Shocks, Information and Structural VARs

Luca Gambetti *

Universitat Autònoma de Barcelona, Barcelona GSE,
Università di Torino and Collegio Carlo Alberto

Summary

Structural Vector Autoregressions (SVAR) have become one of the most popular tool to measure the effects of structural economic shocks. Several new techniques to “identify” economic shocks have been proposed in the literature in the last decades. Identification hinges on the implicit assumption that economic shocks are retrievable from the data. In other words, the data contain enough information to correctly estimate the shocks. SVAR models, however, are small-scale models, only a small number of variables can be handled, and this feature can forcefully limit the amount of information that variables can convey. After discussing the problems for identification arising from narrow information sets, the paper presents some theoretical results and empirical procedures aimed at testing whether information is enough to estimate

*I would like to thank Mario Forni and Luca Sala for helpful discussions and comments and Nicolò Maffei Faccioli for his very skillful research assistance.
economic shocks. Also, possible solutions to the problem of limited information like Factor Augmented VAR or dynamic rotations are discussed.

Keywords: Structural Vector Autoregressions, Shocks, Information, Non-invertibitliy, Non-fundamentalness.
1 Introduction

Understanding the propagation mechanisms of both policy and non-policy shocks still represents one of the major challenges for researchers and policymakers. Since the seminal paper of Sims (1980), Structural Vector Autoregressions (SVAR) have become one of the most popular tools in empirical macroeconomics employed to measure and study the effects of structural economic shocks.

In this class of models the variables are all endogenous and driven by economic shocks of different nature. The model is designed to measure and quantify, with a minimum number of restrictions derived from the economic theory, the effects on the economic system of these random disturbances. The model, in essence, is a multivariate autoregression where each variable depends on its own lags and the lags of the other variables. The residuals of these equations are linear combinations of the underlying structural economic shocks. Economic theory allows to identify these relations, so that the variables can be expressed as dynamic combinations of current and past economic shocks, where the coefficients of these combinations, i.e. the impulse response functions, represent the dynamic response of the variables to economic shocks.

This class of models has been extensively employed to study the effects of monetary
policy shocks, fiscal policy shocks as well as other non-policy shocks.\textsuperscript{1,2,3} Ramey (2016) and Kilian and Lütkepohl (2017) provide an excellent and exhaustive review of methods and results concerning the identification of macroeconomic shocks.


other words, current and past values of the variable in the model span the same space spanned by economic shocks and therefore shocks can be estimated as a linear combinations of the data. SVAR models however have been designed as small-scale models, i.e. a small number of variables can be handed. The feature forcefully limits the amount of information that variables can convey, and, in turn, makes the key assumption that shock are retrievable fragile and possibly not holding.

The problem of limited information sets was originally raised by Hansen and Sargent (1991) and Lippi and Reichlin (1993, 1994). The authors show that when the information set of the econometrician is narrower than that of the economic agents, then shocks cannot be estimated because the underlying model of the variable considered by the econometrician is non-fundamental, i.e. current and past observations of economic variables do not span the same space spanned by the shocks.

In recent years, several papers have shown that fundamentness does not hold in many economic models. For instance, in presence of anticipated fiscal or technology shocks, shocks which do not have immediate effects on fiscal variables, see Leeper, Walker and Yang (2008), or productivity, see Beaudry, Fève, Guay and Portier (2016). The information loss is due to the fact that fiscal variables or productivity become non-informative about current shocks being their effects delayed.

This entry of the Encyclopedia discusses the condition of validity of structural analysis in Vector Autoregressions, with focus on the conditions under which economic shocks are retrievable from a SVAR. I will present theoretical and empirical testable conditions to assess whether a specific VAR contains enough information to estimate
economic shocks. Also, I will discuss several ways of amending the model to include the relevant information if the model turns out to suffer of a lack of information.

The remainder of the paper is organized as follows. Section 2 reviews shocks identification in SVARs. Section 3 discusses the condition under which economic shocks are retrievable. Section 4 presents theoretical and empirical conditions to assess whether shocks are retrievable. Section 5 present two solutions to the problem of narrow information. Section 6 discusses an extension.

2 Shocks identification in SVAR

2.1 The Economy

I begin by discussing the class of economic models which is consistent with SVAR analysis. In the spirit of Frisch (1933) and Slutsky (1927), the macroeconomy is assumed to be the summation of agents’ reactions to random economic disturbances of various types occurring at every point in time. Formally, let $x_t$ be a $n$-dimensional stationary vector of time series with the following representation

$$x_t = F(L)u_t,$$  \hspace{1cm} (1)

where $u_t$ is a $q$-dimensional White Noise vector of orthonormal shocks and $F(L) = \sum_{k=0}^{\infty} F_kL^k$ is an $n \times q$ matrix of polynomials in the non-negative powers of the lag operator $L$. The vector $u_t$ includes the structural economic shocks and the matrix $F(L)$ contains the impulse response functions, the object that captures agents’ responses
to economic shocks. I impose no restrictions on $q$ and $n$, the number of shocks and variables in the economy (equation (1)).

