DISCUSSION PAPERS
Institute of Economics
University of Copenhagen

04-14

A Priori Inequality Restrictions and Bound Analysis in VAR Models

Massimo Franchi

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

A priori inequality restrictions and bound analysis in VAR models

Massimo Franchi*

July 21, 2004

Abstract

The aim of this paper is to use inequality restrictions on the parameters of a structural model to find bounds on impulse response functions which are valid for any structural representation satisfying those restrictions. Economic theories specify signs and bounds of the coefficients which are the same among alternative paradigms: parameters are either positive or negative and propensities are between zero and one. These restrictions can thus provide a core of well established a priori impositions on which one can derive an economically meaningful interpretation of the reduced form system. Unlike just and over-identifying restrictions, inequalities select a set of structural interpretations: for this reason inference on impulse responses is derived as a bound analysis. In the last section we introduce an objective method to compare alternative under-identifying restrictions expressed as inequalities.

JEL classification: C32.

Keywords: VAR, identification, inequality restrictions, impulse response functions.

1 Introduction

The problem of drawing inferences from the probability distribution of the observed variables to the unknown structure that generates them is the very central issue in the econometrics of systems of equations. The study of identification[1] dates back to the beginning of the thirties and to the works of the Cowles Commission. Even though this literature is mainly concerned with linear models and with the problem of estimating structural relations, the framework also applies to more general setups[2] and to the literature on VAR models[3].

Koopmans, Rubin, and Leipnik (1950) and Fisher (1959) suggested and further developed the idea of using inequality restrictions for identification; in this paper we study how to implement this idea in the VAR framework. This line of research is closely related to a very recent literature on sign restrictions: Faust (1998) considers the claim that the monetary policy shock accounts for a small share of the forecast error variance of output and checks all possible identifications for the one that is the worst for the claim, subject to the restriction that the implied economic structure produce reasonable responses to policy shocks; Uhlig (1999) defines a monetary innovation in terms of its qualitative effects on interest rate and

*I would like to thank Søren Johansen, Katarina Juselius, Marco Lippi and Harald Uhlig for very precious comments. All remaining errors are mine. *Address for correspondence:* Dipartimento di Scienze Economiche, Via Cesalpino 12, 00161 Rome, Italy. *E-mail:* massimo.franchi@uniroma1.it.

[1]See Frisch (1934), Haavelmo (1943, 1944), Koopmans (1949), Koopmans and Reiesol (1950), Koopmans, Rubin, and Leipnik (1950), Hurwicz (1950), Wald (1950), Fisher (1959, 1963, 1966) and Rothenberg (1971).

[2]See Chesher (2003) for a discussion about identification in non separable models.

[3]Explicitly or not the rank and order conditions for identification are used in integrated and stationary VAR models, see Johansen (1995) and Favero (2001).

prices, selects the identification schemes that satisfy these priors and checks what is the effect of such innovations on real activity; Canova and De Nicoló (2000) use theoretical information about the pairwise dynamic cross correlations in response to structural shocks in order to discriminate over the set of all possible identifications. A weak feature of these contributions is that in order to characterize the set of identification schemes that satisfy the inequality restrictions they rely on numerical procedures; on the contrary we show that an analytical approach to the problem is feasible and that it leads to a wider interpretability of the results.

In the VAR approach the identification problem comes *after* the estimation of the always identified unrestricted reduced form, and it involves the choice of the economic interpretation of that reduced form. In this case the crucial issue is whether the a priori restrictions are well founded or at least reasonable from a theoretical point of view: in some cases the linear restrictions needed to satisfy the rank condition are available and generally acceptable; in many others the agreement on these restrictions is very weak and alternative theoretical paradigms specify different identifying restrictions.

When the estimation is conducted on the reduced form there is no economic reason why we should restrict ourselves to the imposition of identifying restrictions. Inequality restrictions are much more general than linear restrictions and it is often the case that competing paradigms agree on signs and bounds of some coefficients while they don't agree on exclusions: for example, a propensity is between zero and one, an elasticity is positive or negative, a restrictive monetary innovation raises the short term interest rate. In this sense a system of inequality restrictions can provide a core of well established a priori impositions on which deriving a structural interpretation of the reduced form system.

Before imposing a priori restrictions on the reduced form system we may use it to derive bounds that *any* structural interpretation satisfies by construction: these bounds are calculated directly from the reduced form system and hold independently from the a priori restrictions we impose for identification. They specify the ranges over which the structural parameters or the impulse responses vary when the identifying restrictions change. This is the subject of section 2. In section 3 we motivate the use of inequality restrictions and expose how to implement it; in general a system of inequalities is not identifying and selects a set of structural interpretations: for this reason inference on impulse responses is derived as a bound analysis. In section 4 we propose an objective criterion to compare alternative under-identified systems and in the last section we collect some concluding remarks.

2 Structures, admissible transformations and reduced form bounds for VAR models

Consider the system[4]

$$\boldsymbol{A}_0\boldsymbol{x}_t + \boldsymbol{A}_1\boldsymbol{x}_{t-1} = \boldsymbol{\varepsilon}_t \tag{2.1}$$

where $\boldsymbol{x}_t$ is the $p \times 1$ vector of endogenous variables, $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$ are $p \times p$ matrices of coefficients and $\boldsymbol{\varepsilon}_t$ is the $p \times 1$ vector of unobservables.

