A_+P) satisfies the identifying restrictions for almost any value of unrestricted structural parameters (A_0, A_+). This result is vital to finding efficient algorithms for small-sample estimation and inference of an exactly identified model, as shown in Section VI.

III.1. **Rank conditions for exact identification.** We are now ready to prove the rank conditions for exact identification. The well known order condition for exact identification implied by Rothenberg (1971) states that the total number of restrictions must be equal to $n(n-1)/2$. This necessary condition is commonly used in the SVAR literature. In what follows, we prove that if Rothenberg (1971)'s order condition is satisfied for exact identification, the sufficient condition in Theorem 2 becomes necessary as well.

Theorem 7. Consider an SVAR with restrictions represented by R. The SVAR is exactly identified if and only if the total number of restrictions is equal to $n(n-1)/2$ and the rank condition in Theorem 2 is satisfied.

Proof. The proof is provided in Appendix C. □

Theorem 7 not only implies that our definition of exact identification, Definition 6, is consistent with the traditional definition (such as Rothenberg's order condition), but it also gives us a checkable necessary and sufficient condition for exact identification. We now show that there is a more powerful condition, one that does not even involve checking the rank of any matrix but requires checking only whether the numbers of restrictions satisfy an appropriate order. This important result is stated in the following theorem.

Theorem 8. Consider an SVAR with restrictions represented by R. The SVAR is exactly identified if and only if $q_i = n - j$ for $1 \leq j \leq n$.

Proof. The proof is provided in Appendix C.   □

Note that the rank condition in Theorem 8 is a simple counting exercise. The necessary condition of Rothenberg (1971) simply counts the total number of restrictions. Our necessary and sufficient condition not only counts the number of restrictions but also requires that the restrictions follow a certain pattern equation by equation. As much of the existing literature deals with exactly identified SVARs, the simple counting condition in Theorem 8 applies to a wide range of identifications, including the nonlinear restrictions on (A_0, A_+) described in Section II.4.

In addition to the powerful rank condition for exact identification, Theorem 8 forms the basis for efficient Bayesian and classical small-sample methods for estimation and inference. For an exactly identified SVAR, Theorem 6 states that there exists an unique orthogonal matrix P such that $(A_0 P, A_+ P)$ satisfies the restrictions for almost every unrestricted structural parameter point (A_0, A_+). This result gives us a practical way to find the set of structural parameters that satisfy the identifying restrictions, if one is able to draw unrestricted structural parameters or reduced-form parameters. For each draw of unrestricted structural parameters, one needs to find only an orthogonal matrix P that rotates the unrestricted draw to the one that satisfies the restrictions.[8] If the original draw is for the reduced-form parameters (B, Σ), one can rotate the Cholesky decomposition of Σ to get a draw that satisfies the restrictions. The difficulty in this whole procedure is to build an efficient algorithm to find the needed orthogonal matrix P. As will be shown in Section VI, such an algorithm can be found. We will also show that for systems in which the restrictions on $f(\cdot)$ can be permuted to be triangular, there is an even faster algorithm.

Why is this result important? Take as an example an SVAR with restrictions on impulse responses that is exactly identified. The existing methods in the literature typically solve a system of nonlinear equations. This traditional approach becomes very inefficient if a large number of simulations is required to obtain accurate results for small-sample estimation and inference. When time-varying SVARs are studied (Canova and Gambetti, 2004; Cogley and Sargent, 2005; Primiceri, 2005; Sims and Zha, 2006b), the traditional approach is practically infeasible because the system of

---

[8]This procedure applies to the maximum likelihood estimation as well. One first obtains a maximum likelihood estimate of the parameters in an unidentified system and then uses P to rotate these parameters to get the estimate of the structural parameters that satisfies the identifying restrictions.

nonlinear equations would have to be solved for each possible state of the structural parameters. In contrast, Theorem 8 implies that the restrictions for exactly identified systems have a very special structure. This structure can be exploited to obtain an efficient method for finding the orthogonal matrix P, which in turn leads to efficient small-sample simulation techniques, as will be described in Section VI.

IV. TWO THEORETICAL ISSUES

In this section we address two important theoretical issues. The first issue concerns the difference between global identification and local identification. We consider a model that is locally identified but not globally identified. This model is particularly insightful because it shows that the rank condition in the existing literature for local identification does not provide any guidance as to whether or not the model is globally identified.

The second issue highlights how the identification theory developed in the traditional simultaneous-equation literature cannot be applied to identification of an SVAR. In particular, we show the restrictions on the covariance matrix of structural shocks help identify Hamilton (1994)'s demand-supply model in which the supply equation is not identified in the traditional simultaneous equation framework. Using the theory developed in Sections II and III, we show that the same supply equation is globally identified in the SVAR framework.

IV.1. **Local vs. global identification.** To illustrate how an SVAR can be locally but not globally identified, we consider the following three-variable example taken directly from Sims and Zha (1999) and Fubac, Waggoner, and Zha (2007):[9]

$$A_0 = \begin{bmatrix} a_{11} & 0 & a_{13} \\ a_{21} & a_{22} & 0 \\ 0 & a_{32} & a_{33} \end{bmatrix}. \tag{10}$$

This simple model has no lags, and there is only one restriction on each equation such that $q_1 = q_2 = q_3 = 1$. The model satisfies Rothenberg (1971)'s order condition that the total number of restrictions equals $n(n-1)/2 = 3$, and using the procedure outlined in Hamilton (1994, pages 332-335), we shall show that at a particular

---

[9]Fubac, Waggoner, and Zha (2007) show that this contemporaneous SVAR model can be derived from the three-variable forward-looking New-Keynesian model studied by Cochrane (2006).

parameter point the model is locally identified.[10] According to Theorem 8, however, this model is not exactly identified. Moreover, a deeper analysis of this simple model reveals that the space of reduced form parameters divides into three sets, a set of positive measure on which the model is locally identified but not globally identified, a set of positive measure on which there is no representation of structural parameters satisfying the restrictions implied by (10), and a set of measure zero on which the model is globally identified. This example is instructive because it outlines what can go wrong when the model is locally identified but fails to be globally identified.

To apply Hamilton (1994, pages 332-335)'s procedure, we first need to transform the matrix (10) into Hamilton's notation. This re-parameterization involves two matrices, B_0 and D, given by

$$B_0 = \begin{bmatrix} 1 & a_{21}/a_{11} & 0 \\ 0 & 1 & a_{32}/a_{22} \\ a_{13}/a_{33} & 0 & 1 \end{bmatrix} \text{ and } D = \begin{bmatrix} a_{11}^{-2} & 0 & 0 \\ 0 & a_{22}^{-2} & 0 \\ 0 & 0 & a_{33}^{-2} \end{bmatrix}. \tag{11}$$

For the model to be locally identified, the matrix

$$J = \begin{bmatrix} -2D_n^+(\Omega \otimes B_0^{-1})S_b & D_n^+(B_0^{-1} \otimes B_0^{-1})S_d \end{bmatrix}$$

must be of full column rank, where

$$S_b = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, S_d = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, (D_n^+)' = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 1 \end{bmatrix}$$

and $\Omega = B_0^{-1}D(B_0^{-1})'$. Even in this simple example, it is difficult to determine analytically whether the matrix J is of full column rank, but given any particular choice of the parameter point, it is easy to determine the rank of J using any numerical

---

[10]Alternatively, one can use Rothenberg (1971)'s information matrix to check if the model is locally identified, as suggested by Sargent (1976) and recently employed by Iskrev (2007).

linear-algebra package such as Matlab or Gauss. For instance, if $a_{11} = a_{22} = a_{33} = 1$ and $a_{13} = a_{21} = a_{32} = 2$, then J is of rank 6, and thus the model is locally identified. The model is not, however, globally identified at that point. To see why, consider the orthogonal matrix P

$$P = \begin{bmatrix} 2/3 & 1/3 & -1/3 \\ -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \end{bmatrix}.$$

It is straightforward to show that

$$\tilde{A}_0 = A_0 P = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

is observationally equivalent because $\tilde{A}_0$ satisfies the restrictions.

Is this choice of the parameter point so special that the model may still be globally identified at other parameter points? Given the high dimension of this model, this question cannot be answered by any numerical procedure. To answer this question, we prove below that all the reduced-form parameters can be grouped into three distinct sets. The first set has a positive measure on which none of the elements can be represented by structural parameters satisfying the restrictions. The second set also has a positive measure on which every element has two structural representations that satisfy the restrictions. The third set has measure zero on which every element has a unique structural representation that satisfies the restrictions.

To prove these results, we begin by decomposing the reduced-form covariance matrix as $\Sigma = C'C$, where C is an upper triangular matrix

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & c_{33} \end{bmatrix}.$$

Define b_1, d_1, b_2, and d_2 by

$$\begin{aligned} b_1 &= 2c_{13}^2 c_{22}^2 c_{33}^2 - 2c_{12}c_{13}c_{23}c_{22}^2 c_{33}^2 + c_{22}^2 c_{33}^4, \\ d_1 &= (c_{13}c_{22} - c_{12}c_{23})^2 c_{22}^2 c_{33}^4 \left(c_{13}^2 + c_{23}^2 + c_{33}^2 \right), \\ b_2 &= 2c_{23}^2 c_{22}^2 c_{33}^2 + 2c_{12}c_{13}c_{23}c_{22}^2 c_{33}^2 + c_{22}^2 c_{33}^4, \\ d_2 &= c_{23}^2 \left(c_{12}^2 + c_{22}^2 \right) c_{22}^2 c_{33}^4 \left(c_{13}^2 + c_{23}^2 + c_{33}^2 \right). \end{aligned}$$

It can be verified that $b_1^2 - 4d_1 = b_2^2 - 4d_2$.

Proposition 1. Only one of the three following cases can occur:

(1) If $b_1 < 0$ or $b_2 < 0$ or $b_1^2 - 4d_1 = b_2^2 - 4d_2 < 0$, there exists no matrix A_0 that satisfies the restrictions and $C'C = \Sigma = (A_0 A_0')^{-1}$.

(2) If $b_1 \geq 0$ and $b_2 \geq 0$ and $b_1^2 - 4d_1 = b_2^2 - 4d_2 > 0$, there are exactly two matrices, A_0 and $\tilde{A}_0$, that satisfy the restrictions and $C'C = \Sigma = (A_0 A_0')^{-1} = (\tilde{A}_0 \tilde{A}_0')^{-1}$.

(3) If $b_1 \geq 0$ and $b_2 \geq 0$ and $b_1^2 - 4d_1 = b_2^2 - 4d_2 = 0$, there exists a unique matrix A_0 that satisfies the restrictions and $C'C = \Sigma = (A_0 A_0')^{-1}$.

Proof. If $(A_0 A_0')^{-1} = \Sigma = C'C$, then $(CA_0)' CA_0 = I$, which implies that CA_0 is orthogonal. But

$$CA_0 = \begin{bmatrix} a_{11}c_{11} + a_{21}c_{12} & a_{22}c_{12} + a_{32}c_{13} & a_{13}c_{11} + a_{33}c_{13} \\ a_{21}c_{22} & a_{22}c_{22} + a_{32}c_{23} & a_{33}c_{23} \\ 0 & a_{32}c_{33} & a_{33}c_{33} \end{bmatrix},$$

while any orthogonal matrix with a zero in the first column and third row must be of the form

$$\begin{bmatrix} \gamma & -\alpha\lambda & \beta\lambda \\ \lambda & \alpha\gamma & -\beta\gamma \\ 0 & \beta & \alpha \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \gamma & \alpha\lambda & -\beta\lambda \\ \lambda & -\alpha\gamma & \beta\gamma \\ 0 & \beta & \alpha \end{bmatrix},$$

where $\alpha^2 + \beta^2 = 1$ and $\gamma^2 + \lambda^2 = 1$. The first column and third row of these two representations of CA_0 gives us expressions for α, β, γ, and λ. The upper right hand block of these representations give us equations that we can solve for a_{11}, a_{13}, a_{21}, and a_{22} to obtain

$$a_{11} = \pm \frac{-a_{32}^2 c_{12} c_{13} c_{22} + a_{32}^2 c_{12}^2 c_{23} + a_{33}^2 c_{12}^2 c_{23} + a_{33}^2 c_{22}^2 c_{23}}{a_{32} a_{33} c_{11} c_{22}^2 c_{33}},$$

$$a_{13} = -\frac{a_{32}^2 c_{13} c_{22} + a_{33}^2 c_{13} c_{22} - a_{32}^2 c_{12} c_{23} - a_{33}^2 c_{12} c_{23}}{a_{33} c_{11} c_{22}},$$

$$a_{21} = \pm \frac{a_{32}^2 c_{13} c_{22} - a_{32}^2 c_{12} c_{23} - a_{33}^2 c_{12} c_{23}}{a_{32} a_{33} c_{22}^2 c_{33}},$$

$$a_{22} = -\frac{a_{32}^2 c_{23} + a_{33}^2 c_{23}}{a_{32} c_{22}},$$

where the signs of a_{11} and a_{21} are chosen to be consistent with the normalization rule. Substituting all of this into the equations $\alpha^2 + \beta^2 = 1$ and $\gamma^2 + \lambda^2 = 1$ and

simplifying, gives one linear and one quadratic equation in terms of a_{32} and a_{33}. These can be solved to obtain

$$a_{33}^2 = \frac{b_1 \pm \sqrt{b_1^2 - 4d_1}}{2c_{22}^2 c_{33}^4 \left(c_{13}^2 + c_{23}^2 + c_{33}^2\right)},$$

$$a_{32}^2 = \frac{b_2 \mp \sqrt{b_2^2 - 4d_2}}{2c_{22}^2 c_{33}^4 \left(c_{13}^2 + c_{23}^2 + c_{33}^2\right)}.$$

Because both a_{33} and a_{32} must be real, we must have that $b_1 \geq 0$, $b_2 \geq 0$, and $b_1^2 - 4d_1 = b_2^2 - 4d_2 \geq 0$ in order for a solution to exist. There will be only one solution if $b_1^2 - 4d_1 = b_2^2 - 4d_2 = 0$ and two solutions if $b_1^2 - 4d_1 = b_2^2 - 4d_2 > 0$. $\quad\square$

Clearly, every structural parameter point is locally identified. On the other hand, the set of structural parameters that are globally identified is of measure zero because the constraint $b_1^2 - 4d_1 = b_2^2 - 4d_2 = 0$ must be met. This powerful example shows how a structural model can be locally identified but fail to be globally identified and highlights practical distinctions between local identification and global identification.