Representation (1) is very general. A special case is when the representation is derived as the equilibrium solution of a Dynamic Stochastic General Equilibrium (DSGE) model. Consider the ABCD representation discussed in Fernández-Villaverde et al. (2007)

$$s_t = A s_{t-1} + B u_t$$
$$x_t = C s_{t-1} + D u_t,$$

where $s_t$ is an $m$-dimensional vector of state variables. Representation (1) can be derived as

$$x_t = [D + C(I - AL)^{-1}BL]u_t.$$  

(4)

Given that $x_t$ represents the whole economy, the number of variables $n$ can be large. The typical situation is that where the econometrician has to focus on a subset of variables. Let $z_t$ be a $s$-dimensional ($s \leq n$) subvector of $x_t$. Here we limit our attention to a subset of variables, but in principle $z_t$ could contain combinations of the elements in $x_t$ like principal components or averages. Moreover, we allow $z_t$ to be driven only by a subset of shocks $u_t^z$ of dimension $q_z$, with $q_z \leq q$, which are the object of interest to the econometrician. The structural economic representation for the subset of variables or linear combinations considered by the econometrician is given by

$$z_t = B(L)u_t^z$$  

(5)
where $B(L) = \sum_{k=0}^{\infty} B_k L^k$ is the matrix of structural impulse response functions. Again, I do not make any assumptions about $s$ and $q_z$, but in the subsections below I will discuss three different cases.

Here I do not take any stand on the true underlying economic theory. Any model delivering representation (1) and (5), i.e. a linear (not necessarily square) MA, is compatible with the analysis discussed below.

2.2 Vector Autoregressions

Once the class of economic models under considerations has been clarified, let us focus on the Vector Autoregression representation of the vector $z_t$. By stationarity, the Wold representation of $z_t$ exists, and is given by

$$z_t = C(L) \varepsilon_t$$  \hspace{1cm} (6)

where $\varepsilon_t \sim WN(0, \Sigma)$ is the Wold shock, $C(L) = \sum_{k=0}^{\infty} C_k L^k$ represents the Wold impulse response functions and $C_k$, with $k = 1, 2, ..., \infty$, are matrices of coefficients. If there are no roots on the unit circle, then an infinite VAR representation exists, and can be well approximated with a finite-order VAR

$$A(L) z_t = \varepsilon_t$$ \hspace{1cm} (7)

where $A(L) = I - A_1 L - ... - A_p L^p$ and $A_j$, $j = 1, ..., p$, are matrices of coefficients. Model (7) represents a Vector Autoregression of order $p$, VAR($p$).
2.3 Structural shocks

The main goal of Structural Vector Autoregression (SVAR) analysis is to recover the matrix of structural impulse response functions \( B(L) \) (global identification) or some of the columns of \( B(L) \) (partial identification) and the corresponding structural shocks starting from equations (6) and (7). This, in a nutshell, is implemented in three steps. First, the matrix \( A(L) \) and the innovation \( \varepsilon_t \) are estimated by least squares; second, \( C(L) \) is obtained by inverting \( A(L) \); third, the vector of structural shocks and structural impulse response functions, \( B(L) \), are obtained as a linear combination of the vector of innovations and as combinations of the Wold impulse response functions respectively. Formally

\[
B(L) = C(L)B_0 \tag{8}
\]

where \( B_0 \) is the identifying matrix and

\[
u_t^* = B_0^{-1}\varepsilon_t. \tag{9}\]

In the case of partial identification, where only one shock is identified, the relevant column of \( B(L) \), call it \( b(L) \), can be obtained as

\[
b(L) = C(L)b_0 \tag{10}\]

where \( b_0 \) is an identifying column vector and

\[
u_t^* = \tilde{b}_0\varepsilon_t. \tag{11}\]

where \( \tilde{b}_0 \) is a row of \( B_0^{-1} \). In practice, identification can be implemented by pinning down the orthogonal matrix \( H \) and setting \( B_0 = SH \), \( b_0 = Sh \), \( \tilde{b}_0 = h'S^{-1} \) where \( S \) is the
Cholesky factor of $\Sigma$. Several approaches have been proposed to obtain the structural shocks. A partial list includes zero contemporaneous or long-run restrictions, sign restrictions, maximizing restrictions. See Ramey (2016), Kilian and Lütkepohl (2017) and the entry of this encyclopedia Gambetti (2020) for a discussion and a review of the techniques. Here the focus is not to discuss how identification can be implemented. Rather we take the identifying matrix $H$ or column $h$ as given.

3 When are shocks retrievable?

In this section I discuss the conditions under which the strategy outlined in the previous section can be applied successfully and the structural representation correctly retrieved from the data.

3.1 Main concepts

I begin by discussing the concepts of invertibility, fundamentalness, recoverability, partial invertibility and sufficient information to better understand the conditions under which structural shocks and their effects can be inferred from current and past values of the data, and more specifically from the the Wold representation.

Invertibility is a property of moving averages representations. According to the standard textbook definition, the representation (5) is invertible if $u_t^e$ can be written as a combination of current and past values of $z_t$ with absolutely summable matrices of coefficients. In other words, observing $z_t$ and its past values is equivalent to observing
If (5) is invertible, then $u_t^z$ and $B(L)$ can be obtained by taking the appropriate linear combinations of the Wold shocks and impulse response functions.