Following Koopmans, Rubin, and Leipnik (1950) the concepts of structure and observational equivalence are defined as

Definition 2.1. *A structure a is defined by the coefficient matrices $\boldsymbol{A}_0$ (non singular) and $\boldsymbol{A}_1$ and by the conditional distribution function f_a for the unobservables, that is $a = (\boldsymbol{A}_0, \boldsymbol{A}_1, f_a)$. We say that two structures a and b are observationally equivalent if they imply*

[4]Any system with a finite number of lags may be rewritten as a system with just one lag. The distinction between endogenous and exogenous variables and the presence of deterministic components do not affect the exposition.

the same conditional probability distribution of the endogenous variables, that is

$$a \sim b \overset{def}{\Longleftrightarrow} P(\boldsymbol{x}_t|a, \boldsymbol{x}_{t-1}) = P(\boldsymbol{x}_t|b, \boldsymbol{x}_{t-1})$$

We now investigate the properties of observationally equivalent structures in the VAR framework. The reduced form VAR is

$$\boldsymbol{x}_t = \boldsymbol{\Pi}\boldsymbol{x}_{t-1} + \boldsymbol{u}_t \tag{2.2}$$

where $\boldsymbol{x}_t$ is the $p \times 1$ vector of endogenous variables, $\boldsymbol{\Pi}$ is a $p \times p$ matrix of reduced form coefficients and $\boldsymbol{u}_t$ is the $p \times 1$ vector of unobservables. Notice that no assumption about the stationarity of the process has been made. Since identification is a population problem and it is independent from the finite sample problems related to estimation, in the exposition we will assume the population moments as known; in this way we focus exclusively on the problems related to the identification of the system leaving aside the ones related to its estimation.

We make the two following standard assumptions:

Distributional Assumption: $\boldsymbol{u}_t$ is i.i.d. normal with mean zero and covariance $\boldsymbol{\Omega}$, $\boldsymbol{u}_t \sim i.i.d.\ N(\boldsymbol{0}, \boldsymbol{\Omega})$;

Covariance Restriction: the covariance of the unobservables of each structure is the identity, that is $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \boldsymbol{I}$ for every structure.

The first assumption implies that the conditional joint density of the endogenous variables is completely characterized by the first two conditional moments, mean and variance. The second is the standard restriction that consists in transforming the reduced form into a representation that displays orthogonal innovations.

These two assumptions imply that the generic structural form of (2.2) is written as

$$\boldsymbol{A}_0\boldsymbol{x}_t = \boldsymbol{A}_1\boldsymbol{x}_{t-1} + \boldsymbol{\varepsilon}_t \tag{2.3}$$

where $\boldsymbol{\varepsilon}_t$ is a normal i.i.d. vector with mean zero and covariance matrix $\boldsymbol{I}$, $\boldsymbol{\varepsilon}_t \sim i.i.d. N(\boldsymbol{0}, \boldsymbol{I})$. The matrices $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ and $\boldsymbol{I}$ define the first two conditional moments and thus specify the conditional joint density of the endogenous variables; consequently a structure is characterized by $a = (\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{I})$ where $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ specify the contemporaneous and lagged structural relations among the endogenous variables and $\boldsymbol{I}$ is the covariance matrix of the unobservables. Each structure corresponds to a specific economic interpretation of the reduced form.

Two structures $a = (\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{I})$ and $b = (\boldsymbol{B}_0, \boldsymbol{B}_1, \boldsymbol{I})$ are observationally equivalent if they have the same reduced form, that is if these two conditions hold

$$\boldsymbol{A}_0^{-1}\boldsymbol{A}_1 = \boldsymbol{B}_0^{-1}\boldsymbol{B}_1 = \boldsymbol{\Pi} \qquad \boldsymbol{A}_0^{-1}\boldsymbol{A}_0'^{-1} = \boldsymbol{B}_0^{-1}\boldsymbol{B}_0'^{-1} = \boldsymbol{\Omega} \tag{2.4}$$

When this is the case, the economic explanations implied by the different structural interpretations of the reduced form system cannot be compared on empirical grounds.

We now ask the following: is there a relation among the different economic interpretations of the same reduced form system? Or, equivalently, in which way are the observationally equivalent structures related?

In the standard case when no covariance restrictions are imposed (see Koopmans, Rubin, and Leipnik, 1950) a necessary and sufficient condition for the equivalence of two structures is that they are connected through a linear invertible transformation; in other words the set of admissible transformations is $GL(p)$, the set of $p \times p$ invertible matrices.

As the next proposition shows, in our setup the set of admissible transformations is $O(p)$, the set of $p \times p$ orthogonal matrices:

Proposition 2.2. *A necessary and sufficient condition for the equivalence of two structures $a = (\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{I})$ and $b = (\boldsymbol{B}_0, \boldsymbol{B}_1, \boldsymbol{I})$ is that*

$$\boldsymbol{B}_i = \boldsymbol{O}'\boldsymbol{A}_i \quad \text{for } i = 0, 1$$

and [5] $\boldsymbol{O} \in O(p)$.