IV.2. **Identifying supply and demand.** We use Hamilton (1994, Sections 9.1 and 11.6)'s supply-demand model of the orange market as a study case to highlight the prominent role of restrictions on the covariance matrix of structural disturbances in achieving identification of an SVAR. Let p_t be the log of the price of the good of interest (oranges), q_t be the log of the quantity of such a good, and w_t indicate an exogenous variable (weather) that affects the supply of oranges. For expository illustration, we analyze the following Hamilton model with no lag:

$$a_{0,31} w_t = \varepsilon_t^w, \quad \text{Weather} \tag{12}$$

$$a_{0,12} q_t + a_{0,22} p_t = \varepsilon_t^d, \quad \text{Demand} \tag{13}$$

$$a_{0,13} q_t + a_{0,23} p_t + a_{0,33} w_t = \varepsilon_t^s, \quad \text{Supply} \tag{14}$$

where $-a_{0,22}/a_{0,12} < 0$ (a negatively sloped demand curve) and $-a_{0,23}/a_{0,13} > 0$ (a positively sloped supply curve).

In the traditional simultaneous-equation framework where no restrictions are imposed on the covariance matrix of the structural disturbances ε_t^d and ε_t^s, it is well known that the demand equation is identified. The instrumental variable w_t shifts the supply curve but not the demand curve. The identification of such a demand

equation is a textbook example. As also discussed in most textbooks, however, the supply equation cannot be identified. The reason is that there is no instrumental variable to shift the demand schedule so as to trace out the slope of the supply curve. Another way to understand this result is to replace the supply equation with an arbitrary linear combination of the demand and supply equations, while leaving the demand equation unchanged. This replacement does not violate any restrictions and thus the supply equation is not identified.

In the SVAR framework, such a replacement is no longer harmless. An arbitrary linear combination of the supply and demand equations will not preserve the lack of correlation between the supply and demand disturbances ε_t^d and ε_t^s. Only an orthonormal linear transformation can preserve the orthogonal nature of these structural shocks. Does there exist a unique orthogonal matrix P such that $A_0 P$ satisfy the restrictions implied by (12) - (14)? If so, both the demand and supply equations are exactly identified.

To study this issue, we express this simultaneous-equation model in the form of (1) as

$$A_0 = \begin{bmatrix} 0 & a_{0,12} & a_{0,13} \\ 0 & a_{0,22} & a_{0,23} \\ a_{0,31} & 0 & a_{0,33} \end{bmatrix}. \tag{15}$$

For the identification represented by (15), $n = 3$ and there are a total of three restrictions ($= n(n-1)/2$), satisfying Rothenberg (1971)'s order condition for exact identification. While one can use the rank condition given by Hamilton (1994, page 332-335) to determine whether the model is identified locally around a given point in the parameter space, the question is whether the model is globally identified.

To answer this question, we apply Theorem 8 by writing (15) in the form of (3). The transformation is

$$f(A_0, A_+) = A_0,$$

and the restrictions can be represented by

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Q_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We have ordered the equations so that $q_1 \geq q_2 \geq q_3$, where $q_1 = 2$, $q_2 = 1$, and $q_3 = 0$. Since $q_j = n - j$, it follows directly from Theorem 8 that the non-triangular SVAR represented by (15) is exactly identified. Of course, there are points at which neither

the demand nor the supply equation is identified. As discussed in Hamilton (1994, pages 332-335), the model is not identified at the parameter values with $a_{0,33} = 0$. The set G, defined by (9), contains all such locally unidentified points but has measure zero according to Theorems 6 and 8.

In comparison with the traditional simultaneous-equation framework discussed at the beginning of this subsection, the linear restrictions represented by (15), together with the usual SVAR restrictions that ε_t's are uncorrelated as a second-moment condition, amount to the traditional simultaneous-equation assumption that one of the two variables p_t and q_t in the demand-supply system is predetermined. In other words, identification of the demand equation is equivalent to identifying the demand shock ε_t^d. Since ε_t^d and ε_t^s are uncorrelated, the demand shock can be used as a "shifter" to move the demand curve up and down along the supply curve so as to achieve identification of the supply equation (Hausman and Taylor, 1983).

V. APPLICATION

The theory presented in Sections II and III establishes easy-to-check rank conditions for global identification and for exact identification. Section IV illustrates the importance of using our rank conditions for global identification instead of the existing conditions for local identification and explains why our theoretical results differ from those used for traditional simultaneous-equation models.

Since our theoretical results, particularly about restrictions on impulse responses, are newly developed, it is both essential and instructive to show the reader how our theory can be applied in practice. In this section, we demonstrate how to apply our theory to a number of existing SVARs studied in the literature. For almost all these models, global identification has not been formally established. For the examples of nonlinear restrictions on (A_0, A_+), we show how to use $f()$ to transform these nonlinear restrictions to linear ones in the transformed parameter space.

V.1. **Triangular SVARs.** If restrictions on A_0 are triangular, as in Eichenbaum and Evans (1995) and Christiano, Eichenbaum, and Evans (1996), the model is (obviously) exactly identified. Applying Theorem 8 becomes trivial.

V.2. **Identification through the lag structure.** Restrictions on the lag structure can be used to aid identification. As an illustration, we revisit the supply-demand model in Section IV.2 with the additional restriction $a_{0,33} = 0$. In this case, the model is not

identified either globally or locally because the second and third equations cannot be distinguished. Global identification, however, can be restored through the lag structure. To see this point, we expand the system to include A_1. The restrictions for this one-lag model can be represented as

$$A_0 = \begin{bmatrix} 0 & a_{0,12} & a_{0,13} \\ 0 & a_{0,22} & a_{0,23} \\ a_{0,31} & 0 & 0 \end{bmatrix}, \text{ and } A_1 = \begin{bmatrix} 0 & a_{1,12} & a_{1,13} \\ 0 & a_{1,22} & a_{1,23} \\ a_{1,31} & 0 & a_{1,33} \end{bmatrix}, \quad (16)$$

and the transformation is

$$f(A_0, A_+) = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}. \quad (17)$$

For this model, Q_j's for $j = 1, 2, 3$ are

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } Q_3 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since $a_{1,33} \neq 0$, it follows from the Q_j's that $q_1 = 4$, $q_2 = 2$, and $q_3 = 1$, and thus the total number of restrictions $(\sum_{j=1}^3 q_j)$ is 7, greater than $n(n-1)/2 = 3$. Even though Rothenberg (1971)'s order condition is satisfied, the model may or may not be overidentified. To use our theory to determine the global identifiability of this model, we fill the matrices $M_j(f(A_0, A_+))$ for $j = 1, 2, 3$ according to (16) and (17), and these matrices are

$$\begin{bmatrix} 0 & a_{0,12} & a_{0,13} \\ 0 & a_{0,22} & a_{0,23} \\ 0 & a_{1,12} & a_{1,13} \\ 0 & a_{1,22} & a_{1,23} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} a_{0,31} & 0 & 0 \\ a_{1,31} & 0 & a_{1,33} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \text{ and } \begin{bmatrix} a_{0,31} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (18)$$

It is clear from (18) that even with the restriction $a_{0,33} = 0$, $M_j(f(A_0, A_+))$ has rank 3 for some values of a's for $j = 1, 2, 3$. Hence, even if $a_{0,33} = 0$, the model is globally identified almost everywhere according to Theorems 2 and 4.

V.3. **A monetary SVAR.** To identify the systematic monetary policy behavior, Sims (1986), Gordon and Leeper (1994), Bernanke and Mihov (1998), Leeper and Zha (2003), and Sims and Zha (2006b), among others, propose identifying restrictions along the same line of Hamilton's simultaneous-equation model discussed in Section V.2. This approach focuses directly on an interpretation of the structural equations themselves. In particular, they separate the monetary policy equation from the money demand equation and other non-policy equations. The restrictions require non-triangular relationships between financial variables such as the interest rate and money. The following A_0 gives a particular example of restrictions on the contemporaneous coefficients only:[11]

$$
A_0 = \begin{array}{c} \\ \log Y \\ \log P \\ R \\ \log M \\ \log P_c \end{array}
\begin{array}{c}
\begin{array}{ccccc} \text{PS} & \text{PS} & \text{MP} & \text{MD} & \text{Inf} \end{array} \\
\begin{bmatrix}
a_{11} & a_{12} & 0 & a_{14} & a_{15} \\
0 & a_{22} & 0 & a_{24} & a_{25} \\
0 & 0 & a_{33} & a_{34} & a_{35} \\
0 & 0 & a_{43} & a_{44} & a_{45} \\
0 & 0 & 0 & 0 & a_{55}
\end{bmatrix}
\end{array}, \tag{19}
$$

where the transformation function for this case is

$$
f(A_0, A_+) = A_0.
$$

The five variables in the model are: log GDP ($\log Y$), log GDP deflator ($\log P$), the nominal short-term interest rate (R), log M3 ($\log M$), and log commodity prices ($\log P_c$). The monetary policy (MP) column in (19) represents a central bank's contemporaneous behavior, the information (Inf) column describes the commodity (information) market, the MD column corresponds to the money demand equation, and the block consisting of the first two columns represents the production sector (PS), whose variables are arbitrarily ordered to be upper triangular. For this model, we have $k = n = 5$. To apply Theorem 2, we need to write down the restrictions Q_j for

---

[11]See Zha (1999) for restrictions on the lagged structure as well.

$j = 1, \ldots, 5$. These matrices are

$$
Q_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \ Q_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \ Q_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
$$

$$
Q_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } Q_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.
$$

It follows from the Q_j's for $j = 1, \ldots, 5$ that $q_1 = 4$, $q_2 = 3$, $q_3 = 3$, $q_4 = 1$, and $q_5 = 0$ and the total number of restrictions ($\sum_{j=1}^{5} q_j$) is 11, greater than $n(n-1)/2 = 10$. Therefore, by Rothenberg (1971)'s order condition, the model may be overidentified. Since the order condition is only necessary, we apply the sufficient condition of Theorem 2 by filling the matrices $M_j \left(f(A_0, A_+) \right)$ for $j = 1, \cdots, 5$ as

$$
\begin{bmatrix} 0 & a_{22} & 0 & a_{24} & a_{25} \\ 0 & 0 & a_{33} & a_{34} & a_{35} \\ 0 & 0 & a_{43} & a_{44} & a_{45} \\ 0 & 0 & 0 & 0 & a_{55} \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \ \begin{bmatrix} 0 & 0 & a_{33} & a_{34} & a_{35} \\ 0 & 0 & a_{43} & a_{44} & a_{45} \\ 0 & 0 & 0 & 0 & a_{55} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \ \begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} & a_{15} \\ 0 & a_{22} & 0 & a_{24} & a_{25} \\ 0 & 0 & 0 & 0 & a_{55} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix},
$$

$$
\begin{bmatrix} 0 & 0 & 0 & 0 & a_{55} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.
$$

Clearly there exist values of a's such that matrix $M_j \left(f(A_0, A_+) \right)$ has the rank $n = 5$ for $j = 1, \ldots, 5$. According to Theorems 2 and 4, the model is globally identified almost everywhere in the structural parameter space where the identifying restrictions are satisfied.

V.4. **Open-economy SVARs.** Cushman and Zha (1997), Kim (1999), and Kim and Roubini (2000) extend the non-triangular identification to open economies such as Canada and European countries. Kim (1999) and Kim and Roubini (2000) only consider contemporaneous restrictions, while Cushman and Zha (1997) impose restrictions on the lag structure in addition to the restrictions on the contemporaneous matrix.[12] In this subsection, we only analyze the model of Kim (1999). For the illustrative purpose, this model is relatively small and the restrictions are imposed on the contemporaneous matrix A_0 only. For the SVARs of Cushman and Zha (1997) and Kim and Roubini (2000), one can use the similar operational approach employed in Section V.3 to show that those models are indeed globally identified. We leave the verification to the reader.

Kim (1999) uses monthly data on five variables: the call money rate (R), the monetary aggregate (M), the consumer price index (P), the industrial production (y), and the world export commodity price index (P_c). Except for R, all the other variables are expressed in log. The restrictions for Kim (1999)'s identification can be expressed as

$$
A_0 = \begin{array}{c} \\ R \\ M \\ P \\ y \\ P_c \end{array}
\begin{bmatrix}
\text{PS} & \text{PS} & \text{MP} & \text{MD} & \text{Inf} \\
0 & 0 & a_{13} & a_{14} & a_{15} \\
0 & 0 & a_{23} & a_{24} & a_{25} \\
0 & a_{32} & 0 & a_{34} & a_{35} \\
a_{41} & a_{42} & 0 & a_{44} & a_{45} \\
0 & 0 & a_{53} & 0 & a_{55}
\end{bmatrix} . \tag{20}
$$

The label 'PS' on the top stands for the production sector, 'MP' for the monetary policy equation, 'MD' for the money demand, and 'Inf' for the information equation. For the production sector, the variables P and y are arbitrarily ordered in a triangular form, and other financial variables such as R, M, and P_c do not enter this sector. For the monetary policy equation, the monetary authority does not react to y and P because output and the general price level cannot be observed within the month. The money demand function involves only the four variables R, M, P, and y. The last column labeled as 'Inf' suggests that the commodity prices respond to all the variables in the complete information market.

---

[12]As discussed in Section V.2, restrictions on the lag structure may be crucial in helping achieve identification of an SVAR.

The transformation function for this model is

$$f(A_0, A_+) = A_0.$$

From (20), one can see that $n = 5$ and $q_j = n - j$ for $j = 1, \ldots, 5$. Then, it follows from Theorem 8 that Kim (1999)'s SVAR is exactly identified.

Now suppose that we allow the variable P_c to be treated as part of the production sector so that they enter the columns labelled as 'PS.' With this change, the number of restrictions are now less than $n(n-1)/2 = 10$ and Rothenberg (1971)'s order condition is violated. Thus, the model is not identified. To meet the order condition, at least two additional restrictions are needed. Following Cushman and Zha (1997), we assume that the contemporaneous money demand equation takes the functional form

$$M - P = y - a_{14}/a_{24}R,$$

which is consistent with many DSGE models (see, for example, Blanchard and Fischer (1989)). This alternative identification can be expressed as

$$A_0 = \begin{array}{c} R \\ M \\ P \\ y \\ P_c \end{array} \begin{bmatrix} \text{MD} & \text{PS} & \text{PS} & \text{MP} & \text{Inf} \\ a_{14} & 0 & 0 & a_{13} & a_{15} \\ a_{24} & 0 & 0 & a_{23} & a_{25} \\ -a_{24} & 0 & a_{32} & 0 & a_{35} \\ -a_{24} & a_{41} & a_{42} & 0 & a_{45} \\ 0 & a_{51} & a_{52} & a_{53} & a_{55} \end{bmatrix}, \tag{21}$$

with the same transformation function as the previous identification.