A different but very similar concept is the concept of *fundamentalness*, see Rozanov (1967). The structural shocks are fundamental for $z_t$ if they belong to the linear space spanned by the present and past history of $z_t$. The main difference with invertibility is best seen when (5) is square. If the determinant of the matrix of the impulse response functions has one or more roots with unit modulus (the other roots being larger than 1 in modulus), the representation is not invertible, since $z_t$ does not have a VAR($\infty$) representation; nevertheless, fundamentalness holds, since the residuals of the projections of $z_t$ onto its first $k$ lags converge, as $k$ goes to infinity, to the space generated by the structural shocks. In the remainder of the paper we focus on the concept of invertibility.

Chahrour and Jurado (2019) discusses another concept which is related to the previous ones: *recoverability*. Recoverability, in essence, means that a shock can be obtained from the past, present and future values of the data. For instance the process $z_t = u_{t-1}$ is noninvertible but recoverable.

Invertibility is a global concept in the sense that refers to the recoverability of all of the shocks in $u_t^z$ from current and past values of the observable variables. However, even if invertibility does not hold, still some of the shocks can be obtained as linear combinations of current and past data. If a shock, or a subset of shocks, can be recovered, then the property of *partial invertibility* holds. Partial invertibility has

\[\text{11}\]

Invertibility and partial invertibility are concepts that can be expressed in terms of the informational content conveyed in a set of variables. Forni and Gambetti (2014) introduces the concept of *sufficient information*. Let $v_t$ be any sub-vector of $u^x_t$. The vector $z_t$ and the related VAR is *sufficient* for $v_t$ if and only if there exist a matrix $M$ such that $v_t = M \varepsilon_t$, i.e. there exists a combination of the innovations that delivers the structural shocks. Moreover $z_t$ is *globally sufficient* if it is informationally sufficient for all the elements of $u^x_t$. In this case there exists a matrix $M$ such that $u^x_t = M \varepsilon_t$. Notice that this last property (global sufficiency) coincides with the concept of invertibility: the structural shocks can be found as a linear combination of the current and past values of economic variables. Similarly, sufficient information for a single shock coincides with the concept of partial invertibility.

It is worthy to stress the importance of sufficiency and partial invertibility of the structural moving average representation since it is common practice in the literature to identify just one single shock or a subset of shocks.

I next discuss the above concepts for three different cases, corresponding to different specifications of the vector $u^x_t$.

### 3.2 The standard case: $q_z = s$

Let us assume that the number of shocks in $u^x_t$ is the same as the number of variables in $z_t$, i.e. $B(L)$ is a square matrix. This is the case commonly considered in the
literature. Invertibility requires that all of the roots of the determinant of $B(L)$ have to be strictly larger than one in modulus.\(^5\) When this condition holds, then all of the structural shocks and the structural impulse response functions can be obtained from the Wold representation as \(u_t^* = B_0^{-1}\epsilon_t\) and \(B(L) = C(L)B_0\) respectively (see Section 3 for the details).\(^6\)

As already mentioned above, there might be cases where invertibility does not hold. In this situations, the vector \(u_t^*\) cannot be obtained from the vector of innovations. Recently, several paper have shown that noninvertibility is likely to arise in presence of anticipated shocks, shocks which have delayed effects on the key variables, see for instance Leeper, Walker and Yang (2008). However, even in these situations it might be still possible to obtain a subset of shocks, i.e. \(z_t\) could be partially sufficient for a subvector of \(u_t^*\), see Sims and Zha (2006) and Sims (2012). To illustrate the point, let us consider the following simple example. Consider the model

\[
\begin{pmatrix}
z_{1t} \\
z_{2t}
\end{pmatrix} =
\begin{pmatrix}
0.5 & L \\
0.5 & -L
\end{pmatrix}
\begin{pmatrix}
u_{1t} \\
u_{2t}
\end{pmatrix}
\] (12)

The determinant of the MA matrix, $-L$, vanishes in zero, so that the MA representation is noninvertible and \(z_t\) is not globally sufficient for \(u_t\). Nevertheless, \(z_t\) is sufficient for \(u_{1t}\), because \(z_{1t} + z_{2t} = u_{1t}\), and therefore the model is partially invertible.

We use a simulation to illustrate the point. Consider the monetary policy rule

\[i_t = \phi\pi_t + u_{2t}\] where \(\pi_t\) is inflation and \(u_{2t}\) represents the monetary policy shock. We

\(^5\)Fundamentalness allows the roots also to be on the unit circle.