The set of observationally equivalent structures is thus defined as

$$S_I = \{(\boldsymbol{B}_0, \boldsymbol{B}_1) : \boldsymbol{B}_0 = \boldsymbol{O}'\boldsymbol{A}_0,\ \boldsymbol{B}_1 = \boldsymbol{O}'\boldsymbol{A}_1 \text{ and } \boldsymbol{O} \in O(p)\} \tag{2.5}$$

Two things are worth noticing: the first is that when we restrict all the structures to have the same covariance matrix $\boldsymbol{\Sigma}$ (in this case $\boldsymbol{\Sigma} = \boldsymbol{I}$) the set of admissible transformations is the set of orthogonal transformations $O(p)$ and not the set of invertible transformations $GL(p)$. This implies that the reduced form defines bounds that any $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ satisfies by construction: for any $\boldsymbol{O} \in O(p)$ it holds[6] that $\|\boldsymbol{O}\| = 1$, which is clearly not the case for the linear group $GL(p)$.

The next proposition states that any continuous function of an observationally equivalent structure is bounded[7]:

Proposition 2.3. *Let $f : S_\Sigma \to \mathbb{R}$ be a continuous function; then*

$$|f(\boldsymbol{A}_0, \boldsymbol{A}_1)| \leq c$$

for any $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$.

Impulse response functions, variance decompositions and more generally any continuous transformation of the system in (2.3) will be bounded by some quantity. These bounds are calculated directly from the reduced form in (2.2) before the imposition of identifying restrictions and hold for any structural interpretation of the system. Any set of identifying restrictions implies results that lie between these bounds. We use proposition 2.3 to calculate the reduced form bounds on the short-run structure of a cointegrated VAR in section 2.1 and on impulse response functions of a stationary system in section 2.2.

Second we notice that since $\boldsymbol{A}_0$ is invertible the set of observationally equivalent structures and the orthogonal group are isomorphic[8]: for any $\boldsymbol{O} \in O(p)$ there is one $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ and viceversa for any $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ there is one $\boldsymbol{O} \in O(p)$. The set of observationally equivalent structures has thus the same cardinality of the orthogonal group, that is there is a continuum of structures with a given $\boldsymbol{\Sigma}$ that can be derived from one reduced form system. Moreover, the existence of this bijection implies that we may equivalently think about identification as the process of selecting one structure in S_Σ or one orthogonal matrix in $O(p)$. This equivalence will be used extensively in section 3 when discussing the transformation of the reduced form system into a structural representation.

[5] $O(p)$ is the orthogonal group, the set of $p \times p$ orthogonal matrices. For the proof see propositions A.1 and A.2 in the appendix. Notice that the same result holds for any positive definite covariance matrix, not only for the identity.

[6] $\|\cdot\|$ indicates the L_2-norm for vectors and the spectral norm for matrices.

[7] For the proof see proposition A.3 in the appendix.

[8] The relation is a bijective even when the set of admissible transformations is the set of invertible transformations $GL(p)$, see Fisher (1966).

In what follows we will present a distinct exposition for integrated and stationary systems using the error correction representation for cointegrated systems and the moving average representation for stationary processes. The reason for this distinction resides exclusively in the different scope of the analysis and not in the different nature of the problem of observational equivalence in the two frameworks.

2.1 Reduced form bounds for cointegrated systems

In the cointegrated VAR framework Engle and Granger (1987), Johansen (1996) a great importance is given to the structural interpretation of the long-run relations and to the process of adjustment to them. In this setup we discuss the short-run dynamics and the long-run equilibrium relations of a given structural interpretation of the reduced form. In what follows the cointegration vectors are fixed and identified with the usual procedure (see Johansen, 1996); we deal exclusively with the identification of the short-run structure.

The reduced form error correction representation of (2.2) is

$$\Delta \boldsymbol{x}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{x}_{t-1} + \boldsymbol{u}_t \tag{2.6}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $p \times r$ matrices and $\boldsymbol{u}_t \sim i.i.d.\ N(\boldsymbol{0}, \boldsymbol{\Omega})$.

The main focus of the analysis is on the matrices $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ which are assumed to have rank $r < p$ and which define respectively the long-run economic relations (the cointegration vectors) and the process of adjustment to them.

Let $\boldsymbol{D}$ be a decomposition of $\boldsymbol{\Omega}$, that is let $\boldsymbol{\Omega} = \boldsymbol{D}\boldsymbol{D}'$; then the structural form of (2.6) implied by $\boldsymbol{D}$ is

$$\boldsymbol{A}_0 \Delta \boldsymbol{x}_t = \boldsymbol{A}_1 \boldsymbol{\beta}'\boldsymbol{x}_{t-1} + \boldsymbol{\varepsilon}_t \tag{2.7}$$

where $\boldsymbol{A}_0 = \boldsymbol{D}^{-1}$, $\boldsymbol{A}_1 = \boldsymbol{D}^{-1}\boldsymbol{\alpha}$ and $\boldsymbol{\varepsilon}_t \sim i.i.d.\ N(\boldsymbol{0}, \boldsymbol{I})$.

The set of observationally equivalent structures implied by the reduced form system in (2.6) is defined as

$$S_I = \{(\boldsymbol{A}_0, \boldsymbol{A}_1) : \boldsymbol{A}_0 = \boldsymbol{O}'\boldsymbol{D}^{-1},\ \boldsymbol{A}_1 = \boldsymbol{O}'\boldsymbol{D}^{-1}\boldsymbol{\alpha} \text{ and } \boldsymbol{O} \in O(p)\} \tag{2.8}$$

Notice that the set of equivalent structures is parameterized by the cointegrating space $\boldsymbol{\beta}$ (through $\boldsymbol{\alpha}$): the identification of the short-run structure involves the dynamic adjustment to the attractor defined by $\boldsymbol{\beta}$ and not the attractor itself. The structural coefficients in $\boldsymbol{A}_0, \boldsymbol{A}_1$ characterize the dynamic properties of the structural equations: their signs and magnitudes specify whether the structural system is error correcting or not and the velocities of adjustment to the long-run relations. It is with respect to these issues that the identification of the short-run structure is a crucial step.