Note that we have reordered the equations so that the convention given by (4) is satisfied. Since the total number of restrictions for this alternative identification is now equal to $n(n-1)/2 = 10$, Rothenberg (1971)'s order condition for exact identification is met. The question is whether the model, in fact, exactly identified under this identification. Without utilizing our theory, it is impossible to answer the question directly from (21). To apply Theorem 8, we express the restrictions Q_j for

$j = 1, \ldots, 5$ implied by (21) as

$$
Q_1 = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \; Q_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \; Q_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
$$

$$
Q_4 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } Q_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.
$$

It is straightforward to see that $q_1 = 3$, $q_2 = 3$, $q_3 = 2$, $q_4 = 2$, and $q_5 = 0$. As a result, the rank condition of Theorem 8 is violated and therefore the model described by (21) is not exactly identified. This example is informative because the alternative identification is as economically plausible as Kim (1999)'s original identification, and yet without our theory one may conclude that the model is exactly identified while it is not.

V.5. **Restrictions on impulse responses.** It has become increasingly popular that identifying restrictions are imposed directly on impulse responses (Sims, 2005). One of the most important advantages of our theory is that it allows one to determine whether nonlinear restrictions on the structural parameters resulting from restrictions on impulse responses identify the model globally. In Section II.4 we show how we can write transformation functions $f(\cdot)$ to handle restrictions on impulse responses at various horizons. In this section we show how to apply our theory to a particular SVAR with short-run and long-run restrictions on impulse responses as introduced by Galí (1992). This kind of restriction has been widely used to obtain stylized facts for DSGE modeling.

Following Peersman and Smets (2003), we consider a four-variable SVAR with three contemporaneous and three long-run restrictions on impulse responses. The four endogenous variables are quarterly output growth ($\Delta \log Y$), quarterly inflation (ΔP), the nominal short-term interest rate (R), and a quarterly change of the nominal exchange rate euro/dollar ($\Delta \log Ex$). The short-run restrictions are:

- Monetary policy shocks have no contemporaneous effect on output.
- Exchange rate shocks have no contemporaneous effect on output.
- Exchange rate shocks have no contemporaneous effect on the interest rate.

TABLE 1. Restrictions implying that the model is identified

$$f(A_0, A_+) = \begin{bmatrix} L_0 \\ L_\infty \end{bmatrix} = $$

	Ex	P	D	S
$\Delta \log Y$	0	0	×	×
$\Delta \log P$	×	×	×	×
R	0	×	×	×
$\Delta \log Ex$	×	×	×	×
$\Delta \log Y$	0	0	0	×
$\Delta \log P$	×	×	×	×
R	×	×	×	×
$\Delta \log Ex$	×	×	×	×

The long-run restrictions on impulse responses are:

- Aggregate demand shocks have no long-run effect on output.
- Monetary policy shocks have no long-run effect on output.
- Exchange rate shocks have no long-run effect on output.

The transformation function of the original parameters and the identifying restrictions on the transformed parameters are represented in Table 1. On the top row of the table, 'Ex' stands for a shock to the exchange rate market, 'P' for a monetary policy shock, 'D' for a demand shock,, and 'S' for a supply shock. The symbol '×' means that no restriction is imposed, and '0' means an exclusion restriction.

From this table one can see that $n = 4$, $k = 2n = 8$, $q_1 = 3$, $q_2 = 2$, $q_3 = 1$, and $q_4 = 0$. The total number of restrictions ($\sum_{j=1}^{4} q_j$) is equal to $n(n-1)/2 = 6$ and Rothenberg (1971)'s order condition for exact identification holds. Because $q_j = n - j$ for $j = 1, \ldots, 4$, this model is exactly identified according to Theorem 8.

To emphasize the importance of Theorem 8, we consider the assumption of supply shocks having no contemporaneous effect on inflation because of the price stickiness, in place of the original assumption that exchange rate shocks have no contemporaneous effect on output. This alternative identification implies the set of restrictions on $f(A_0, A_+)$ as represented in Table 2.

For this alternative set of restrictions, the total number of restrictions is still equal to 6 and therefore Rothenberg (1971)'s order condition for exact identification holds. But it is straightforward to show that Theorem 8 is not satisfied, because $q_1 = 2$ (the number of restrictions in the 'Ex' column) is the same as $q_2 = 2$ (the number

TABLE 2. Restrictions implying that the model is not identified

$$f\left(A_{0}, A_{+}\right)=\left[\begin{array}{c} L_{0} \\ L_{\infty} \end{array}\right]=$$

	Ex	P	D	S
$\Delta \log Y$	$\times$	0	$\times$	$\times$
$\Delta \log P$	$\times$	$\times$	$\times$	0
R	0	$\times$	$\times$	$\times$
$\Delta \log Ex$	$\times$	$\times$	$\times$	$\times$
$\Delta \log Y$	0	0	0	$\times$
$\Delta \log P$	$\times$	$\times$	$\times$	$\times$
R	$\times$	$\times$	$\times$	$\times$
$\Delta \log Ex$	$\times$	$\times$	$\times$	$\times$

of restrictions in the 'P' column) and $q_3 = 1$ (the number of restrictions in the 'D' column) is the same as $q_4 = 1$ (the number of restrictions in the 'S' column). Thus, this is another example where, if we had naively applied the order condition, we would have wrongly concluded that the model is exactly identified.

VI. ALGORITHMS FOR ESTIMATION AND SMALL-SAMPLE INFERENCE

In Sections II and III we have developed a general theory to determine whether a wide class of restrictions identify SVARs globally. In Section V we have used our theory to establish formally that many widely-used SVARs are globally identified. We have shown that slight modifications in restrictions may render the model unidentified. We believe that the development of this theory is important because global identification of an SVAR must be a first object to establish in the SVAR analysis.

After global identification has been determined, the next step is to perform small-sample estimation and inference. For the maximum likelihood estimate or the posterior estimate when a prior is used, the existing estimation method for SVARs with restrictions on impulse responses is inefficient. This inefficiency can become more serious for small-sample inference, because Bayesian MCMC methods or classical bootstrap procedures often require expensive computation of randomly sampling structural parameters (Kilian, 1998; Geweke, 1999; Inoue and Kilian, 2002; Geweke, 2005; Pesavento and Rossi, 2006). These existing methods become extremely expensive when time-varying SVARs are studied (Uhlig, 1997; Canova and Gambetti, 2004; Cogley and Sargent, 2005; Primiceri, 2005; Sims and Zha, 2006b; Gambetti, Pappa,

and Canova, forthcoming). Take as an example an exactly identified model with drifting parameters and with restrictions directly imposed on impulse responses. The conventional method of Galí (1992) involves solving a system of nonlinear equations for every draw of the parameters and at each time t for the time-varying SVAR. For a twelve-lag SVAR of more than three variables, it quickly becomes computationally infeasible to have as many as millions of draws that are often required to achieve accurate small-sample inferences.

To address this practical problem, in this section we first build on Theorem 6 and develop an algorithm to achieve computational efficiency for exactly identified models. Second, we show that for a triangular system an even faster algorithm is feasible. Third, we derive a computationally efficient algorithm designed for sign restrictions. This algorithm improves considerably on the existing methods, and it is important because sign restrictions have been widely used in the recent literature. Finally, we describe a class of priors that allow us to use these algorithms in the Bayesian framework.

VI.1. **Algorithms for exactly identified models.** Assume that the model is exactly identified. Let $f(A_0, A_+)$ be the associated transformation function and let $Q_1, Q_2, \cdots, Q_n$ represent the identifying restrictions. Theorem 6 tells us that for any value of (A_0, A_+), either an estimate or a particular draw, there is a unique orthogonal matrix P such that $(A_0 P, A_+ P)$ satisfies the identifying restrictions. The matrix P is sometimes called the rotation matrix. The core of our argument is that, instead of solving a complicated system of nonlinear equations as in Galí (1992), we can find the rotation matrix P in a very efficient manner. The following algorithm gives a step-by-step description of how to find this rotation matrix efficiently. Recall that we follow the convention that the transformation function $f(\cdot)$ has been so chosen that $rank(Q_j) = q_j = n - j$ for $j = 1, \cdots, n$ as in Theorem 8.

Algorithm 1. Let an SVAR be exactly identified and (A_0, A_+) be any value of the unrestricted structural parameters.

(Step 1) Set $j = 1$.

(Step 2) Form the matrix

$$
\tilde{Q}_j = \begin{bmatrix} Q_j f\left(A_0, A_+\right) \\ p_1' \\ \vdots \\ p_{j-1}' \end{bmatrix}.
$$

If $j = 1$, then $\tilde{Q}_j = Q_j f\left(A_0, A_+\right)$.

(Step 3) Let p_j be any unit-length vector such that $\tilde{Q}_j p_j = 0$. Such a vector exists because $rank\left(Q_j\right) = n - j$ and hence $rank\left(\tilde{Q}_j\right) < n$. Use the LU decomposition of $\tilde{Q}_j$ to find this unit-length vector p_j.

(Step 4) If $j = n$, stop; otherwise, set $j = j + 1$ and go to Step 2.

The above algorithm produces the orthogonal matrix

$$
P = \begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}
$$

that is guaranteed by Theorem 6. As shown in Section II.3, the restrictions represented by $f\left(A_0, A_+\right)$ and $Q_1, Q_2, \cdots, Q_n$ are very general. This generality makes our algorithm useful for a large set of identifying restrictions.

How does the algorithm work for small-sample estimation and inference? Suppose that one wishes to find the ML estimate (or the estimate at the posterior peak) of the restricted model. Assume one is able to get the ML estimate (or the estimate at the posterior peak) for the unrestricted structural parameters or the reduced-form parameters.[13] Algorithm 1 provides us an orthogonal matrix P that rotates the unrestricted estimate to the estimate that satisfies the identifying restrictions. If the original estimate is for the reduced-form parameters, one can use Algorithm 1 to rotate the Cholesky decomposition of Σ to get the estimate of structural parameters that satisfy the restrictions.

Suppose now that one wishes to perform small-sample inference by using the bootstrap procedure or the Bayesian MCMC method to construct confidence intervals of structural parameters. Denote a draw of the unrestricted structural parameters by $\left(A_0^*, A_+^*\right)$. For such a draw, one uses Algorithm 1 to find the rotation matrix P such that $Q_j f\left(A_0^* P, A_+^* P\right) e_j = 0$ for all $1 \leq j \leq n$. If one works on a draw of

---

[13]Such an estimate of the unrestricted structural parameters may not be unique, but it gives the same likelihood or posterior value as other estimates at the peak of the likelihood or the posterior density.

TABLE 3. Short-run and long-run restrictions

$$
f\left(A_0, A_+\right) =
\begin{bmatrix} L_0 \\ L \end{bmatrix} =
\begin{array}{c}
\Delta \log Y \\
R \\
\log P \\
\Delta \log Y \\
R \\
\log P
\end{array}
\begin{array}{ccc}
P & D & S \\
\left[\begin{array}{ccc}
0 & \times & \times \\
\times & \times & \times \\
\times & \times & \times \\
0 & 0 & \times \\
\times & \times & \times \\
\times & \times & \times
\end{array}\right]
\end{array}
$$

the reduced-form parameters, one can obtain (A_0^*, A_+^*) from the Cholesky decomposition of the reduced-form covariance matrix. Given (A_0^*, A_+^*), one can then use Algorithm 1 to find the rotation matrix P such that $Q_j f\left(A_0^* P, A_+^* P\right) e_j = 0$ for all $1 \le j \le n$.

VI.2. **An example.** To illustrate how Theorem 6 and Algorithm 1 work in practice, we present a simple example of finding the rotation matrix P using Algorithm 1. To maximize clarity of the exposition, we consider a simple three-variable standard SVAR with one lag so that $A_+ = A_1$, the analysis of which can be easily extended to more variables and more lags. The three variables are output growth ($\Delta \log Y$), the interest rate (R), and inflation ($\Delta \log P$). There are three identifying restrictions: demand shocks have no long-run effect on output, and monetary policy shocks have neither a short-run nor a long-run effect on output. These restrictions on impulse responses can be expressed as the restrictions on the columns of the transformed matrix $f(\cdot)$. Table 3 presents this transformation and the restrictions on the transformed parameters. In the table, the symbol '$\times$' indicates no restrictions, '0' indicates an exclusion restriction, 'P' stands for policy shocks, 'D' for demand shocks, and 'S' for supply shocks.

The first and foremost step is to determine whether this system is identified. From Table 3 one can see that $n = 3$, $q_1 = 2$, $q_2 = 1$, and $q_3 = 0$. It follows from Theorem 8 that this system is exactly identified. Therefore, for almost any value of (A_0, A_+), there exists a unique rotation matrix P such that $Q_j f\left(A_0 P, A_+ P\right) e_j = 0$ for all $1 \le j \le n$.