\(^6\)Under invertibility \(B(L)^{-1}z_t = u_t^*\). By pre multiplying by \(B_0\), \(\tilde{B}(L)z_t = B_0u_t^*\), where \(\tilde{B}(0) = I\), so that \(\varepsilon_t = B_0u_t\).
assume the following model for inflation and the interest rate satisfying the above policy rule

$$
\begin{pmatrix}
\pi_t \\
i_t
\end{pmatrix} =
\begin{pmatrix}
0.5L + 0.2L^2 & -0.4L - 0.7L^2 \\
1.5(0.5L + 0.2L^2) & 1.5(-0.4L - 0.7L^2) + 1
\end{pmatrix}
\begin{pmatrix}
u_{1t} \\
u_{2t}
\end{pmatrix}
$$

(13)

Notice that zero is one of the root of the determinant of the above moving average, so the model is noninvertible. However the monetary shocks and its impulse response functions can be estimated since $u_{2t} = i_t - \phi \pi_t$. We simulate 1000 time series for $\pi_t$ and $i_t$ and for each realization a VAR model is estimated with OLS and a Cholesky identification scheme implemented. Notice that the second shock in the Cholesky representation satisfies the restrictions of the monetary policy shock: nonzero contemporaneous effects on $i_t$ and zero contemporaneous effect on $\pi_t$. Panel (a) of Figure 1 plots the responses to $u_{1t}$ and Panel (b) the response to $u_{2t}$ of plots the median, the 2.5th and 97.5th percentile of the responses, together with the true theoretical response (solid blue lines). The effects of $u_{1t}$ are very badly estimated, but the impulse response functions of $u_{2t}$ are correctly captured. This is the case because the model is non-invertible (no all of the shock can be retrieved), but the variables are informationally sufficient for the monetary policy shock $u_{2t}$.

### 3.3 The nonstandard case: $q_z \neq s$

In the nonstandard case, the number of variables and shocks do not coincide. Let us start with $q_z < s$, the number of shocks is smaller than the number of variables so that the MA representation (5) is tall. A tall MA, except in very special cases, is
always invertible. Indeed, noninvertibility requires that the determinant of all of the $q_z \times q_z$ submatrices of $B(L)$ share exactly the common root on or inside the unit circle. That case is avoided with a minimum of heterogeneity in the response functions. The resulting VAR representation exists, it is of finite order and has reduced dynamic rank, see Anderson and Deistler (2008). The VAR can still be consistently estimated using standard techniques, but the resulting covariance matrix of the VAR innovations will be of reduced rank. This has to be taken into account when the structural shocks are obtained. The importance of tall systems and singular VARs for estimating structural shocks is discussed in Forni et al. (2009).

We now move to the case $q_z > s$. In this case the model is never invertible and
globally sufficient, there is no matrix delivering $u_t^*$ as a linear combination of the vector of innovations. In other words, the variables do not contain enough information to disentangle structural shocks because there are more shocks than variables. However, the same logic discussed above applies here. Some of the shocks can be recovered.

For instance, consider again the above Taylor rule $i_t = \phi \pi_t + u_{3t}$ where $\pi_t$ is inflation and $u_{3t}$ represents the monetary policy shock. Clearly, in this case the monetary policy shock is always recoverable independently on the number of other shocks driving the two variables. Suppose there are three shocks driving inflation and the interest rate. The representation of the two variables is

$$\begin{pmatrix} \pi_t \\ i_t \end{pmatrix} = \begin{pmatrix} b_1(L) & b_2(L) & b_3(L) \\ \phi b_1(L) + 1 & \phi b_2(L) & \phi b_3(L) \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}. \quad (14)$$

The model is noninvertible and not sufficient globally, so the vector $u_t$ cannot be recovered from the data. But $u_{3t}$ can be obtained from the combination $u_{3t} = i_t - \phi \pi_t$.

I extend the simulation above to better illustrate the point. I assume that $b_1(L) = -0.5 - 0.2L$, $b_2(L) = 0.8$, $b_3(L) = -0.4L - 0.7L^2$, and $\phi = 1.5$. Under this parametrization the monetary policy shock, $u_{3t}$, can be correctly identified as the second shock of the Cholesky decomposition of a VAR with the two variables since $b_3(0) = 0$. I generate 1000 dataset with $u_t \sim N(0, I)$. For each dataset a VAR with AIC lags is estimated. Panel (c) of Figure 1 plot the median, the 2.5th and 97.5th percentile of the responses, together with the true theoretical response. The median and the theoretical response perfectly overlap, showing that a VAR can correctly es-
timate the responses even when the variables are less than the shocks. Obviously the first shock of the Cholesky representation will not have any structural interpretation since it will be combination of the current and past values of the remaining structural shocks.

3.4 Summary

The implications discussed in the previous subsection are summarized in Table 1. Partial invertibility and partial sufficiency can always hold, independently on the number of shocks and independently on whether the root condition is satisfied or not in the square case. On the contrary, invertibility and global sufficiency hold only when $q_z < s$ or when the root condition is satisfied in the case of $q_z = s$. The discussion points out the importance of the concepts of partial invertibility and sufficiency since in empirical analyses most of the time just one shock is identified.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Global Invertibility/Sufficiency</th>
<th>Partial Invertibility/Sufficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_z &lt; s$</td>
<td>Always</td>
<td>Possible</td>
</tr>
<tr>
<td>$q_z = s$</td>
<td>Root condition</td>
<td>Possible</td>
</tr>
<tr>
<td>$q_z &gt; s$</td>
<td>Never</td>
<td>Possible</td>
</tr>
</tbody>
</table>

Table 1: Summary of the findings for different model specifications.
3.5 Deficient information