We now apply proposition 2.3 to derive bounds for the short-run structural matrices $\boldsymbol{A}_0, \boldsymbol{A}_1$ in (2.7):

Proposition 2.4. *Let $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_I$ and denote with a^k_{ij} the generic element of the structural matrix $\boldsymbol{A}_k$, $k = 0, 1$; then for any i, j and k there exists a $c > 0$ such that*

$$|\, a^k_{ij}| \leq c \tag{2.9}$$

Any structure in the parameter space satisfies these bounds by construction; this means that any set of identifying restrictions defines values for the structural coefficients that lie between these bounds.

2.1.1 Example 1: A cointegrated system

We illustrate the ideas exposed in the previous section with the aid of one numerical example based on generated data[9]. In order to allow for the graphical representation of the set of

[9]See Appendix B for the numerical values employed. $\boldsymbol{\Sigma}$ is a diagonal matrix.

admissible transformations we set the dimension of the system to two; consequently the dimension of the cointegrating space is set to one, that is $p = 2$ and $r = 1$; it follows that $\boldsymbol{A}_0$ is a 2×2 matrix and $\boldsymbol{A}_1$ is a 2×1 vector.

Let $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ and denote with a^k_{ij} the generic element of the structural matrix $\boldsymbol{A}_k$; given the specific numerical values of $\boldsymbol{\alpha}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}$, the reduced form bounds defined in proposition 2.4 are reported in the next table:

Reduced form bounds				
$\boldsymbol{A}_0$	$\lvert a^0_{11}\rvert \le 1.36$	$\lvert a^0_{12}\rvert \le 1.84$	$\lvert a^0_{21}\rvert \le 0.66$	$\lvert a^0_{22}\rvert \le 0.89$
$\boldsymbol{A}_1$	$\lvert a^1_{11}\rvert \le 1.33$	$\lvert a^1_{21}\rvert \le 0.64$	-	-

Table 1: Reduced form bounds for the coefficients of $(\boldsymbol{A}_0, \boldsymbol{A}_1)$.

The table shows that as we let the structure $(\boldsymbol{A}_0, \boldsymbol{A}_1)$ vary in S_Σ, a^0_{11} varies in $[-1.36, 1.36]$, a^0_{12} in $[-1.84, 1.84]$, a^0_{21} in $[-0.66, 0.66]$ and a^0_{22} in $[-0.89, 0.89]$. Regarding $\boldsymbol{A}_1$, a^1_{11} varies in $[-1.33, 1.33]$ and a^1_{21} in $[-0.64, 0.64]$.

From the knowledge of $\boldsymbol{\alpha}$ and $\boldsymbol{D}$ (that is from the knowledge of $\boldsymbol{\Pi}$ and $\boldsymbol{\Omega}$) the calculation of these bounds is straightforward: from (2.8) the generic element of $\boldsymbol{A}_k$ is written as $a^k_{ij} = \boldsymbol{o}_i{}'\boldsymbol{m}_j$ where $\boldsymbol{o}_i$ is the i-th row of $\boldsymbol{O}$ and $\boldsymbol{m}_j$ is the j-th column of $\boldsymbol{M} = \boldsymbol{D}^{-1}$ for $k = 0$ and $\boldsymbol{M} = \boldsymbol{D}^{-1}\boldsymbol{\alpha}$ for $k = 1$; then $|a^k_{ij}| \le \|\boldsymbol{o}_i{}'\|\,\|\boldsymbol{m}_j\| = \|\boldsymbol{m}_j\| = c$.

No matter which a priori restrictions we impose to identify one structure, its coefficients lie in these bounds by construction: while just-identified systems are points in these segments under-identifying restrictions will select subsegments. Notice that the orthogonality of the columns of $\boldsymbol{O}$ imply that imposing restrictions on given coefficients transmit to others as well.

Even though $(\boldsymbol{A}_0, \boldsymbol{A}_1)$ have a direct structural interpretation we usually normalize the system with respect to some variables; that is $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$ become

$$\boldsymbol{N}_0 = \begin{bmatrix} 1 & a^0_{12}/a^0_{11} \\ a^0_{21}/a^0_{22} & 1 \end{bmatrix} = \begin{bmatrix} 1 & n^0_{12} \\ n^0_{21} & 1 \end{bmatrix} \text{ and } \boldsymbol{N}_1 = \begin{bmatrix} a^1_{11}/a^0_{11} \\ a^1_{21}/a^0_{22} \end{bmatrix} = \begin{bmatrix} n^1_{11} \\ n^1_{21} \end{bmatrix} \tag{2.10}$$

and discuss the structural interpretability in terms of these matrices[10]. We will use this formulation when imposing economic restrictions in section 3.1.1.

2.2 Reduced form bounds for stationary systems

In stationary VAR models inference is usually derived as impulse responses and variance decompositions of innovations which have been normalized to have unitary variance. Different decompositions of $\boldsymbol{\Omega}$ imply different impulse response functions; just think about the number of possible Cholesky decompositions: we have $p\,!$ possible orderings and thus the same number of impulse response functions. Of course the number is much larger if we don't restrict ourselves to triangular decompositions.