To show how to find such a rotation matrix using Algorithm 1, we express the restrictions described by Table 3 by the matrices Q_j's as follows

$$Q_1 = \begin{bmatrix} 1\ 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 1\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0\ 0\ 0\ 1\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0 \end{bmatrix}, \text{ and } Q_3 = \begin{bmatrix} 0\ 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0 \end{bmatrix}.$$

By deleting the rows of zeros out of the above Q_j's, we have

$$\bar{Q}_1 = \begin{bmatrix} 1\ 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 1\ 0\ 0 \end{bmatrix} \text{ and } \bar{Q}_2 = \begin{bmatrix} 0\ 0\ 0\ 1\ 0\ 0 \end{bmatrix}.$$

Since all the rows in Q_3 are zeros, there is no $\bar{Q}_3$. Working with $\bar{Q}_j$ is operationally easier than working with Q_j, since $\tilde{Q}_j$ in Algorithm 1 derived from $\bar{Q}_j$ will always be an $(n-1) \times n$ matrix.[14]

For the purpose of walking through Algorithm 1, suppose that the estimates of the reduced-form parameters B and Σ are

$$B = \begin{bmatrix} 0.5 & -1.25 & -1 \\ 0.5 & 0.25 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} 1 & 0.5 & 1 \\ 0.5 & 4.25 & 2.5 \\ 1 & 2.5 & 3 \end{bmatrix}.$$

We compute A_0 from the Cholesky decomposition of Σ^{-1} (in Matlab, $A_0 = chol\left(\Sigma^{-1}\right)'$), and $A_+ = BA_0$. Since $L'_0 = \left(A_0^{-1}\right)$ and $L'_\infty = (A_0 - A_+)^{-1}$, we have

$$f(A_0, A_+) = \begin{bmatrix} L_0 \\ L_\infty \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{22}$$

and

$$\tilde{Q}_1 = \bar{Q}_1 f(A_0, A_+) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

---

[14]In most applications it is obvious how to form $\bar{Q}_j$, but one can always use the Matlab function orth() and define

$$\bar{Q}_j = orth\left(Q_j'\right)'.$$

15

The first step in Algorithm 1 is to find a unit length vector p_1 such that $\tilde{Q}_1 p_1 = 0$. The most computationally efficient method of finding this vector is to employ the LU decomposition of $\tilde{Q}_1$. However, it is often more convenient to employ the QR decomposition of $\tilde{Q}_1'$.[16] Let $\tilde{Q}_1' = QR$ where Q is orthogonal and R is upper triangular. If we choose p_1 to be the last row of Q, then

$$\tilde{Q}_1 p_1 = R'Q'p_1 = R' \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0,$$

since the last column of R' is zero. Therefore p_1 is equal to

$$p_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

To obtain p_2, we form

$$\tilde{Q}_2 = \begin{bmatrix} \bar{Q}_2 f\,(A_0, A_+) \\ p_1' \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

As before, take p_2 to be the last row of the orthogonal component of the QR decomposition of $\tilde{Q}_2'$ to get

$$p_2 = \begin{bmatrix} 0.7071 \\ -0.7071 \\ 0 \end{bmatrix}.$$

To obtain p_3, we form

$$\tilde{Q}_3 = \begin{bmatrix} p_1' \\ p_2' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0.7071 & -0.7071 & 0 \end{bmatrix}.$$

---

[16]In Matlab, the function qr() applied to an $n \times (n-1)$ matrix returns an $n \times n$ orthogonal matrix Q and an $n \times (n-1)$ upper triangular matrix R. In other software packages, however, the "orthogonal" matrix Q may be $n \times n - 1$ and the triangular matrix R may be $n - 1 \times n - 1$. If those packages are used, one needs to pad the matrix $\tilde{Q}_j$ with a row of zeros before proceeding further. In either case, the last row of R will be zero.

Again, letting p_3 be the last row of the orthogonal matrix of the QR decomposition of $\tilde{Q}'_3$, we get

$$p_3 = \begin{bmatrix} -0.7071 \\ -0.7071 \\ 0 \end{bmatrix}.$$

Combining these results, one obtains the orthogonal matrix

$$P = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix} = \begin{bmatrix} 0 & 0.7071 & -0.7071 \\ 0 & -0.7071 & -0.7071 \\ 1 & 0 & 0 \end{bmatrix}.$$

It is straightforward to verify that

$$Q_j f \left(A_0 P, A_+ P \right) e_j = 0$$

for all $1 \le j \le 3$ or

$$\bar{Q}_j f \left(A_0 P, A_+ P \right) e_j = 0$$

for all $1 \le j \le 2$.

VI.3. **Algorithms for triangular systems.** While Algorithm 1 gives us an efficient way to find the rotation matrix P for exactly identified models, in this section we present a much faster algorithm for triangular systems that are defined as follows.

Definition 7. The restrictions in the form of (3) and (4) are triangular if and only if there exists an invertible matrix P_1 such that the matrix $P_1 f \left(A_0, A_+ \right) P_0$ is lower triangular, where

$$P_0 = \begin{bmatrix} 0 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 0 \end{bmatrix}.$$

According to Definition 7, identifying restrictions are triangular if they can be transformed into a lower triangular system. For a triangular system, Algorithm 1 can be so improved that the orthogonal matrix given by Theorem 6 can be found using only a single (instead of successive) QR decomposition as described in the following theorem.

Theorem 9. Suppose the restrictions are triangular as defined in Definition 7. Let P_1 be the matrix that makes the restrictions triangular, (A_0, A_+) be any structural

parameters, $(P_1 f(A_0, A_+))' = QR$ using the QR decomposition (where Q is an orthogonal matrix and R is upper triangular), and $P = QP_0$. Then

$$Q_j f(A_0 P, A_+ P) e_j = 0 \text{ for } 1 \leq j \leq n.$$

Proof. Since $P_1 f(A_0, A_+) = R'Q'$,

$$
\begin{aligned}
R' &= P_1 f(A_0, A_+) Q P_0 P_0 \\
&= P_1 f(A_0 Q P_0', A_+ Q P_0) P_0.
\end{aligned}
$$

Since R' is lower triangular, $(A_0 Q P_0', A_+ Q P_0')$ satisfies the restrictions by Definition 7. □

In the example of Section VI.2, the restrictions implied by Q_j or $\bar{Q}_j$ are of the exclusion type. This type simply sets particular parameters to zero. Formally, identifying restrictions in the form of (3) and (4) are of the exclusion type if, for all $1 \leq j \leq n$, each row of Q_j has zeros and ones, with a single one at most. It follows from Definition 7 that if the restrictions are of the exclusion type, P_1 becomes a permutation matrix and thus the system can be permuted into a lower triangular system. To show how such a permutation task can be accomplished, we first postmultiply the matrix $f(A_0, A_+)$ by P_0 to reverse the order of its columns and then use P_1 to permute the rows of $f(A_0, A_+) P_0$ into a triangular form by swapping the first and fourth rows. These operations imply

$$
P_1 = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.
$$

Now that we have proven that the restrictions used in Section VI.2 are triangular by Definition 7, we can use Theorem 9 to find the rotation matrix P. As an illustration, we use the same values of the reduced-form parameters B and Σ as in Section VI.2.

The numerical value of the matrix $f(A_0, A_+)$ is given by (22). The QR decomposition of $(P_1 f(A_0, A_+))'$ gives

$$
Q = \begin{bmatrix} -0.7071 & -0.70710 & 0 \\ -0.7071 & 0.7071 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
$$

The required rotation matrix is equal to

$$
P = QP_0 = \begin{bmatrix} 0 & -0.70710 & -0.7071 \\ 0 & 0.7071 & -0.7071 \\ 1 & 0 & 0 \end{bmatrix}.
$$

The reader can easily verify that $f(A_0 P, A_+ P)$ satisfies the identifying restrictions, i.e., $Q_j f(A_0 P, A_+ P) e_j = 0$ for $1 \leq j \leq n$.

VI.4. **Sign restrictions.** The identifying restrictions described in Section II.3 and the algorithms developed in Sections VI.1 and VI.3 are based on linear restrictions on transformed structural parameters. One objective in employing this class of restrictions is to identify structural shocks. According to the conventional wisdom (and many DSGE models), for example, a contractionary monetary policy shock should raise the interest rate and lower output and prices. Thus, a successful identification should produce impulse responses that conform to this conventional wisdom. Some restrictions of the identification type described in Section II.3, such as the triangular identification, may not generate impulse responses that have the desired signs. In this situation, Faust (1998), Canova and De Nicoló (2002), and Uhlig (2005) propose an alternative approach.[17] Their basic idea is to use sign restrictions directly imposed on impulse responses such that, for example, the interest rate rises while money, output, and prices fall in response to a contractionary monetary shock.

The algorithms established in Sections VI.1 and VI.3 cannot be applied to small-sample estimation and inference of an SVAR with sign restrictions. The reason is that an SVAR with sign restrictions on impulse responses is *not* exactly identified. According to Theorem 6, a necessary and sufficient condition for an SVAR to be exactly identified is that for any value of (A_0, A_+), there exists a unique rotation matrix P such that $(A_0 P, A_+ P)$ satisfies the restrictions. For sign restrictions, however, there exist a number of such P's (Fry and Pagan, 2007).

---

[17]The algorithms of Faust, Canova and De Nicoló, and Uhlig are briefly described in Appendix E.

Canova and DeNicoló (2002) propose an algorithm that is designed to find one of these P's. Suppose that such a P exists (i.e., the impulse responses generated through the rotation matrix P satisfy the sign restrictions), an important question is whether their algorithm always finds it (Fry and Pagan, 2007). The answer is positive, as formally established by the following theorem.

Theorem 10. Let P be an $(n \times n)$ orthogonal matrix. There exists $\theta_{i,j}$ for $1 \leq i < j \leq n$ such that (1) $0 \leq \theta_{i,j} < \pi$ for $i < j < n$, (2) $0 \leq \theta_{i,n} < 2\pi$, and (3)[18]

$$P = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} Q_{i,j}\left(\theta_{i,j}\right),$$

or

$$P = S \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} Q_{i,j}\left(\theta_{i,j}\right),$$

where

$$S = \begin{bmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & -1 \end{bmatrix},$$

and

$$Q_{i,j}\left(\theta_{i,j}\right) = \begin{bmatrix} & & \text{col } i & & \text{col } j & & \\ & & \downarrow & & \downarrow & & \\ 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ \text{row } i \rightarrow \ 0 & \cdots & \cos\left(\theta_{i,j}\right) & \cdots & -\sin\left(\theta_{i,j}\right) & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ \text{row } j \rightarrow \ 0 & \cdots & \sin\left(\theta_{i,j}\right) & \cdots & \cos\left(\theta_{i,j}\right) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}.$$

Proof. See Appendix E.2 for the proof. □

The algorithm of Canova and DeNicoló (2002), as discussed in Appendix E.2, becomes computationally infeasible when an SVAR system is moderately large (e.g.,

---

[18]In Canova and De Nicoló (2002), the notation $Q_{i,j}\left(\theta\right)$ is used where θ is implicitly assumed to vary with different i and j.

$n > 4$). To solve this inefficiency problem, we develop a new algorithm, based on the following theorem, for small-sample estimation and inference of an SVAR with sign restrictions.

Theorem 11. Let $\tilde{X}$ be an $n \times n$ random matrix with each element having an independent standard normal distribution. Let $\tilde{X} = \tilde{Q}\tilde{R}$ be the QR decomposition of $\tilde{X}$ with the diagonal of $\tilde{R}$ normalized to be positive. Then $\tilde{Q}$ has the uniform (or Haar) distribution.

Proof. The proof follows directly from Stewart (1980).[19]   □

Theorem 11 gives us a convenient way of implementing a random selection of orthogonal matrices to obtain impulse responses that satisfy the sign restrictions. The following algorithm describes this implementation.

Algorithm 2. Let (A_0, A_+) be any value of the *unrestricted* structural parameters.

(Step 1) Draw an independent standard normal $n \times n$ matrix $\tilde{X}$ and let $\tilde{X} = \tilde{Q}\tilde{R}$ be the QR decomposition of $\tilde{X}$ with the diagonal of $\tilde{R}$ normalized to be positive.

(Step 2) Let $P = \tilde{Q}$ and generate impulse responses from $A_0 P$ and $B = A_+ A_0^{-1}$.

(Step 3) If these impulse responses do not satisfy the sign restrictions, return to Step 1.

If (A_0, A_+) is the estimate of unrestricted structural parameters, $(A_0 P, A_+ P)$ obtained via Algorithm 2 is the estimate of structural parameters satisfying the sign restrictions. If (A_0, A_+) is a draw of unrestricted parameters, $(A_0 P, A_+ P)$ obtained via Algorithm 2 is a draw of structural parameters satisfying the sign restrictions.[20] If B and Σ are a draw of the reduced-form parameters, we use B and the Choleski decomposition of Σ to obtain (A_0, A_+) and then use Algorithm 2 to obtain $(A_0 P, A_+ P)$ that satisfies the sign restrictions.

Our algorithm should be viewed as a generalized version of Uhlig (2005)'s algorithm. If several structural shocks must be identified, Uhlig's algorithm searches

---

[19]Stewart (1980) has even more efficient algorithms for generating uniform random orthogonal matrices, but they are less straightforward and more difficult to implement.

[20]In theory the algorithm is not guaranteed to terminate. In practice, we set a maximum number of iterations to be 100,000 for Steps (2)-(4) to be repeated. If the maximum is reached, the algorithm should move to Step (1) to draw another orthogonal matrix $\tilde{Q}$. In our MCMC experiments, the maximum of iterations was never reached for millions of simulations.

for the orthogonal matrix column by column recursively. During this search, the orthogonal matrix may not be found for some draws, either from the bootstrap procedure or from the posterior distribution. Our algorithm, based on the Householder-transformation methodology, differs from Uhlig (2005)'s algorithm in two aspects: (1) all the posterior draws are kept in practice, and (2) the orthogonal matrix is simply a draw from the uniform (or Haar) distribution with only a single operation of the QR decomposition. These differences make our algorithm more efficient when several shocks are to be identified. Especially for a time-varying SVAR system with more than three or four structural shocks to be identified, the efficiency gain can be as high as the 10-1 ratio when our algorithm is compared with the algorithms of Uhlig and Canova and De Nicoló (see, for example, Benati and Mumtaz (2006) for an application of our algorithm in their DSGE model).[21]

VI.5. **A reference prior.** Sections VI.1, VI.3, and VI.4 develop new algorithms usable to small-sample estimation and inference of a large class of SVARs. If we are to obtain the ML estimate (or the estimate at the posterior peak) of the restricted model, the algorithms use the estimate of the unrestricted parameters, denoted by (A_0, A_+), to find an orthogonal matrix P such that the transformed parameters $(A_0 P, A_+ P)$ satisfy the restrictions. The same procedure applies to small-sample inference.

In the Bayesian framework, an additional property is needed: the transformed parameters $(A_0 P, A_+ P)$ must have the same prior distribution as the original parameters (A_0, A_+). This property ensures that our algorithms are valid for Bayesian estimation and inference as well as for computation of the marginal data density for model comparison.[22]

Do the existing priors used in the Bayesian SVAR literature have this property? In this subsection, we show that the reference prior of Sims and Zha (1998) preserves the property that (A_0, A_+) and $(A_0 P, A_+ P)$, where P is an orthogonal matrix, share the same prior distribution. The Sims and Zha (1998) prior is general and encompasses the popular Minnesota prior. The prior density for the unrestricted parameters takes the following form

$$a_0 = \text{vec}(A_0) \sim N(0, I_n \otimes H_0), \tag{23}$$

---

[21]Since the particular application studied by Uhlig (2005) concerns only a monetary policy shock (not any other shocks), the efficiency gain from our algorithm is negligible.