Noninvertibility and partial noninvertibility of representation (5) can arise for two reasons. First, the economy itself is noninvertible, i.e. a left inverse of $F(L)$ in the nonnegative powers of $L$ in equation (1) does not exist. In this situation it is obvious that there are no subvectors of $x_t$ that can be used to estimate the structural shocks. Noninvertible economies can arise because of imperfect information as we will discuss in section 5.3. However, this case is largely ignored in the literature and the economy (1) is typically assumed to be invertible and $x_t$ sufficient for $u_t$ (and therefore for $u^z_t$). If the economy is invertible, then noninvertibility of (5) can only arise, this is the second reason, because $z_t$ does not contain enough information to estimate $u^z_t$. In other words, it arises because the information set of the econometrician ($z_{t-j}$, $j = 0, 1, ...$) is narrower than that of the agent ($x_{t-j}$, $j = 0, 1, ...$) and this makes the shocks non-retrievable. This is the situation discussed in the seminal paper of Hansen and Sargent (1991).\(^7\) Thus, invertibility should be seen as a direct consequence of a deficient information set.

In recent years, several papers have shown that invertibility does not hold in many economic models. For instance, in presence of anticipated fiscal or technology shocks, shocks which do not have immediate effects on fiscal variables (see the next subsection) or productivity respectively. The information loss is due to the fact that these variables become non-informative about current shocks being their effects delayed.\(^8\) I will present an example in the next section.

\(^7\)See also Lippi and Reichlin (1993, 1994) for early reference about the problem of noninvertibility.

\(^8\)News shocks represent a typical example where noninvertibility can arise, see Beaudry, Fève,
4 Theoretical conditions and diagnostics

4.1 An economic example: fiscal foresight

Leeper, Walker and Yang (2008) shows that noninvertibility in VAR models naturally arises in an economy with fiscal foresight.\(^9\) Starting with a standard growth model with log preferences and inelastic labor supply, the authors obtain the equilibrium capital accumulation equation

\[
k_t = \alpha k_{t-1} + a_t - \kappa \sum_{i=0}^{\infty} \theta^i E_t \tau_{t+i+1}
\]

where \(\kappa = (1 - \theta) \left( \frac{\tau}{1 - \tau} \right)\), \(\tau\) being the steady state tax rate and \(\theta < 1\), and \(k_t, a_t\) and \(\tau_t\) are the log deviations from the steady state of capital, technology and the tax rate, respectively. Under the assumption that that the effect of fiscal policy on taxes is delayed by two periods, technology and taxes follow

\[
a_t = u_{A,t} \\
\tau_t = u_{\tau,t-2}
\]

where \(u_{\tau,t}\) and \(u_{A,t}\) are i.i.d. shocks that economic agents can observe. The solution with two periods of foresight for the model variables is

\[
\begin{pmatrix}
a_t \\
k_t \\
\tau_t
\end{pmatrix} =
\begin{pmatrix}
0 & 1 \\
\frac{-\kappa L + \theta}{1 - \alpha L} & \frac{1}{1 - \alpha L} \\
L^2 & 0
\end{pmatrix}
\begin{pmatrix}
u_{\tau,t} \\
u_{A,t}
\end{pmatrix}.
\]

Guay and Portier (2016).

\(^9\)Simple examples of noninvertibility in economic models can also be found in Lippi and Reichlin (1993) and Fernández-Villaverde, Rubio-Ramírez, Sargent and Watson (2007).
Equation (16) represents a special case of model (1). Now, let us consider the square subsystem given by the first two rows (technology and capital): the determinant \( \frac{\alpha(z+\theta)}{1-\alpha_L} \) vanishes for \( z = -\theta \), which is less than 1 in modulus. Similarly, the determinant of the submatrix given by the first and the last rows (technology and taxes) is \(-z^2\), which vanishes for \( z = 0 \). Finally, the determinant of the subsystem formed by the second and the last row (capital and taxes) also vanishes for \( z = 0 \). In conclusion, \((u_{\tau,t}, u_{A,t})'\) is non-invertible for any pair of variables on the left-hand side, implying that standard VAR techniques are unable to correctly estimate the fiscal shock.

However it should be noted that the two shocks can be obtained using the three variables all together, since the three different \(2 \times 2\) submatrices have different roots. This is a case of \( s > q^z \) and therefore the model is fundamental.

### 4.2 Theoretical conditions for invertibility and sufficient information

This subsection focuses on the relation between DSGE models and VARs, see Giacomini (2013), Franchi and Vidotto (2013), Franchi and Paruolo (2014) and Pagan and Robinson (2016). More specifically, we review the conditions under which a DSGE model admits a VAR representation, and the conditions under which SVAR models can be employed to estimate the effects of DSGE shocks, see Ravenna (2007).
Consider the following (log) linear solution of a DSGE model

\[
    s_t = A s_{t-1} + B u_t \tag{17}
\]
\[
    z_t = C s_{t-1} + D u_t \tag{18}
\]

where \( s_t \) is an \( r \)-dimensional vector of stationary state variables, \( z_t \) again is the vector considered by the econometrician, \( q \) is the number of shocks, \( q \leq r \leq n \), \( A, B, C \) and \( D \) are conformable matrices of parameters and \( B \) has a left inverse \( B^{-1} \) such that \( B^{-1}B = I_q \). Also

\[
    u_t = B^{-1} s_t - B^{-1} A s_{t-1}.
\]