Let $\boldsymbol{D}$ be a decomposition of $\boldsymbol{\Omega}$, that is let $\boldsymbol{\Omega} = \boldsymbol{D}\boldsymbol{D}'$; then the impulse propagation scheme implied by $\boldsymbol{D}$ is

$$\boldsymbol{R}_s = \boldsymbol{\Pi}^s\boldsymbol{D} \tag{2.11}$$

where the element of row i and column j of $\boldsymbol{R}_s$ describes the response of variable i to the j-th innovation s periods after its occurrence.

[10] Notice that this normalization excludes the structures with a zero a^0_{11} and a zero a^0_{22} and that the off diagonal elements are not bounded anymore.

If $\boldsymbol{D}$ decomposes $\boldsymbol{\Omega}$ then any $\boldsymbol{C} = \boldsymbol{D}\boldsymbol{O}$ with $\boldsymbol{O} \in O(p)$ is such that $\boldsymbol{C}\boldsymbol{C}' = \boldsymbol{\Omega}$; then we write

$$\boldsymbol{R}_s = \boldsymbol{T}_s \boldsymbol{O} \tag{2.12}$$

where $\boldsymbol{T}_s = \boldsymbol{\Pi}^s \boldsymbol{D}$ and $\boldsymbol{O} \in O(p)$.

A given impulse propagation sequence $r = \{\boldsymbol{R}_s\}_{s \geq 0}$ is derived by applying the sequence of linear transformations $\{\boldsymbol{T}_s\}_{s \geq 0}$ on a specific $\boldsymbol{O} \in O(p)$; if we let $\boldsymbol{O}$ vary in $O(p)$ we describe the space of impulse propagation sequences

$$R = \{r : \boldsymbol{R}_s = \boldsymbol{T}_s \boldsymbol{O},\ \boldsymbol{O} \in O(p) \text{ and } s = 0,\ 1, \cdots\} \tag{2.13}$$

As an application of proposition 2.3 we derive the reduced form bounds for the impulse responses:

Proposition 2.5. *Let $r \in R$ and denote with r^s_{ij} the generic element of the response matrix $\boldsymbol{R}_s$, $s = 0,\ 1, \cdots$; then for any i, j and s there exists a $c > 0$ such that*

$$|\, r^s_{ij}| \leq c \tag{2.14}$$

No matter which theoretical restrictions we impose to identify the system and no matter which innovation we consider, the impulse propagation scheme lies between these bounds.

2.2.1 Example 2: A stationary system

The dimension of the system[11] is set to three, so that $\boldsymbol{R}_s$ is a 3×3 matrix for each s. Let $r \in R$ and denote with r^s_{ij} the generic element of the response matrix $\boldsymbol{R}_s$, $s = 0, 1, \cdots$.

In the next figure we display the reduced form bounds defined in proposition 2.5:

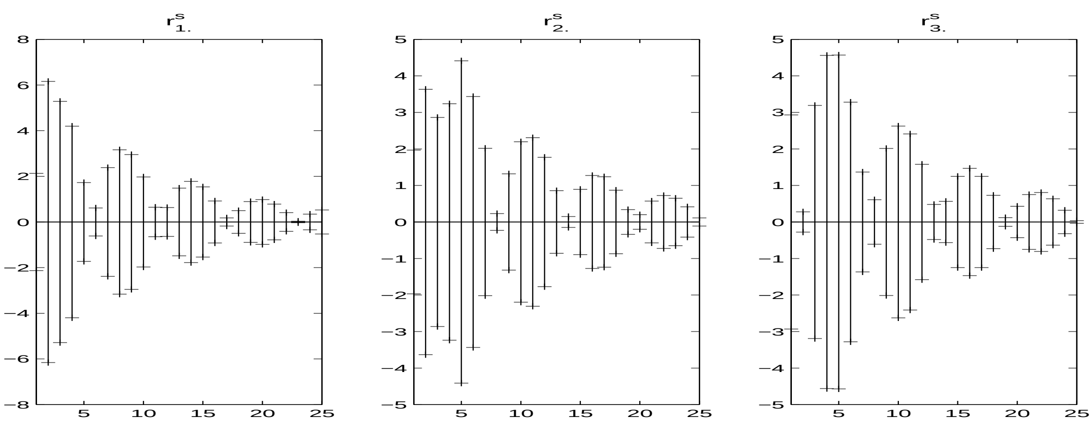

Figure 1: The reduced form bounds for the response of variable x_1, x_2 and x_3 to *any* shock implied by *any* equivalent structure.

We know that any set of just and under-identifying restrictions implies an impulse response that belongs to these intervals; thus in one glance we have a clear picture of where any impulse propagation scheme will lie. Moreover, as in this case, if the eigenvalues of $\boldsymbol{\Pi}$ are complex the difference of response implied by alternative structures at given periods is very small so that the different identification priors don't matter that much; take for example periods 6, 11, 12, 17 for variable one and 8, 14, 19, 20 for the second variable.

[11]See Appendix B for the numerical values employed.

3 A priori inequality restrictions and structural bounds

The choice of one $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$, the identification of the structure, is the very crucial and central choice in the analysis of a system of simultaneous equations; the results we derive from that system are reliable if and only if the identifying restrictions are correct. This is indeed a very delicate issue since the information in the data determines a set of uncountably many observationally equivalent structures and the criterion of choice has to be found out of the observational information in the field of the a priori theoretical impositions.