[22]We thank a referee for bringing up this important point.

and

$$a_+ \mid a_0 = \text{vec}(A_+) \mid a_0 \sim N\left(\text{vec}(\bar{S}A_0), I_n \otimes H_+\right),\tag{24}$$

where H_0 and H_+ are symmetric positive definite matrices and $\bar{S}$ is typically set as

$$\begin{bmatrix} I \\ {}_{n \times n} \\ \mathbf{0} \\ {}_{m \times n} \end{bmatrix}$$

to represent a random-walk component of the prior.

An essential part of the algorithms developed in Sections VI.1, VI.3, and VI.4 is about finding a rotation P. An orthogonal rotation of A_0 and A_+ leads to $\tilde{A}_0 = A_0 P$ and $\tilde{A}_+ = A_+ P$, where $\tilde{A}_0$ and $\tilde{A}_+$ are restricted structural parameters satisfying the identifying restrictions. The key question is whether the Sims and Zha prior is invariant to such a rotation. The following two identities prove that this invariance result holds.

$$\tilde{a}_0 = \text{vec}(A_0 P) = (P' \otimes I_n)a_0 \sim N(0, I_n \otimes H_0),$$

and

$$\begin{aligned}
\tilde{a}_+ \mid \tilde{a}_0 &= \text{vec}(A_+ P) \mid \tilde{a}_0 \\
&= (P' \otimes I_m)a_+ \mid \tilde{a}_0 \\
&\sim N\left((P' \otimes I_m)\text{vec}(\bar{S}A_0), I_n \otimes H_+\right) = N\left(\text{vec}(\bar{S}A_0), I_n \otimes H_+\right).
\end{aligned}$$

In other words, (A_0, A_+) and $(A_0 P, A_+ P)$ share the same prior density as long as P is an orthogonal matrix. Since an orthogonal rotation of (A_0, A_+) has no effect on the likelihood function, the posterior density for $(A_0 P, A_+ P)$ is the same as that of (A_0, A_+).

VII. Conclusion

SVARs are widely used for policy analysis and to provide stylized facts for economic theory. Before one proceeds to perform an empirical analysis with a particular SVAR, however, it is essential that the model be checked to ascertain its identifiability. Otherwise, the empirical results would be misleading.

In this paper we have made contributions towards closing important theoretical gaps in the SVAR literature by providing a general theory of global identification and practical algorithms for small-sample estimation and inference. In particular,

we have established general rank conditions for global identification of SVAR models. These conditions can be checked as a simple matrix-filling and rank-checking exercise.

The usual rank conditions for (local) identification in the literature involve computing the rank of certain derivatives (Fisher, 1966; Rothenberg, 1971; Hausman and Taylor, 1983). We are able to obtain much simpler conditions by exploiting the orthogonal structure given by Theorem 1 and thus do not have to explicitly compute any derivatives. Consequently, for exactly identified SVARs, our necessary and sufficient condition for identification involves simply counting restrictions and checking the pattern of these restrictions, an exercise no more complicated than Rothenberg (1971)'s order condition.

Our theoretical results apply to a large class of identifying restrictions, including nonlinear restrictions on the parameters such as widely-used short-run and long-run restrictions imposed on impulse responses. Our theory is valid for both classical and Bayesian analysis of SVAR models. Moreover, the efficient algorithms developed in this paper provide essential tools for researchers to study a variety of empirical SVARs for comparative and robustness analysis.

APPENDIX A. DIFFERENTIABLE MANIFOLDS AND SETS OF MEASURE ZERO

We have used differentiable manifolds through out this paper. A trivial example of a differentiable manifold is any open subset of a Euclidean space, so both $\mathbb{P}^S$, the set of all structural parameters, and U, the set of admissible structural parameters are differentiable manifolds. More interesting examples are $\mathbb{P}^R$, $\hat{U}$, and $O(n)$, the set of all reduced-form parameters, the set of admissible reduced-form parameters, and the set of all $n \times n$ orthogonal matrices, respectively. In this appendix, we highlight a few properties of differentiable manifolds and develop some results about measure zero subsets of differentiable manifolds. See Spivak (1965, Chapter 5) for a further discussion of differentiable manifolds. We will henceforth refer to differentiable manifolds as simply manifolds. A k dimensional manifold in $\mathbb{R}^n$ is defined as follows.

Definition 8. A subset M of $\mathbb{R}^n$ is a k-dimensional manifold if for every $x \in M$ there exists an open set $V \subset \mathbb{R}^n$ containing x, an open set $W \subset \mathbb{R}^k \times \mathbb{R}^{n-k}$, and a continuously differentiable function $h : V \to W$ with continuously differentiable inverse such that

$$
\begin{aligned}
h(V \cap M) &= W \cap \left(\mathbb{R}^k \times \{0\} \right) \\
&= \{(y_1, \cdots, y_n) \in W \mid y_{k+1} = \cdots = y_n = 0\}.
\end{aligned}
$$

The pair (V, h) defines a n dimensional coordinate system about x. If $\pi : \mathbb{R}^n \to \mathbb{R}^k$ is defined by $\pi(x_1, \cdots, x_n) = (x_1, \cdots, x_k,)$, then the n-dimensional coordinate system restricts to a k-dimensional coordinate system $(V \cap M, \pi \circ h)$. Thus a manifold is locally a Euclidean space. We exploit this local Euclidean structure to characterizer sets of measure zero on manifolds.

While measure zero sets are often studied in the context of Lebesgue measure, for Euclidean spaces we do not need the full power of this machinery. We define sets of measure zero as follows.

Definition 9. A set $A \subset \mathbb{R}^k$ is of measure zero if and only if for every $\varepsilon > 0$ there exist countably many closed k-dimensional rectangles R_i of volume v_i such that $A \subset \bigcup_i R_i$ and $\sum_i v_i < \varepsilon$.

This definition of measure zero is equivalent to definition arising from Lebesgue measure. We can use the local coordinate systems to extend the notion of sets of measure zero to manifolds.

Definition 10. Let A be a subset of the k-dimension manifold M in $\mathbb{R}^n$. The set A is of measure zero in M if and only if for every coordinate coordinate system (V, h), the set $(\pi \circ h)(A \cap V)$ is of measure zero in $\mathbb{R}^k$.

For this definition to be meaningful, it must be the case that it is independent of the choice of coordinate system. Suppose that (V_1, h_1) and (V_2, h_2) are two coordinate systems about $x \in M$. It must be the case that for any $A \subset M$ the set $(\pi \circ h_1)(A \cap V_1 \cap V_2)$ is of measure zero if and only if $(\pi \circ h_2)(A \cap V_1 \cap V_2)$ is of measure zero. The following lemma guarantees this.

Lemma 1. Let W and be an open subset of $\mathbb{R}^k$, let $h : W \to \mathbb{R}^k$ be a continuously differentiable function and let A be a subset of $h(W)$. If $h^{-1}(A)$ is of measure zero, then A is of measure zero.

Proof. Let B be the set $x \in h^{-1}(A)$ such that $\det h' \neq 0$ and let C be the set of all $x \in h^{-1}(A)$ such that $\det h' = 0$. By Sard's theorem,[23] the measure of $h(C)$ is zero. So it suffices to show that if B is of measure zero, $g(B)$ is of measure zero. Since $\det h' \neq 0$ on B, by the inverse function theorem,[24] for every $x \in B$ there exists an open set $V \subset W$ containing x such that h restricted to V is invertible. Since h is continuously differentiable, we can choose the open set V so that the absolute value of the derivative of h on V is bounded. Choose c so that $|\det h'| < c$ on V. Because $\mathbb{R}^k$ has a countable basis, open balls with rational center and rational radius, there exist countably many such V that cover B. So, to complete the proof, it suffices to show that if $B \cap V$ is of measure zero, then $h(B \cap V)$ will be of measure zero. If $B \cap V$ is of measure zero, then given any $\varepsilon > 0$ there exists a set $D = \bigcup_i R_i$ containing $B \cap V$ such that each R_i is a closed rectangle in V of volume v_i and $\sum_i v_i < \varepsilon$. Let χ_D and $\chi_{h(D)}$ be the indicator functions on D and $h(D)$, respectively. We can make a change

---

[23]See Spivak (1965, page 72) for a statement and proof of Sard's theorem.

[24]See Spivak (1965, page 35) for a statement and proof of the inverse function theorem.

of variables[25] to integrate the indicator function $\chi_{h(D)}$, which gives

$$\int_{h(V)} \chi_{h(D)} = \int_V \left(\chi_{h(D)} \circ h \right) |\det h'|$$

$$= \int_V \chi_D |\det h'| < c\varepsilon.$$

Because $h(B \cap V) \subset h(D)$ and D can be chosen to make $\int_{h(V)} \chi_{h(D)}$ arbitrarily small, it must be the case that $h(B \cap V)$ is of measure zero. $\qquad\square$

Define $\iota : \mathbb{R}^k \rightarrow \mathbb{R}^n$ by $\iota(x_1, \cdots, x_k) = (x_1, \cdots, x_k, 0, \cdots, 0)$. If (U_1, h_1) and (U_2, h_2) are two coordinate systems about $x \in M$ and $A \subset M$, then

$$\left(\pi \circ h_2 \circ h_1^{-1} \circ \iota \right) (\pi \circ h_1) (A \cap U_1 \cap U_2) = (\pi \circ h_2) (A \cap U_1 \cap U_2)$$

and

$$\left(\pi \circ h_1 \circ h_2^{-1} \circ \iota \right) (\pi \circ h_2) (A \cap U_1 \cap U_2) = (\pi \circ h_1) (A \cap U_1 \cap U_2).$$

Since both $\pi \circ h_2 \circ h_1^{-1} \circ \iota$ and $\pi \circ h_1 \circ h_2^{-1} \circ \iota$ are continuously differentiable, it follows from Lemma 1 that $(\pi \circ h_1) (A \cap U_1 \cap U_2)$ will be of measure zero if and only if $(\pi \circ h_2) (A \cap U_1 \cap U_2)$ is of measure zero. Thus our definition of a measure zero set in a manifold is independent of the choice of coordinate systems.

We can easily extend Lemma 1 to manifolds, but first we need to define what we mean by a differentiable map between manifolds.

Definition 11. Let M_1 be a k_1-dimensional manifold and let M_2 be a k_2-dimensional manifold. A function $g : M_1 \rightarrow M_2$ is continuously differentiable at x if and only if for any coordinate systems (V_1, h_1) about x and (V_2, h_2) about $g(x)$, the composition $pi \circ h_2 \circ f \circ h_1^{-1} \circ \iota$ is continuously differentiable.

Lemma 2. Let M_1 and M_2 be k-dimensional manifolds, let $g : M_1 \rightarrow M_2$ be a continuously differentiable function, and let A be a subset of $h(M_1)$. If $h^{-1}(A)$ is of measure zero, then A is of measure zero.

Proof. Given any $x \in M_1$ and coordinate systems (V_1, h_1) about x and (V_2, h_2) about $g(x)$, the result follows by applying Lemma 1 to the function $\pi \circ h_2 \circ h_1^{-1} \circ \iota$. $\qquad\square$

This lemma was implicitly used in the proof of Theorem 6, and will be explicitly invoked in the proofs of Theorems 7 and 8. Next we analyze the set R in light of the results of this appendix.

---

[25]See Spivak (1965, page 67) for a discussion of integration by substitution in higher dimensions.

While R will not, in general, be a manifold, the set of un-normalized restricted structural parameters will be. We denote this set by $\mathfrak{R}$ and

$$\mathfrak{R} = \left\{ (A_0, A_+) \in U \mid Q_j f (A_0, A_+) e_j = 0 \text{ for } 1 \le j \le n \right\}. \qquad \text{(A1)}$$

The set R will be equal to $\mathfrak{R} \cap N$. Related to $\mathfrak{R}$ is the set $\tilde{\mathfrak{R}}$ defined by

$$\tilde{\mathfrak{R}} = \left\{ X \in f (U) \mid Q_j X e_j = 0 \text{ for } 1 \le j \le n \right\}. \qquad \text{(A2)}$$

The set $\tilde{\mathfrak{R}}$ is a open subset of a linear subspace of the set of $k \times n$ matices and $f (\mathfrak{R}) = \tilde{\mathfrak{R}}$. The dimension of $\tilde{\mathfrak{R}}$ is $nk - \sum_j q_j$. The implicit function theorem[26] implies that for every $u \in U$ there exist open sets $V_1 \subset f (U)$ containing $f (u)$, $V_2 \subset \mathbb{R}^{(m+n-k)n}$, and $V_3 \subset U$ containing u, and a continuously differentiable function $h : V_1 \times V_2 \to V_3$ with continuously differentiable inverse such that $f (h (v_1, v_2)) = v_1$ for every $(v_1, v_2) \in V_1 \times V_2$. From this representation, it is easy to see that h^{-1} maps $\mathfrak{R} \cap V_3$ to $(\tilde{\mathfrak{R}} \cap V_1) \times V_2$. While V_3 and h^{-1} do not quite satisfy the requirements of Definition 8, it is the case that there is is some linear transformation g of $V_1 \times V_2$ onto itself such that V_3 and $g \circ h^{-1}$ do. First, this implies that $\mathfrak{R}$ is a $(m + n) n - \sum_j q_j$ manifold in $\mathbb{R}^{(m+n)n}$. Second, if $\tilde{A} \subset \tilde{\mathfrak{R}}$ and $A = f^{-1} (\tilde{A})$, then $A \cap V_3$ will be of measure zero in $\mathfrak{R}$ if and only if $(\tilde{A} \cap V_1) \times V_2$ is of measure zero in $\tilde{\mathfrak{R}} \times V_2$. But $(\tilde{A} \cap V_1) \times V_2$ is of measure zero in $\tilde{\mathfrak{R}} \times V_2$ if and only if $\tilde{A} \cap V_1$ is of measure zero in $\tilde{\mathfrak{R}}$. Thus $A \subset \mathfrak{R}$ will be of measure zero in $\mathfrak{R}$ if and only if $\tilde{A} \subset \tilde{\mathfrak{R}}$ is of measure zero in $\tilde{\mathfrak{R}}$. We record this useful result in the following lemma.

Lemma 3. Let $\tilde{A} \subset \tilde{\mathfrak{R}}$. The set $A = f^{-1} (\tilde{A})$ is of measure zero in $\mathfrak{R}$ if and only if the set $\tilde{A}$ is of measure zero in $\tilde{\mathfrak{R}}$.