Substituting in \( z_t \) we have

\[
    z_t = \left[ DB^{-1} - (DB^{-1}A - C)L \right] s_t
\]

In the square case \( q = s \) we have

\[
    z_t = DB^{-1} \left[ I - (A - BD^{-1}C)L \right] s_t.
\]

The shocks can be obtained as a square summable combination of the present and past of \( z_t \) if and only if the eigenvalues of \( (A - BD^{-1}C) \) are strictly less than one in modulus, the so called “poor man’s condition” discussed in Fernández-Villaverde et al (2007). When this condition holds a VAR representation in the structural shocks exists

\[
    z_t = \sum_{j=0}^{\infty} (A - BD^{-1}C)^j BD^{-1} z_{t-j} + Du_t.
\]
4.3 Theoretical conditions for partial sufficiency

In many cases, however, the condition is too restrictive since the researcher might be interested in investigating the effects of a single shock. The theoretical conditions under which sufficiency holds have been for the first time discussed in Sims and Zha (2006). Let us consider the projection of $u^*_t$ onto the entries of $\varepsilon_t$

$$u^*_t = M\varepsilon_t + e_t$$  \hspace{1cm} (19)

where $\varepsilon_t$ is the innovation of the VAR for $z_t$ and the fraction of unexplained variance in the above regression (recall $\sigma^2_{u^*_t} = 1$)

$$\delta_i = \sigma^2_{\varepsilon_i}.$$  \hspace{1cm} (20)

When $\delta_i = 0$ then partial informational sufficiency holds for shock $i$. The idea is that in this case the structural shock is an exact linear combination of the innovations. On the contrary, a large value of $\delta_i$ means that the structural shock cannot be obtained from the innovations and therefore from a VAR. Notice that, in many cases $\delta_i$ might be nonzero but small. In these cases information sufficiency is only approximate but still, as shown in Forni et al. (2019), the shocks can be estimated with very good approximation.

4.4 Empirical diagnostics

From an empirical point of view the key question is whether $z_t$ includes enough information to estimate the structural shocks and their impulse response functions. Several papers have suggested procedures in order to address the question.
Chen, Choi and Escanciano (2017) shows that, with non-Gaussian i.i.d. disturbances, the Wold innovations are a martingale difference sequence if and only if the structural shocks are fundamental. Using this theoretical result, the authors propose a testing procedure to assess whether the Wold innovations are martingale difference sequence. Soccorisi (2016) proposes a method to measure nonfundamentalness.

Giannone and Reichlin (2006), based on a result in Forni and Reichlin (1996), uses a Granger causality test to assess whether a set of variables has an invertible representation. Forni and Gambetti (2014) proposes a Granger causality test for global sufficiency and an orthogonality test for sufficiency of a single shock. The main difference between the two procedures is that Forni and Gambetti (2014) uses the principal components since the test is constructed on the basis of a theoretical necessary and sufficient condition of information sufficiency.

Assume that the economy has the state-space representation (17)-(18). Under this assumption the testing procedure is the following.

1. Consider a large data set $x_t$, capturing all of the relevant macroeconomic information.
2. Establish and compute a maximum number of principal components $P$ of $x_t$.\textsuperscript{10}
3. Perform a Granger causality test to see whether such principal components Granger cause $z_t$. If the null of no Granger causality is not rejected, $z_t$ is informationally sufficient. Otherwise, sufficiency is rejected.

\textsuperscript{10}The principal components are nothing but a consistent estimator of a combination of current and lagged values of the state variables of the model, see Forni and Gambetti (2014).
Notice that the above discussed tests are designed to test global sufficiency. Sufficiency of a single shocks can be tested similarly. The test simply reduces to an $F$-test of orthogonality of the estimated shock with respect to the lags of the estimated factors.

Canova and Sahneh (2018) suggests to use the Geweke, Mese and Dent (1983) version of the Granger causality test applied on the VAR residuals. This type of test has been used in Ramey (2011) to show that the government spending shock obtained with a SVAR $a la$ Perotti (2007) is predicted by the forecast of public expenditure from the survey of professional forecasters and therefore cannot represent the government spending shock.

5 Solutions

Here I discuss a few solutions to the problem of invertibility.

5.1 Extending the information set

If sufficiency is rejected, a solution to the problem is to enlarge the information set. The simplest way would be to add variables in $z_t$. The problem with this is that in general is not clear what variables should be included.\textsuperscript{11} An alternative is to augment the vector $z_t$ in the VAR with the relevant factors and estimating a Factor-Augmented VAR model (FAVAR), see Bernanke Boivin and Eliasz (2005). The rationale is that the factors are the objects that convey all of the relevant information about the dynamics of the economy. When the data are generated by a DSGE model, the

\textsuperscript{11}Jarocinski and Mackoviak (2018) propose a procedure to select the correct variables.
factors span the space spanned by the state variables of the model. The factors can be consistently estimated with the principal components of a large dataset, see Stock and Watson (2002). Thus, the factors add to the model the relevant information coming from many economic series, and this can solve the deficient information problem.

FAVAR models have been extensively used over the last years precisely to cope with the problem of narrow information sets and it has been shown that they can solve many existing puzzles.\textsuperscript{12} The factors can be consistently estimated with the principal component estimator and the number of factor to be included can be established by repeating the same test (Granger causality or orthogonality depending on the goal of the application) with the vector of variables augmented by the factors.