In economics the most settled assumptions are in the form of inequalities; there are explicit statements about signs and bounds of structural coefficients about which competing economic theories agree: parameters are either positive or negative and propensities are between zero and one; there are even intuitive indications about the qualitative effects of innovations that may be employed: for example, a restrictive monetary innovation raises short term interest rates. This qualitative knowledge may be analytically imposed in order to give a structural interpretation of the reduced form system. Conditionally on these priors we discriminate between structural and non structural representations.

If we denote with $S_P = \{(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma :$ the restrictions are satisfied$\}$ the set containing the structural representations and with S_P^C the set that collects the non structural interpretations of the system it follows that $S_P \cup S_P^C = S_\Sigma$ and $S_P \cap S_P^C = \emptyset$. Different economic priors imply different partitions of the set of observationally equivalent structures.

When we impose just-identifying restrictions on S_Σ, the set of structural representations is a singleton $S_P = \{(\boldsymbol{A}_0, \boldsymbol{A}_1)\}$ and consequently the set of non structural interpretations consists in all the space but one point $S_P^C = S_\Sigma \setminus \{(\boldsymbol{A}_0, \boldsymbol{A}_1)\}$. On the contrary, systems of inequality restrictions select more than one structural interpretation; the set S_P is thus larger and its complement S_P^C smaller than their counterparts defined by just-identifying restrictions. Given the isomorphism between S_Σ and $O(p)$ implied by proposition 2.2 we define the set of structural orthogonal matrices as the inverse image of S_P in $O(p)$, that is we define $O_P = \{\boldsymbol{O} \in O(p) : (\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_P\}$.

As in the standard cases of just and over-identifying restrictions, impulse responses are derived from the set of structural representations S_P; unlike those the results are derived from a set of structures and presented as bounds on the coefficients of interest. We refer to them as structural bounds in order to stress their difference from the reduced form bounds defined in section 2. They are calculated as follows

$$c_{ij} = \min_{\boldsymbol{O} \in O_P} \boldsymbol{e}_i' \boldsymbol{M} \boldsymbol{e}_j \qquad \text{and} \qquad d_{ij} = \max_{\boldsymbol{O} \in O_P} \boldsymbol{e}_i' \boldsymbol{M} \boldsymbol{e}_j \tag{3.1}$$

where $\boldsymbol{e}_i$ is the $p \times 1$ vector with zeros everywhere except at the i-th element, $\boldsymbol{M}$ is the $p \times l$ matrix which specifies the coefficients of interest and $\boldsymbol{e}_j$ is the $l \times 1$ vector with zeros everywhere except at the j-th element. Notice that even though O_P has been defined imposing restrictions on *given* magnitudes, by the orthogonality among the columns of $\boldsymbol{O}$ these restrictions affect the other coefficients as well.

As before, we will use the error correction representation for cointegrated systems and the moving average representation for stationary processes; the reason resides in the different scope of the analysis and not in the different nature of the problem in the two frameworks.

For cointegrated systems we compute the structural bounds for the matrices $\boldsymbol{A}_k$, $k = 0, 1$ in (2.8); in this case $\boldsymbol{M} = \boldsymbol{A}_k$, $k = 0, 1$, that is $\boldsymbol{M} = \boldsymbol{O}'\boldsymbol{D}^{-1}$ for $k = 0$ and $\boldsymbol{M} = \boldsymbol{O}'\boldsymbol{D}^{-1}\boldsymbol{\alpha}$ for $k = 1$. For stationary systems we compute the structural bounds for the response matrices $\boldsymbol{R}_s$, $s = 0,\ 1, \cdots$ in (2.13); in this case $\boldsymbol{M} = \boldsymbol{T}_s\boldsymbol{O}$ for $s = 0,\ 1, \cdots$.

3.1 Restrictions on signs and magnitudes of structural coefficients

This line of research was already indicated but never pursued in the Cowles Commission approach: "A further class of a priori restrictions that can often be based on economic considerations is inequalities. Frequently, the sign of the coefficients $\boldsymbol{\beta}_{gi\tau}$ or $\gamma_{gk\tau}$ (*the sign of the structural coefficients*) is known beforehand. (...) In the present article we do not study the question of how to give effect to restrictions of this kind." (Koopmans, Rubin, and Leipnik, 1950).

Let $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ and denote with a^k_{ij} the generic element of the structural matrix $\boldsymbol{A}_k$, $k = 0, 1$; we define economically plausible a structure that satisfies the following system of inequalities

$$l^k_{ij} \leq \; a^k_{ij} \leq u^k_{ij} \tag{3.2}$$

for given i, j and k where $l^k_{ij} \leq u^k_{ij} \in \mathbb{R}$ are the bounds implied by the theory[12].

According to these priors the set of structural representations is defined as $S_P = \{(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma : (3.2)$ is satisfied for given i, j and $k\}$. Given the bijection between S_Σ and $O(p)$ we define the set of structural orthogonal matrices as the inverse image of S_P in $O(p)$, that is we define $O_P = \{\boldsymbol{O} \in O(p) : (\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_P\}$.