Finally, as has been noted, the set R is equal to $\mathfrak{R} \cap N$. When we say, as was done in Theorem 4, that a subset A of R is of measure zero in R, we take this to mean that A is of measure zero in the manifold $\mathfrak{R}$.

APPENDIX B. PROOF OF THEOREM 4

Before proceeding with the proof of Theorem 4, we prove the following lemma.

Lemma 4. For $1 \le j \le n$, let V_j be a linear subspace of $\mathbb{R}^m$ and let $V = V_1 \times \cdots \times V_n$. Define S to be the set of all $(v_1, \cdots, v_n) \in V$, whose span is of dimension strictly less than n. Either $S = V$ or S is a set of measure zero in V.

---

[26]See Spivak (1965, page 43) for a statement and proof of this theorem.

Proof. Let χ_S be the indicator function of S. To prove that S is of measure zero, it suffices to prove that χ_S is measurable and that $\int_V \chi_S = 0$. To show the latter, we divide $V_1 \times \cdots \times V_{n-1}$ into two sets A and B. We will show that A is of measure zero and that $\int_{V_n} \chi_S (v_1, \cdots, v_{n-1}, v_n) = 0$ for every $(v_1, \cdots, v_{n-1}) \in B$. Using Tonelli's Theorem,[27] this implies that

$$
\begin{aligned}
\int_V \chi_S &= \int_{V_1 \times \cdots \times V_{n-1}} \int_{V_n} \chi_S (v_1, \cdots, v_{n-1}, v_n) \\
&= \int_A \int_{V_n} \chi_S (v_1, \cdots, v_{n-1}, v_n) + \int_B \int_{V_n} \chi_S (v_1, \cdots, v_{n-1}, v_n) \\
&= 0
\end{aligned}
$$

as desired.

First we show that χ_S is a measurable function, which is equivalent to showing that S is measurable. We shall show that S is a closed subset of V, and hence measurable. The volume of the parallelepiped spanned by the vectors $v_1, \cdots, v_n$ is zero if and only $(v_1, \cdots, v_n) \in S$. Since the volume function is continuous and S is the inverse image of the closed set $\{0\}$, it is closed.

We now proceed with the heart of the proof, which proceeds by induction on n. When $n = 1$, $V = V_1$ and $S = \{0\}$. If the dimension of V is zero, then $S = V$, and if the dimension of V is positive, then S is of measure zero in V.

Now assume that the lemma holds for $n - 1$. Suppose that $S \neq V$, so there exists $(\tilde{v}_1, \cdots, \tilde{v}_n) \in V$ whose span is of dimension n. Define U to be the $m - 1$ dimensional subspace of $\mathbb{R}^m$ that is perpendicular $\tilde{v}_n$ and let ρ be the projection mapping onto U. For $1 \leq j \leq n - 1$, define $W_j = \rho (V_j)$ and let $W = W_1 \times \cdots \times W_{n-1}$. Define S_{n-1} to be the elements of W whose span is of dimension less than $n - 1$. By the induction hypothesis, either $S_{n-1} = W$ or S_{n-1} is of measure zero in W. Since the span of $\rho (\tilde{v}_1), \cdots, \rho (\tilde{v}_{n-1})$ is of dimension $n - 1$, $(\rho (\tilde{v}_1), \cdots, \rho (\tilde{v}_{n-1})) \notin S_{n-1}$ and $S_{n-1} \neq W$. Thus S_{n-1} is of measure zero in W. Let A be the set of all $(v_1, \cdots, v_{n-1}) \in V_1 \times \cdots \times V_{n-1}$ such that $(\rho (v_1), \cdots, \rho (v_{n-1})) \in S_{n-1}$. Because S_{n-1} is of measure zero in W, A must be of measure zero in V, as desired. All that remains to be shown is that $\int_{V_n} \chi_S (v_1, \cdots, v_{n-1}, v_n) = 0$ for every $(v_1, \cdots, v_{n-1}) \in B$, where B is the complement of A in $V_1 \times \cdots \times V_{n-1}$.

Fix $(v_1, \cdots, v_{n-1}) \in B$ and let C be the span of $v_1, \cdots, v_{n-1}$. Because $(v_1, \cdots, v_{n-1}) \in B$, the dimension of C is $n - 1$, and hence $\chi_S (v_1, \cdots, v_{n-1}, v_n) \neq 0$ if and only if

---

[27]See Royden (1968), page 270, for a statement and proof of Tonelli's Theorem.

$v_n \in C$. Thus $\int_{V_n} \chi_S (v_1, \cdots, v_{n-1}, v_n) = 0$ if and only if $V_n \cap C$ is of measure zero in V_n. Since $V_n \cap C$ is either all of V_n or a set of measure zero in V_n, it suffices to show that there exists an element of V_n that is not in C. Again, since $(v_1, \cdots, v_{n-1}) \in B$, the vector $\tilde{v}_n$ is not in C. $\qquad \square$

We are now ready to complete the proof of Theorem 4.

Proof of Theorem 4. Let W be the complement of K in R. In terms of W, Theorem 4 states that either $W = R$ or W is of measure zero in R. Let $\tilde{W} = f(W)$ and $\tilde{R} = f(R)$. From the definition of K given by (7) and the definition of R given by (5), one can easily check that $f^{-1}(\tilde{W}) = W$ and $f^{-1}(\tilde{R}) = R$. So, by Lemma 3, we have that if $\tilde{W}$ is of measure zero in $\tilde{\mathfrak{R}}$, then W is of measure zero in $\mathfrak{R}$. Thus, if suffices to prove that either $\tilde{W} = \tilde{R}$ or $\tilde{W}$ is of measure zero in $\tilde{\mathfrak{R}}$.

Let $\tilde{V}$ be the set of all $k \times n$ matrices X such that $Q_j X e_j = 0$ for $1 \le j \le n$, and let $\tilde{W}_i$ be the set of all matrices $X \in \tilde{V}$ such that $rank (M_i (X)) < n$. Since $\tilde{R} \subset \tilde{V}$ and $\tilde{W} = \bigcap_{i=1}^n \tilde{W}_i \cap \tilde{R}$, it suffices to show that either $\tilde{W}_i = \tilde{V}$ or $\tilde{W}_i$ is of measure zero in $\tilde{V}$. Because of the block structure of $M_i(\cdot)$, if $X = [x_1, \cdots, x_n]$, then $X \in \tilde{W}_i$ if and only if $rank (Q_i [x_{i+1}, \cdots, x_n]) < n - i$. The results now follow from Lemma 4 with V_j defined to be the set of all vectors of the form $Q_i x$ where $Q_{i+j} x = 0$. $\qquad \square$

From the proof of Theorem 4, it easily follows that if

$$K_j = \{ (A_0, A_+) \in R \mid rank (M_j (f (A_0, A_+))) = n \}, \qquad (A3)$$

then either K_j is empty or the complement of K_j in R is of measure zero in R. We record this in the following lemma that will be used in Appendix C.

Lemma 5. Either K_j is empty or the complement of K_j in R is of measure zero in R.

APPENDIX C. PROOF OF THEOREMS 7 AND 8

We prove Theorem 8 first, and then Theorem 7. We proceed via a sequence of lemmas.

Lemma 6. If $q_j = n - j$ for $1 \le j \le n$, then for every $(A_0, A_+) \in U$, there exists a $P \in O(n)$ such that $(A_0, A_+) \in R$.

Proof. Let $(A_0, A_+) \in U$ and let $X = f(A_0, A_+)$. Because the rank of $Q_1 X$ is at most $q_1 = n - 1$, there exists a vector $p_1 \in \mathbb{R}^n$ of length one such that $Q_1 X p_1 = 0$. Now

assume that $p_1, \cdots, p_{i-1}$ are orthonormal vectors in $\mathbb{R}^n$ such that $Q_j X p_j = 0$ for $1 \leq j < i$. Define

$$
\hat{Q}_i = \begin{bmatrix} Q_i X \\ p_1' \\ \vdots \\ p_{i-1}' \end{bmatrix}.
$$

Since the rank of $\hat{Q}_i$ is at most $q_i + i - 1 = n - 1$, there exists a $p_i \in \mathbb{R}^n$ of length one such that $\hat{Q}_i p_i = 0$. Thus $p_1, \cdots, p_i$ are orthonormal vectors in $\mathbb{R}^n$ such that $Q_j X p_j = 0$ for $1 \leq j \leq i$. So, we can continue until we have constructed an orthogonal matrix $P = \begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}$ such that $XP \in \tilde{\mathfrak{R}}$. From the definition of the normalization rule N, we know that there is a diagonal matrix D with plus or minus ones along the diagonal such that $XPD \in \tilde{R}$ or $(A_0 PD, A_+ PD) \in R$ as required. $\qquad \square$

Lemma 7. If $q_j = n - j$ for $1 \leq j \leq n$, then there exists $(A_0, A_+) \in R$ such that $M_j(A_0, A_+)$ is of rank n for $1 \leq j \leq n$.

Proof. In light of Lemma 5, it suffices to construct a matrix $X_i \in \tilde{\mathfrak{R}}$ such that $M_i(X_i)$ is of rank n. It follows from Lemma 6 that $\tilde{\mathfrak{R}}$ is non-empty, so let $Y = [y_1, \cdots, y_n]$ be any matrix in $\tilde{\mathfrak{R}}$. Let $Y_i^j = [y_i, \cdots, y_j]$. The matrix $M_i(Y)$ is of rank n if and only if the matrix $Q_i Y_{i+1}^n$ is of rank $n - i$. Let V_j be the column space of $Q_i Y_{i+1}^{i+j}$. If the dimension of V_j is j for $1 \leq j \leq n - i$, then $M_i(Y)$ is of rank n. If this is not the case, let j be the first index such that the dimension of V_j is less than j. Because the dimension of V_{j-1} is $j - 1$ and the dimension of the null space of Q_i is i, the dimension of the set of all vectors $v \in \mathbb{R}^n$ such that $Q_i v \in V_{j-1}$ is $i + j - 1$. Since the dimension of the null space of Q_{i+j} is $i + j$, there is a non-zero vector v in the null space of Q_{i+j} such that $Q_i v$ is not in V_{j-1}. Because $f(U)$ is an open set and $Y \in \tilde{\mathfrak{R}}$, there exists $\varepsilon > 0$ such that if we replace y_{i+j} by $y_{i+j} + \varepsilon v$, then the resulting matrix will be an element of $\tilde{\mathfrak{R}}$ and have the property that the dimension of the column space of $Q_i [y_{i+1}, \cdots, y_{i+j-1}, y_{i+j} + \varepsilon v]$ will be of dimension j. So, starting with any $Y \in \tilde{\mathfrak{R}}$, we can sequentially modify Y until we arrive at a matrix $X_i \in \tilde{\mathfrak{R}}$ such that $M_i(X_i)$ is of rank n. $\qquad \square$

Given these two lemmas, the following is now easy to show.

Lemma 8. If $q_j = n - j$ for $1 \leq j \leq n$, then the SVAR model is exactly identified.

Proof. Let G be defined by (9), and let W be the set of all $(A_0, A_+) \in R$, which are not globally identified. We must show that G is of measure zero. It follows from Lemma 5, Theorem 4 and Theorem 2, that W is of measure zero in $\mathfrak{R}$. Since $q_j = n - j$, the dimension of $\mathfrak{R}$ is $(m+n)n - n(n-1)/2$ and so the dimension of $\mathfrak{R} \times O(n)$ is $(m+n)n$. Let $h : \mathfrak{R} \times O(n) \to U$ be the continuously differentiable function which maps $(A_0, A_+) \times P$ to $(A_0 P, A_+ P)$. It follows from Lemma 6 that $h^{-1}(G) = W \times O(n)$. Since $W \times O(n)$ is of measure zero, by Lemma 2 the set G will be of measure zero. $\qquad\qquad\square$

This finishes one direction in the proof of Theorem 8. To prove the other direction, the following lemma is key.

Lemma 9. Let $i \leq n \leq k$. If $V_1, \cdots, V_i$ are subspaces of $\mathbb{R}^k$ with the dimension of V_j equal to d_j for $1 \leq j \leq i$ and for all $k \times n$ matrices X there exist orthonormal vectors $p_1, \cdots, p_i$ in $\mathbb{R}^n$ such that $Xp_j \in V_j$ for $1 \leq j \leq i$, then there exists a j such that $1 \leq j \leq i$ and $d_j \geq k - n + i$.

Before giving the formal proof of this lemma, we explain the geometry behind the result. Consider the case $i = n = k = 3$. Since the implication of the lemma is that at least one of the subspaces must be all of $\mathbb{R}^3$, suppose that each of subspaces V_j is at most a plane. It is easy to see that there exists a line L in $\mathbb{R}^3$ that intersects each V_j only at the origin. Let K be the plane through the origin that is perpendicular to L and let Y be the 3×3 matrix which projects $\mathbb{R}^3$ onto K along L. While Y is not invertible, there are invertible matrices that are arbitrarily close to Y. If X is an invertible matrix such that X^{-1} is close to Y, then the subspace $X^{-1}V_j$ will almost lie in K. This depends crucially on the fact that L intersects V_j only at the origin. Since it is not possible to have three orthonormal vectors almost lie in a plane, it cannot be the case that for all 3×3 matrices X there exist orthonormal vectors p_j such that $p_i \in X^{-1}V_j$. Thus the theorem holds for $i = n = k = 3$. The formal proof simply makes rigorous what is geometrically intuitive in this special case.

Proof. The proof will use the following three facts.

(1) For $0 \leq d \leq k$, there exists a subspace U of $\mathbb{R}^k$ of dimension d such that the dimension of of $V_j \cap U$ is equal to $\max\{0, d + d_j - k\}$ for $1 \leq j \leq i$.

(2) Let W be a $i-1$ dimensional subspace of $\mathbb{R}^n$. There exists a $\delta > 0$ such that there cannot be i orthonormal vectors in the set

$$S_{W,\delta} = \left\{ w + u \in \mathbb{R}^k \mid w \in W \text{ and } \|u\| < \delta \right\}.$$

(3) Let U and V be subspaces of $\mathbb{R}^n$ such that $U \cap V = \{0\}$ and let W be the perpendicular complement of U. For every $\varepsilon > 0$, let $X_{W,\varepsilon}$ be the linear transformation that fixes W and maps each $u \in U$ to εu. For every $\delta > 0$ there exists a $\gamma > 0$ such that for all $\gamma > \varepsilon > 0$ if $X_{W,\varepsilon}^{-1} v \in V$ and $\|v\| = 1$, then $v \in S_{W,\delta}$.