An alternative to FAVAR models to solve the problem of deficient information is represented by large Bayesian VAR, see Banbura, Giannone and Reichlin (2010). The Bayesian approach allows to handle large dataset and include hundreds of variables in the model. The curse of dimensionality problem is solved by setting the degree of prior tightness in relation to the model dimension. A potential drawback is that in large BVAR the number of restrictions to identify economic shocks increases substantially. Large BVAR have been extensively used for both forecasting and structural analysis, see, among others, Ellahie and Ricco (2017). Dynamic Factor Models represent another alternative and they have been extensively used for structural analysis

\textsuperscript{12}See, for example, Bianchi, Muntaz and Surico (2009), Boivin, Giannoni and Mihov (2009), Ludvigson and NG (2009), Moench (2008) and Muntaz and Surico (2008), among others. Bernanke and Boivin (2003) investigates the mapping between large-N models and DSGE models.
5.2 Dynamic rotations

A different approach to confront with the problem of noninvertibility and nonfundamentalness consists of directly estimating the nonfundamental representation by applying dynamic rotations by means of Blaschke matrices. The approach was originally proposed by Lippi and Reichlin (1993, 1994) and has been recently used in several applications.

First of all let us consider the definition of Blaschke matrix. $M(z)$ is a Blaschke matrix if (i) has no poles less or equal to one in modulus and (ii) $M(z)^{-1} = M^*(z^{-1})$, where $M^*$ is the conjugate transpose. A property of a Blaschke matrix is that if $u_t$ is an orthonormal white noise then also $v_t = M(z)u_t$ is a orthonormal white noise. Let

$$R(\alpha_i, z) = \begin{pmatrix} \frac{z-\alpha_i}{1-\alpha_iz} & 0 \\ 0 & I \end{pmatrix}$$

where $|\alpha_i| < 1$ and $I$ is the $(n-1)$ identity matrix. Then the $n \times n$ Blaschke matrix is given by

$$M(z) = R(\alpha_1, z)K_1R(\alpha_2, z)K_2...R(\alpha_r, z)K_r$$

where $K_i$ are orthogonal matrices and $|\alpha_i| < 1$. The Blaschke matrix can be used to derive the nonfundamental representation starting from a fundamental representation.

---

13Panel VAR, see Canova and Ciccarelli (2013) for a survey, or Global VAR, see Pesaran, Schuermann and Weiner (2004), represent alternative approaches which are especially employed with multicountry or multisector models.
Let

\[ x_t = N(L)v_t \]

be the fundamental representation of \( x_t \). Let \( w_t = M(L)^{-1}v_t \), where \( M(L) \) is a Blaschke matrix. The nonfundamental representation is given by

\[ x_t = P(L)w_t \]

where \( P(L) = N(L)M(L) \). Notice that \( v_t \) is not contained in the space spanned by the present and past values of \( x_t \), but instead lies in the future of \( u_t \).

Now, I will use a simple example to illustrate how Blaschke matrices can be used in a VAR to estimate the impulse response functions to structural shocks when the structural model is nonfundamental. Consider the following simple structural model expressed in terms of orthonormal shocks:

\[
\begin{pmatrix}
  x_{1t} \\
  x_{2t}
\end{pmatrix}
= \begin{pmatrix}
  \sigma_1 L & 0 \\
  \sigma_1 & \sigma_2
\end{pmatrix}
\begin{pmatrix}
  w_{1t} \\
  w_{2t}
\end{pmatrix}
\]

where \( w_t = [w_{1t} \ w_{2t}]' \sim WN(0, I) \). The process is non-fundamental since the determinant is zero in zero. The inverse Blaschke matrix is

\[
M(L)^{-1} = K^{-1}R(L)^{-1}
\]

\[
= \begin{pmatrix}
  \frac{\sigma_2}{\sigma} & \frac{\sigma_1}{\sigma} \\
  -\frac{\sigma_1}{\sigma} & \frac{\sigma_2}{\sigma}
\end{pmatrix}
\begin{pmatrix}
  L^{-1} & 0 \\
  0 & 1
\end{pmatrix}
\]

\[
= \begin{pmatrix}
  \frac{\sigma_2}{\sigma}L^{-1} & \frac{\sigma_1}{\sigma} \\
  -\frac{\sigma_1}{\sigma}L^{-1} & \frac{\sigma_2}{\sigma}
\end{pmatrix}
\]  (21)
Notice that $R(L)$ is obtained by setting $\alpha_i = 0$, the root smaller than one of the determinant. The fundamental representation can be therefore obtained as

$$
\begin{pmatrix}
x_{1t} \\
x_{2t}
\end{pmatrix} = \begin{pmatrix}
\sigma_1 L & 0 \\
\sigma_2 & \sigma_2
\end{pmatrix} \begin{pmatrix}
\frac{\sigma_2}{\sigma} L^{-1} & \frac{\sigma_1}{\sigma} \\
-\frac{\sigma_2}{\sigma} L^{-1} & \frac{\sigma_2}{\sigma}
\end{pmatrix} \begin{pmatrix}
v_{1t} \\
v_{2t}
\end{pmatrix}
$$

$$
\begin{pmatrix}
x_{1t} \\
x_{2t}
\end{pmatrix} = \begin{pmatrix}
\frac{\sigma_1 \sigma_2}{\sigma} & \sigma_1 \sigma_2 L \\
0 & \sigma
\end{pmatrix} \begin{pmatrix}
v_{1t} \\
v_{2t}
\end{pmatrix}
$$