3.1.1 Example 1 continued

Suppose our theoretical model specifies the following restrictions on the sign of the elements of $\boldsymbol{A}_0$ and on its normalized version $\boldsymbol{N}_0$ in (2.10):

$$\boldsymbol{A}_0 = \begin{bmatrix} + & - \\ - & + \end{bmatrix} \qquad \boldsymbol{N}_0 = \begin{bmatrix} 1 & -1 \leq n^0_{12} \leq 0 \\ n^0_{21} & 1 \end{bmatrix} \tag{3.3}$$

meaning that $a^0_{11} > 0$, $a^0_{12} < 0$, $a^0_{21} < 0$, $a^0_{22} > 0$ and $a^0_{11} + a^0_{12} \geq 0$; we leave n^0_{21} and all the coefficients in $\boldsymbol{N}_1$ unrestricted; the set of structural representations S_P is defined as $S_P = \{(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma : (3.3)$ are satified$\}$ and the associated set of structural transformations O_P is depicted in the next figure[13]:

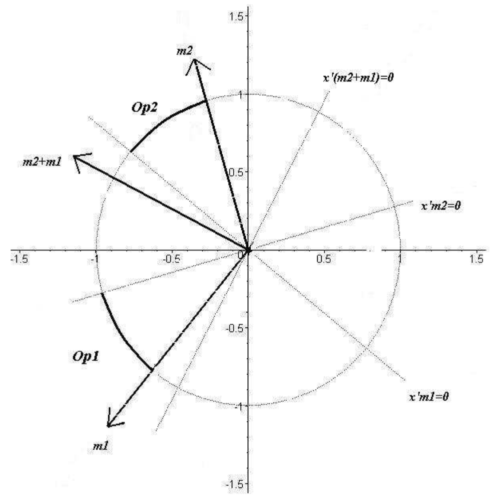

Figure 2: The geometric representation of the set of economically plausible transformations.

From fig.2 we visualize how restrictive the theoretical impositions that define S_p are: any point on the circumference is a mathematically admissible transformation but only the ones lying in O_{P1} and O_{P2} fulfill the economic criterions we have imposed to select a structural interpretation. This fact will be discussed in detail in section 4 when introducing

[12] Letting $l^k_{ij} = u^k_{ij} = 0$ we have the usual exclusion restrictions.

[13] O_{P1} and O_{P2} are the regions where the first and the second column of $\boldsymbol{O}$ are restricted to lie.

a criterion to compare alternative restrictions. Here we only consider the implications of these restrictions on all the coefficients of the system.

Let $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_P$ where S_P is defined by the theoretical restrictions in (3.3) and denote with a_{ij}^k the generic element of the matrix $\boldsymbol{A}_k$; the structural form bounds defined in (3.1) are reported in the next table:

Structural form bounds				
$\boldsymbol{A}_0$	$1.11 \leq a_{11}^0 \leq 1.36$	$-1.05 \leq a_{12}^0 \leq 0$	$-0.37 \leq a_{21}^0 \leq 0$	$0.73 \leq a_{22}^0 \leq 0.89$
$\boldsymbol{A}_1$	$0.23 \leq a_{11}^1 \leq 0.94$	$-0.63 \leq a_{21}^1 \leq -0.46$	-	-

Table 2: Structural form bounds for the coefficients of $(\boldsymbol{A}_0, \boldsymbol{A}_1)$.

In the set of structural representations, a_{11}^0 is restricted to vary in $[1.11, 1.36]$, a_{12}^0 in $[-1.05, 0]$, a_{21}^0 in $[-0.37, 0]$ and a_{22}^0 in $[0.73, 0.89]$. Recall that the reduced form bounds (see table 1) were $[-1.36, 1.36]$, $[-1.84, 1.84]$, $[-0.66, 0.66]$ and $[-0.89, 0.89]$ respectively: the theoretical restrictions imposed in (3.3) are clearly able to decrease the length of the bounds significantly. With regard to $\boldsymbol{A}_1$, the restrictions in (3.3) constrain a_{11}^1 to vary in $[0.23, 0.94]$ instead that in $[-1.33, 1.33]$ and a_{21}^1 in $[-0.63, -0.46]$ instead that in $[-0.64, 0.64]$. Notice that the restrictions were exclusively imposed on $\boldsymbol{A}_0$ and $\boldsymbol{N}_0$. Any interpretation of the reduced form system that satisfies that specific priors imply that the adjustment coefficient of the first variable to the cointegrating relation is positive and that of the second negative. Furthermore we notice that for a_{21}^1 the bounds are very tight and thus very informative about the velocity of the error correction mechanism. Thus the very loose restrictions in (3.3) allow for a clear interpretation of the short-run dynamical behavior of the system.

In the next table we summarize the results:

Bounds	a_{11}^0	a_{12}^0	a_{21}^0	a_{22}^0	a_{11}^1	a_{21}^1
Reduced	±1.36	±1.84	±0.66	±0.89	±1.33	±0.64
Structural	[1.11,1.36]	[-1.05,0]	[-0.37,0]	[0.73,0.89]	[0.23,0.94]	[-0.63,-0.46]

Table 3: A summary of the analysis: reduced and structural bounds.

3.2 Restrictions on the effects of innovations

Let $r \in R$ and denote with r_{ij}^s the generic element of the response matrix $\boldsymbol{R}_s$, $s = 0, 1, \cdots$; we define economically plausible an impulse propagation scheme that satisfies the following system of inequalities

$$l_{ij}^s \leq r_{ij}^s \leq u_{ij}^s \tag{3.4}$$

for given i, j and s where $l_{ij}^s \leq u_{ij}^s \in \mathbb{R}$ are the bounds we impose on the responses[14].