Using (1), we see that there exists a subspace U of $\mathbb{R}^k$ of dimension n such that the dimension of $V_j \cap U$ is of dimension $\tilde{d}_j = \max \{0, n + d_j - k\}$. Let X be an $k \times n$ matrix whose column space if equal to U. Let

$$\tilde{V}_j = \left\{ x \in \mathbb{R}^n \mid Xx \in V_j \right\}.$$

The dimension of $\tilde{V}_j$ is $\tilde{d}_j$, and if Y is any $n \times n$ matrix, then there exist orthonormal vectors $p_1, \cdots, p_i$ in $\mathbb{R}^n$ such that $XYp_j \in V_j$, or $Yp_j \in \tilde{V}_j$ for $1 \leq j \leq i$. If the lemma were true for $n = k$, then this would imply that there would exist a j such that $\tilde{d}_j \geq i$. This would imply that $n + d_j - k \geq i$ or $d_j \geq k - n + i$. Hence it suffices to prove the lemma for $n = k$.

If $n = k$ and $d_j < i$ for $1 \leq j \leq i$, then (1) would imply that there exists a subspace U of $\mathbb{R}^n$ of dimension $n - i + 1$ such that $V_j \cap U = \{0\}$ for $1 \leq j \leq i$. If W is the subspace of dimension $i - 1$ that is perpendicular to U, then (3) would imply that for every $\delta > 0$, there exists an $\gamma_j > 0$ such that for all $\gamma_j > \varepsilon > 0$ if $X_{W,\varepsilon}^{-1} v \in V_j$ and $\|v\| = 1$, then $v \in S_{W,\delta}$. But then (2) would contradict the fact that there must exist orthonormal vectors $p_1, \cdots, p_i$ in $\mathbb{R}^n$ such that $X_{W,\varepsilon}^{-1} p_j \in V_j$ for $1 \leq j \leq i$. So $d_j \geq i$ for some j as required by the lemma when $n = k$.

All that remains to be shown is (1), (2), and (3).

(1) Because $\dim (U + V_j) = \dim (U) + \dim (V_j) - \dim (V_j \cap U)$, (1) is equivalent to showing that there exists a subspace U of dimension d such that

$$\dim (U + V_j) = \min \{ \dim (U) + \dim (V_j), k \}. \tag{A4}$$

When U is of dimension 0, (A4) is trivially true. Assume that there exists a subspace U of dimension $d - 1$ for which (A4) holds. We construct a subspace $\tilde{U}$ of dimension d for which (A4) holds. For those j for which $\dim (U + V_j) < k$, the subspace $U + V_j$

is of measure zero in $\mathbb{R}^k$. So, there exists a $u \notin U$ such that $u \notin U + V_j$ for all j such that $\dim(U + V_j) < k$. Define $\tilde{U}$ to be the subspace spanned by U and u. For those j for which $\dim(U + V_j) < k$, we have that $\dim(\tilde{U} + V_j) = \dim(U + V_j) + 1$ and for those j for which $\dim(U + V_j) = k$, we have that $\dim(\tilde{U} + V_j) = k$. Since $\dim(\tilde{U}) = \dim(U) + 1$, it is easy to verify that (A4) holds for $\tilde{U}$.

(2) Choose $\delta < 1/(i\sqrt{n})$. Suppose there were $v_1, \cdots, v_i$ in $S_{W,\delta}$ that were orthonormal. Since the v_j are in $S_{W,\delta}$, write $v_j = w_j + \delta u_j$ where $w_j \in W$, u_j is perpendicular to W, and $\|u_j\| < 1$. Let X be the $n \times i$ matrix $[w_1, \cdots, w_i]$, let Y be the $n \times i$ matrix $[v_1, \cdots, v_j]$, and let Z be the $n \times i$ matrix $[u_1, \cdots, u_i]$. Because the w_j are in W and the u_j are perpendicular to W, $X'Z = 0$ and $Z'X = 0$. Because the v_j are orthonormal, $Y'Y = I_i$. So, $I_i = X'X + \delta^2 Z'Z$. Because the w_j are in a $i - 1$ dimensional space, the matrix $X'X$ is singular and so there is a $v \in \mathbb{R}^i$ of length one such that $v'X'Xv = 0$. Because elements of Z and v are less than or equal to one in absolute value, each element of Zv is less than or equal to i in absolute value. Thus $1 = v'v = \delta^2 z' Z' Zv \le \delta^2 n i^2 < 1$, which is a contradiction. Thus there cannot be $v_1, \cdots, v_i$ in $S_{W,\delta}$ that are orthonormal.

(3) If this were not true, then there would exist a $\delta > 0$ and a sequence of v_ℓ and ε_ℓ such that the ε_ℓ tend to zero and $X_{W,\varepsilon_\ell}^{-1} v_\ell \in V$, $\|v_\ell\| = 1$, and $v_\ell \notin S_{W,\delta}$. We can write $v_\ell = u_\ell + w_\ell$ where $u_\ell \in U$ and $w_\ell \in W$. Since $X_{W,\varepsilon_\ell}^{-1} v_\ell = \frac{1}{\varepsilon_\ell} u_\ell + w_\ell \in V$, we have that $u_\ell + \varepsilon_\ell w_\ell \in V$. Since $\|v_\ell\| = 1$, u_ℓ and w_ℓ are orthogonal, and $v_\ell \notin S_{W,\delta}$, we have $\|w_\ell\| \le 1$ and $\delta \le \|u_\ell\| \le 1$. Thus, $\lim_{\ell \to \infty} \varepsilon_\ell w_\ell = 0$, and hence some subsequence of the u_ℓ converges to a non-zero element of $U \cap V$, which is a contradiction. $\qquad \square$

Lemma 10. If the SVAR model is exactly identified, then $q_j = n - j$ for $1 \le j \le n$.

Proof. We proceed by contradiction. Suppose that it is not the case that $q_j = n - j$ for $1 \le j \le n$. Since the model is exactly identified, Rothenberg (1971)'s necessary condition for local identification implies that the total number of restrictions, $\sum_{i=1}^{n} q_i$, must be at least $(n-1)n/2$. Since we are assuming that it is not the case that $q_j = n - j$ for $1 \le j \le n$, this implies that there must be at least one index i with $q_i > n - i$. Since the q_j are in decreasing order, this implies that $q_j > n - i$ for $1 \le j \le i$. If we define V_j to be the null space of the matrix Q_j, then the dimension of V_j will be $n - q_j < i$. By Lemma 9, there exists a $k \times n$ matrix X for which there are no orthonormal vectors $p_1, \cdots, p_i$ in $\mathbb{R}^n$ such that $X p_j \in V_j$ for $1 \le j \le i$. Let $M(k, n)$

be the set of all $k \times n$ matrices and let

$$\tilde{H} = \left\{ X \in M(k,n) \mid XP \notin \tilde{\mathfrak{R}} \text{ for all } P \in O(n) \right\}. \tag{A5}$$

We have shown that $\tilde{H}$ is non-empty. The proof will be complete if we can show that $\tilde{H}$ is open. To see this, note since $f(U)$ is dense, $\tilde{H} \cap f(U)$ would be a non-empty open set as would $f^{-1}(\tilde{H})$. Since open sets are of positive measure, this would contradict the fact that the SVAR is exactly identified.

To show that $\tilde{H}$ is open, it suffices to show that if the sequence $X_j \notin \tilde{H}$ converges to X, then $X \notin \tilde{H}$. If $X_j \notin \tilde{H}$ then there would exist $P_j \in O(n)$ such that $X_j P_j \in \tilde{\mathfrak{R}}$. Since $O(n)$ is a compact set, some subsequence of the P_j converges to some orthogonal matrix P. Since $\tilde{\mathfrak{R}}$ is a closed subset, this implies that $XP \in \tilde{\mathfrak{R}}$ as desired. □

Proof of Theorem 8. This result follows directly for Lemmas 8 and 10. □

Proof of Theorem 7. If the SVAR is exactly identified, then by Theorem 8 $q_j = n - j$ for $1 \le j \le n$. Clearly, the number of restrictions, $\sum_{i=1}^{n} q_i$, is equal to $(n-1)n/2$ and by Lemma 7, the rank condition in Theorem 2 is satisfied for some $(A_0, A_+) \in R$.

On the other hand, since $rank(M_j(X)) \le q_j + j$, if rank condition in Theorem 2 is satisfied for some $(A_0, A_+) \in R$, it must be the case that $q_j \ge n - j$ for $1 \le j \le n$. If it is also the case that the total number of restrictions is equal to $(n-1)n/2$, then $q_j = n - j$ for $1 \le j \le n$. So, by Theorem 8, the SVAR is exactly identified. □

APPENDIX D. PROOF OF THEOREM 5

For linear restrictions on the contemporaneous structural coefficients, the system will be triangular if there is an ordering of the equations and a linear transformation of the variables such that the A_0 is triangular. Since our convention is that $q_1 \ge \cdots \ge q_n$, the equations (columns) will be ordered correctly, and so the system will be triangular if there is a linear transformation of the variables such that A_0 is upper triangular. The following lemma gives a characterization that is more useful for our purposes.

Lemma 11. Suppose the transformation $f(\cdot)$ is given by $f(A_0, A_+) = A_0$. Let

$$V_j = \left\{ v \in \mathbb{R}^n \mid Q_j v = 0 \right\}, \tag{A6}$$

for $1 \leq j \leq n$. The SVAR will be triangular if and only if $V_1 \subset \cdots \subset V_n$ and the dimension of V_j is j.

Proof. Suppose that $V_1 \subset \cdots \subset V_n$ and that the dimension of V_j is j. This implies that there exists vectors $v_1, \cdots, v_n$ such that $v_1, \cdots, v_j$ forms a basis for V_j. If $T = [v_1, \cdots, v_n]$, then transforming the variables by T^{-1} will produce an upper triangular system, and thus the system will be triangular.

Now suppose that the system is triangular and that the transformation that produces an upper triangular system is T^{-1}. If $T = [v_1, \cdots, v_n]$, then a basis for V_j will be $v_1, \cdots, v_j$. From this, it follows that $V_1 \subset \cdots \subset V_n$ and that the dimension of V_j is j. $\qquad\square$

We are now ready to prove Theorem 5.

Proof of Theorem 5. Consider an exactly identified SVAR with restrictions on the contemporaneous coefficients given by the matrices $Q_1, \cdots, Q_n$, which are decreasing in rank. Let G be defined by (9). Assume that G is empty. Let V_j be defined by (A6). We show that $V_1 \subset \cdots \subset V_n$. If this were not the case, then we would construct an A_0 and a non-diagonal orthogonal matrix P such that both A_0 and $A_0 P$ satisfy the restrictions, which would contradict the fact that G is empty. So assume that it is not the case that $V_1 \subset \cdots \subset V_n$ and let $\hat{k}$ be the first k such that V_k is not a subset of V_{k+1}. Note that $\hat{k} < n - 1$ because $V_n = \mathbb{R}^n$. We first recursively construct the A_0 and the then the P.

Let a_1 be any non-zero element of V_1. Now assume that $a_1, \cdots, a_k$ have been constructed such that $a_j \in V_j$ for $1 \leq j \leq k$ and $a_i' a_j = 0$ for $1 \leq i < j \leq k$. There exists a non-zero a_{k+1} such that $Q_{k+1} a_{k+1} = 0$ and $a_i \prime a_{k+1} = 0$ for $1 \leq i \leq k$. Such a vector exists because we have imposed at most $q_{k+1} + k = n - 1$ restrictions on $\mathbb{R}^n$. So we recursively constructed $A_0 = [a_1, \cdots, a_n]$.

We now recursively construct the orthogonal matrix P. For $1 \leq j \leq \hat{k}$, let $p_j = e_j$, where e_j is the jth column of the $n \times n$ identity matrix. Since $V_1 \subset \cdots \subset V_{\hat{k}}$, $a_1, \cdots, a_{\hat{k}}$ forms a basis for $V_{\hat{k}}$. Let $V^\perp$ be the subspace of $V_{\hat{k}+1}$ that is perpendicular to $V_{\hat{k}}$. Since $V_{\hat{k}}$ is not a subset of $V_{\hat{k}+1}$, the dimension of $V_{\hat{k}}$ is $\hat{k}$, and the dimension of $V_{\hat{k}+1}$ is $\hat{k}+1$, the dimension of $V^\perp$ must be at least two. The vector $a_{\hat{k}+1} \in V^\perp$. Let $\hat{a}_{\hat{k}+1}$ be any element of $V^\perp$ of length one that is not a multiple of $a_{\hat{k}+1}$. Because $a_{\hat{k}+1}, \cdots, a_n$ forms a basis for the space perpendicular to $a_1, \cdots, a_{\hat{k}}$, $\hat{a}_{\hat{k}+1}$ is a linear combination of $a_{\hat{k}+1}, \cdots, a_n$. This implies that there exists a $p_{\hat{k}+1}$ such that $\hat{a}_{\hat{k}+1} = A_0 p_{\hat{k}+1}$, the

first $\hat{k}$ elements pf $p_{\hat{k}+1}$ are zero, and at least one of the last $n - \hat{k} - 1$ elements of $p_{\hat{k}+1}$ is non-zero. Now assume that an orthonormal set $p_1, \cdots, p_k$ has been constructed, where $\hat{k} + 1 \leq k \leq n$. Choose $p_{k+1} \in \mathbb{R}^n$ of length one so that $Q_{k+1}A_0 p_{k+1} = 0$ and $p_j' p_{k+1} = 0$ for $1 \leq j \leq k$. Such an element exists because we have imposed at most $q_{k+1} + k = n - 1$ restrictions. The matrix $P = [p_1, \cdots, p_n]$ is the required non-diagonal orthonormal matrix. $\qquad \square$

APPENDIX E. EXISTING ALGORITHMS FOR SVARS WITH SIGN RESTRICTIONS

Faust (1998), Canova and De Nicoló (2002), and Uhlig (2005) propose an alternative approach to SVAR modeling by imposing sign restrictions on impulse responses themselves. Although Faust (1998), Canova and De Nicoló (2002), and Uhlig (2005) have the same basic idea, their algorithms for implementation are distinctively different. In this section, we briefly review each of these three algorithms and highlight the problem related to computational efficiency.