(22)

Notice also that (22) corresponds to the Cholesky representation of $z_t$. To estimate the impulse response functions of the nonfundamental shocks one can therefore proceed in two steps. First the Cholesky representation (22) is estimated by standard methods. Second, the Cholesky representation is postmultiplied by the Blaschke matrix $M(L)^{14}$. Notice that the ratios $\frac{\sigma_1}{\sigma}, i = 1, 2$, can be obtained from the coefficients of the Cholesky representation. This is a very simple example. In general, the main drawback of the approach stems from the difficulty in identifying the parameters of the Blaschke matrix.

A few recent papers have used this approach to study the effects of nonfundamental shocks. For instance, Mertens and Ravn (2010) applies this approach to estimate the effects of anticipated fiscal policy shocks.

\[14\text{The Blaschke matrix is given by}
M(L) = \begin{pmatrix}
\frac{\sigma_2}{\sigma} L & -\frac{\sigma_1^2}{\sigma} L \\
\frac{\sigma_1}{\sigma} & \frac{\sigma_2}{\sigma}
\end{pmatrix}\]
6 A deeper nonfundamentalness: Agents’ limited information sets

So far we have focused on the problem of nonfundamentalness, noninvertibility, and insufficiency arising from deficient econometrician’s information set. In this section we discuss a situation where even when the econometrician’s information set coincides with that of the agents, still the shocks are not retrievable because the agents themselves do not have enough information. This situation generates a deeper problem in the sense that the data themselves, which are ultimately the outcomes of agents decisions, do not contain enough information. In this case the standard solution to the nonfundamentalness problem, that is extending the dataset, would not be effective. I will give the intuition below of what could be a possible solution.

In recent years several paper have challenged the plausibility of the assumption that agents perfectly observe the economic shocks hitting the economy\textsuperscript{15}. Rather it is more realistic to think that agents only observe a noisy signal of the shock

\[ s_t = u_t + e_t \]

where \( u_t \) is now a scalar shock and \( e_t \) is the noise. The implications of imperfect information for SVAR analysis are deep and problematic. If agents cannot observe the shocks, then the econometrician, using current and past values of the data, which are ultimately the outcome of agents’ decisions, will not be able to estimate the

structural shocks. If she could, so would do the agent, contradicting the initial assumption. Technically speaking, this is a case of nonfundamentalness much deeper than the standard case arising when the agents have richer information sets than the econometrician. In this case, not even the agents are able to recover structural disturbances. At a first glance it seems that under imperfect information SVAR models are fated to fail, see Blanchard et al. (2013) and Barsky and Sims (2012).

Let’s see the problem with a simple example. Assume that the economic fundamental $a_t$ evolves according to

$$a_t = a_{t-1} + u_{t-1}$$

and suppose the agent observe the above signal $s_t$ but not $u_t$ itself. The MA representation of the two variables is given by

$$
\begin{pmatrix}
\Delta a_t \\
 s_t
\end{pmatrix} =
\begin{pmatrix}
 L & 0 \\
 1 & 1
\end{pmatrix}
\begin{pmatrix}
 u_t \\
 \epsilon_t
\end{pmatrix},
$$

which is noninvertible being the determinant zero in zero. Using the Blaschke matrix discussed above, one can obtain the fundamental representation

$$
\begin{pmatrix}
\Delta a_t \\
 s_t
\end{pmatrix} =
\begin{pmatrix}
 1 & L\sigma_u^2/\sigma_s^2 \\
 0 & 1
\end{pmatrix}
\begin{pmatrix}
 \epsilon_t \\
 s_t
\end{pmatrix},
$$

(23)

where

$$
\begin{pmatrix}
\epsilon_t \\
 s_t
\end{pmatrix} =
\begin{pmatrix}
 L\sigma_u^2/\sigma_s^2 & -L\sigma_u^2/\sigma_s^2 \\
 1 & 1
\end{pmatrix}
\begin{pmatrix}
 u_t \\
 \epsilon_t
\end{pmatrix}.
$$

(24)

Notice that the shocks $\epsilon_t$ and $s_t$ are jointly white noise, and (23) corresponds to the Cholesky representation of $\Delta a_t$ and $s_t$. This implies that the econometrician
estimating a VAR for $\Delta a_t$ and $s_t$ would recover two shocks which are combinations of present and past values of the structural shocks and standard identification techniques cannot deliver the structural shocks.

However, the structural impulse response functions can be estimated as discussed earlier by applying the Blaschke matrix and the noninvertible shocks can be recovered by taking combination of the future of the invertible shocks as seen above. Recent advances in SVAR models under imperfect information are discussed in Forni et al (2017a, 2017b) and Chahrour and Jurado (2018, 2019).
7 Further readings


References


Princeton.


Statistics, 85, 777-792.