E.1. **Faust's algorithm.** Faust (1998) presents a way to check the robustness of any claim from an SVAR. All possible identifications are checked against the claim, subject to the restriction that the SVAR produces the impulse response functions with "correct" signs.

Faust (1998) shows that this problem is equivalent to solving an eigenvalue problem $\sum_{i=0}^{M} \frac{R!}{i!(R-i)!}$ times, where R is the number of sign restrictions and $M = \max(n - 1, R)$. As Faust (1998) recognizes, this algorithm may be infeasible for a large VAR system.

E.2. **Canova and De Nicoló's algorithm.** Canova and De Nicoló (2002) also study SVARs with sign restrictions imposed on impulse response. Their algorithm is based on Theorem 10, whose proof is provided below.

Proof. The proof of Theorem 10 is provided as follows.

The matrix $Q_{i,j}\left(\theta_{i,j}\right)$ is a Givens rotation. The proof is simply a careful application of the algorithm for obtaining the QR decomposition via Givens rotations. We follow Algorithm 5.2.2 of Golub and Van Loan (1996). The basic idea is that multiplying an $n \times n$ matrix X on the left by a Givens rotation $Q_{i,j}\left(\theta_{i,j}\right)$ has the effect of rotating the ith and j^{th} rows of X in the counter-clockwise direction by $\theta_{i,j}$ radians and leaving all the other rows of X fixed. The rotation takes place in the plane spanned by the ith and j^{th} rows. We can choose $\theta_{i,j}$ so that the jth row is rotated so

that its ith element becomes zero. The exact details of choosing such a $\theta_{i,j}$ are given in Section 5.1.8 of Golub and Van Loan (1996). Thus by successively multiplying on the left by the appropriate Givens rotations, we can transform, column by column, any matrix into an upper triangular matrix. The $\theta_{i,j}$ are not unique since if a rotation of $\theta_{i,j}$ radians will place a zero in the ith position of the jth row, then so will a rotation by an additional π radians (180 degrees). Thus we can choose $\theta_{i,j}$ to be between 0 and π when $i < j < n$. When $j = n$, we choose the rotation that not only makes the ith coordinate of the jth row zero, but also makes the ith coordinate of the ith row non-negative. Applying this algorithm to the matrix P' allows us to write

$$\left(\prod_{i=1}^{n-1} \prod_{j=i+1}^{n} Q_{i,j}\left(\theta_{i,j}\right) \right) P' = S,$$

where S is upper triangular. Since P and the $Q_{i,j}\left(\theta_{i,j}\right)$ are all orthogonal, so is S. The only upper triangular orthogonal matrices are diagonal with plus or minus one along the diagonal. Because of our choice of rotations, all the diagonal elements, except the last, must be non-negative and hence equal to one. The results now follow by multiplying P on the right and $S = S^{-1}$ on the left. □

Based on this theorem, Canova and De Nicoló (2002) propose the following algorithm for an SVAR with sign restrictions.

Algorithm 3.

(1) Draw a set of unrestricted parameters (A_0, A_+) from the posterior distribution.

(2) For each draw of (A_0, A_+), compute (B, Σ) and perform the Choleski decomposition of Σ to get A_0^*.

(3) Determine a grid on $\theta_{i,j}$ for the set of all orthogonal matrices $Q_{i,j}(\theta_{i,j})$ in Theorem 10.

(4) Perform a grid search to find an orthogonal matrix P such that the impulse responses generated from $A_0^* P$ and B satisfy all the sign restrictions.

Theorem 10 allows for different ways to design a grid. Because the space of all orthogonal $n \times n$ matrices has the dimension $n(n-1)/2$, any grid that divides the interval $[-\pi/2, \pi/2]$ with M points (or on the interval $[-\pi, \pi]$ with $2M$ points) implies a search over $2M^{n(n-1)/2}$ points in the space of all orthogonal $n \times n$ matrices. Thus, it is infeasible to perform this grid search for a large value of n.

E.3. **Uhlig's algorithm.** Uhlig (2005)'s algorithm for estimating SVARs with sign restrictions on impulse responses is stated as follows.

Algorithm 4.

(1) Draw a set of unrestricted parameters (A_0, A_+) from the posterior distribution.

(2) For each draw of (A_0, A_+), compute (B, Σ).

(3) Compute the eigenvectors of Σ normalized so as to form an orthonormal basis of $\mathbb{R}^n$. Denote these eigenvectors by x_j for $j = 1, \ldots, n$ and let λ_j for $j = 1, \ldots, n$ be the corresponding eigenvectors.

(4) Draw

$$(\alpha_1, \ldots, \alpha_{n-1}), \text{ with } \sum_{j=1}^{n} \alpha_j^2 = 1$$

from a uniform distribution over the $(n-1)$-dimensional sphere.

(5) Construct the impact impulse response vector a to a particular structural shock under study according to

$$a = \sum_{j=1}^{n} \left(\alpha_j \sqrt{\lambda_j} \right) x_j.$$

(6) Construct a matrix C such that $CC' = \Sigma$ and a is a column of C.

(7) Use C and B to generate the impulse responses.

(8) If these impulse responses satisfy the sign restrictions, keep the draw; otherwise, repeat Steps (1)-(7).

This algorithm works well for sign restrictions on the impulse responses to *one* structural shock. If sign restrictions concern impulse responses to a number of structural shocks, as studied by Gambetti, Pappa, and Canova (forthcoming), one has to construct different a's recursively and the algorithm quickly becomes inefficient. Our algorithm based on the Householder transformation is more efficient in dealing with sign restrictions on impulse responses to a number of structural shocks.

References

BEKKER, P. A., AND D. POLLOCK (1986): "Identification of Linear Stochastic Models with Covariance Restrictions," *Journal of Econometrics*, 31, 179–208.

BENATI, L., AND H. MUMTAZ (2006): "U.S. Evolving Macroeconomic Dynamics: A Structural Investigation," Unpublished Manuscript, European Central Bank and Bank of England.

BERNANKE, B. S., AND I. MIHOV (1998): "Measuring Monetary Policy," *Quarterly Journal of Economics*, 113(3), 869–902.

BLANCHARD, O. J., AND S. FISCHER (1989): *Lectures on Macroeconomics*. MIT Press, Cambridge, Massachusetts.

BLANCHARD, O. J., AND D. QUAH (1993): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 83(3), 655–673.

CANOVA, F., AND G. DENICOLÓ (2002): "Monetary Disturbances Matter for Business Fluctuations in the G-7," *Journal of Monetary Economics*, 49(6), 1121–1159.

CANOVA, F., AND L. GAMBETTI (2004): "Structural Changes in the US Economy: Bad Luck or Bad Policy?," Manuscript, Universitat Pompeu Fabra.

CHRISTIANO, L., M. EICHENBAUM, AND C. EVANS (2005): "Nominal Rigidities and the Dynamics Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (1996): "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics*, 78(1), 16–34.

COCHRANE, J. H. (2006): "Identification and Price Determination with Taylor Rules: A Critical Review," Manuscript, University of Chicago.

COGLEY, T., AND T. SARGENT (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *Review of Economic Dynamics*, 8, 262–302.

CUSHMAN, D. O., AND T. ZHA (1997): "Identifying Monetary Policy in a Small Open Economy under Flexible Exchange Rates," *Journal of Monetary Economics*, 39, 433–448.

DHRYMES, P. J. (1978): *Introductory Econometrics*. Springer-Verlag, New York.

——— (1994): *Topics in Advanced Econometrics*. Springer-Verlag, New York.

DUFOUR, J.-M., AND C. HSIAO (forthcoming): "Identification," in *The New Palgrave Dictionary of Economics*, ed. by L. Blume, and S. Durlauf. Palgrave Macmillan, Basingstoke, Hampshire, England, forthcoming.

EICHENBAUM, M., AND C. EVANS (1995): "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates," *Quarterly Journal of Economics*, 110, 975–1009.

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

FERNÁNDEZ-VILLAVERDE, J., AND J. F. RUBIO-RAMÍREZ (2004): "Comparing dynamic equilibrium models to data: a Bayesian approach," *Journal of Econometrics*, 123, 153–187.

FISHER, F. M. (1966): *The Identification Problem in Econometrics*. McGraw-Hill, Inc.

FRY, R., AND A. PAGAN (2007): "Some Issues in Using Sign Restrictions for Identifying Structural VARs," National Centre for Econometric Research Working Paper # 14.

FUBAC, M., D. F. WAGGONER, AND T. ZHA (2007): "Local and Global Identification of DSGE Models: A Simultaneous-Equation Approach," Manuscript, Federal Reserve Bank of Atlanta.

GALÍ, J. (1992): "How Well Does the IS-LM Model Fit Postwar U.S. Data?," *Quarterly Journal of Economics*, 107(2), 709–738.

GALÍ, J. (1999): "Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?," *American Economic Review*, 89(1), 249–271.

GAMBETTI, L., E. PAPPA, AND F. CANOVA (forthcoming): "The Structural Dynamics of US Output and Inflation: What Explains the Changes?," *Journal of Money, Credit, and Banking*.

GEWEKE, J. (2005): *Contemporary Bayesian Econometrics and Statistics*. John Wiley & Sons, Inc., Hoboken, New Jersey.

GEWEKE, J. F. (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication," *Econometric Reviews*, 18, 1–126.

GIANNINI, C. (1992): *Topics in Structural VAR Econometrics*. Springer-Verlag, New York, NY.

GOLUB, G. H., AND C. F. VAN LOAN (1996): *Matrix Computations (Third Edition)*. The Johns Hopkins University Press, Baltimore and London.

GORDON, D. B., AND E. M. LEEPER (1994): "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification," *Journal of Political Economy*, 102(6), 1228–1247.

GREENE, W. H. (1993): *Econometric Analysis*. Macmillan Publishing Company, New York, second edn.

HAMILTON, J. D. (1994): *Times Series Analsis*. Princeton University Press, Princeton, NJ.

HAMILTON, J. D., D. F. WAGGONER, AND T. ZHA (2007): "Normalization in Econometrics," *Econometric Reviews*, 26(2-4), 221–252.

HAUSMAN, J. A., AND W. E. TAYLOR (1983): "Identification in Linear Simultaneous Equations Models with Covariance Restrictions: An Instrumental Variables Interpretation," *Econometrica*, 51(5), 1527–1549.

HSIAO, C. (1983): "Identification," in *Handbook of Econometrics I*, ed. by Z. Griliches, and M. M. Intrilligator, vol. 1, pp. 224–283. North-Holland.

———— (2001): "Identification and Dichotomization of Long- and Short-Run Relations of Cointegrated Vector Autoregressive Models," *Econometric Theory*, 17(5), 889–912.

INOUE, A., AND L. KILIAN (2002): "Bootstrapping Autoregressive Processes with Possible Unit Roots," *Econometrica*, 70(1), 377–391.

ISKREV, N. (2007): "How much do we learn from the estimation of DSGE models? A case study of identifrcation issues in a New Keynesian business cycle model," Manuscript, University of Michigan.

KILIAN, L. (1998): "Small-Sample Confidence Intervals for Impulse Response Functions," *Review of Economics and Statistics*, 80, 186–201.

KIM, S. (1999): "Does Monetary Policy Matter in the G-7 Countries? Using Common Identifying Assumptions about Monetary Policy across Countries," *Journal of International Economics*, 48, 387–412.

KIM, S., AND N. ROUBINI (2000): "Exchange rate anomalies in the industrial countries: A solution with a structural VAR approach," *Journal of Monetary Economics*, 45(3), 561–586.

KING, R. G., C. I. PLOSSER, J. H. STOCK, AND M. W. WATSON (1991): "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81(4), 819–840.

LEEPER, E. M., C. A. SIMS, AND T. ZHA (1996): "What Does Monetary Policy Do?," *Brookings Papers on Economic Activity*, 2, 1 – 78.

LEEPER, E. M., AND T. ZHA (2003): "Modest Policy Interventions," *Journal of Monetary Economics*, 50(8), 1673–1700.

PEERSMAN, G., AND F. SMETS (2003): *The Monetary Transmission Mechanism in the Euro Area: More Evidence from VAR Analysis*. Cambridge University Press, Cambridge.

PESAVENTO, E., AND B. ROSSI (2006): "Small-Sample Confidence Intervals for Multivariate Impulse Response Functions at Long Horizons," *Journal of Applied Econometrics*, 21(8), 1135–1155.

PRIMICERI, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

ROTHENBERG, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39(3), 577–591.

RUBIO-RAMÍREZ, J. F., D. F. WAGGONER, AND T. ZHA (2005): "Markov-Switching Structural Vector Autoregressions: Theory and Application," *Federal Reserve Bank of Atlanta, Working Paper 2005-27*.

SARGENT, T. J. (1976): "The Observational Equivalence of Natural and Unnatural Rate Theories of Macroeconomics," *Journal of Political Economy*, 84(3), 631–640.

SIMS, C. A. (1986): "Are Forecasting Models Usable for Policy Analysis?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 10, 2–16.

——— (2005): "Dummy Observation Priors Revisited," Manuscript, Princeton University.

SIMS, C. A., AND T. ZHA (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39(4), 949–968.

——— (1999): "Error Bands for Impulse Responses," *Econometrica*, 67(5), 1113–1155.

——— (2006a): "Does Monetary Policy Generate Recessions?," *Macroeconomic Dynamics*, 10, 231–272.

——— (2006b): "Were There Regime Switches in US Monetary Policy?," *American Economic Review*, 96, 54–81.

SMETS, F., AND R. WOUTERS (2003): "An estimated stochastic dynamic general equilibrium model of the euro area," *Journal of European Economic Association*, 1(5), 1123–1175.

SPIVAK, M. (1965): *Calculus on Manifolds: A Modern Approach to Classical Thoerems of Advanced Calculus*. The Benjamin/Cummings Publishing Company, New York, New York.

STEWART, G. W. (1980): "The Efficient Generation of Random Orthogonal Matrices with an Application to Condition Estimators," *SIAM Journal on Numerical Analysis*, 17(3), 403–409.

UHLIG, H. (1997): "Bayesian Vector Autoregressions with Stochastic Volatility," *Econometrica*, 65, 59–73.

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

WAGGONER, D. F., AND T. ZHA (2003): "Likelihood Preserving Normalization in Multiple Equation Models," *Journal of Econometrics*, 114(2), 329–347.

ZHA, T. (1999): "Block Recursion and Structural Vector Autoregressions," *Journal of Econometrics*, 90(2), 291–316.