The set of admissible impulse response schemes is thus defined as $R_P = \{r \in R : (3.4) \text{ is satisfied for given } i, j \text{ and } s\}$; in terms of orthogonal matrices we have $O_P = \{\boldsymbol{O} \in O(p) : r \in R_P\}$.

Notice that this is actually the generalization of the standard way of proceeding to achieve identification: we may impose restrictions on the impact effect of the innovations on the variables (Sims, 1980), on the long-run (Blanchard and Quah, 1989) or at any other point of time (Faust, 1998, Uhlig, 1999, Canova and De Nicoló, 2000).

[14]This is how Uhlig (1999) defines a monetary policy innovation, one which has a non positive effect on prices and a non negative effect on short term interest rate for some periods.

3.2.1 Example 2 continued

Suppose we define economically interesting an innovation that has a contemporaneous non positive effect on the second and on the third variables in the system and we are interested in studying its effect on the first variable; that is we impose that $r_2^0 \leq 0$ and $r_3^0 \leq 0$ and we leave r_1^s unrestricted at every s. Then the set of economically admissible impulse response schemes R_P is defined as $R_P = \{r \in R : r_2^0 \leq 0 \ \text{ and } \ r_3^0 \leq 0\}$ and the associated set of structural transformations O_P is depicted in the next figure:

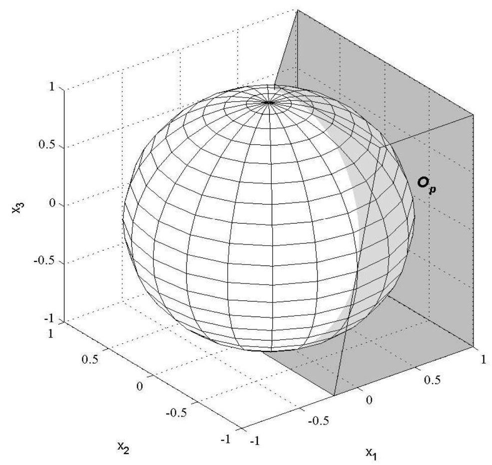

Figure 3: The geometric representation of the set of economically plausible transformations.

Almost half of the transformations satisfy the restrictions (grey part of fig.3).

It is striking to see how these very loose assumptions are able to clearly determine the propagation scheme of that innovation (see the next figure):

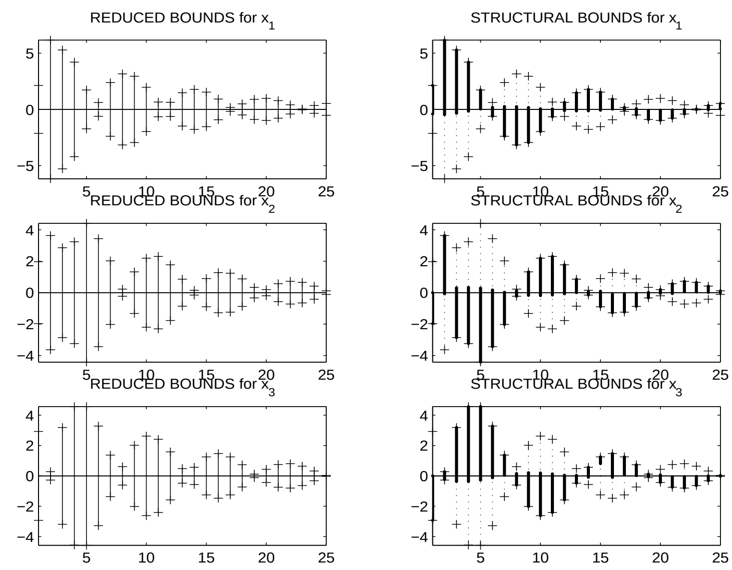

Figure 4: The comparison between the reduced and the structural bounds.

Notice the difference in the plots on the left hand side (reduced form bounds) and the ones on the right (structural bounds) for the response of each variable to an innovation that belongs to R_P: the imposition of a negative *contemporaneous* reaction of variable 2 and 3 is able to trace out a clear propagation scheme for that innovation for all s and for any variable. The response of the first and third variables to this specific innovation is positive for the first five periods, then negative for the following six periods and then it oscillates towards zero as time passes. When the innovation occurs the response of the second variable is negative, positive after one period and then it oscillates to die out. The very loose assumptions that define the innovation are clearly able to determine the sign of the impulse propagation scheme at every period for the three variables.

4 Can we compare alternative inequality restrictions?

By definition every observationally equivalent structure has the same reduced form and thus the same likelihood: this implies that there is no possibility of testing just and under-identified structures. Thus we have to reason differently: suppose that the true structure that has generated the observed data belongs to the set of observationally equivalent structures; in other words consider satisfied the assumptions about linearity, additivity and about the distribution of the unobservables made in section 2. From a strictly mathematical point of view and exit economic theory the fact that any $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ has the same likelihood means that there is no information in the data that can be used to discriminate among equivalent structures; this implies that *before* the imposition of economic restrictions each structure has the same probability of being the true one. For this reason we argue that the set of observationally equivalent structures can be conceived as a measure space with a uniform measure: it contains uncountably many structures which have the same probability of being the true one[15]. Identifiable structures are observationally equivalent and as such they are as plausible as any other element of S_Σ.

Technically, we supplement $O(p)$ with the Borel algebra on $\mathbb{R}^p$ and with the uniform measure on the surface of the sphere[16] and for convenience we normalize it to have measure one. The set of observationally equivalent structures is thus described as the probability space derived from $(\mathbb{R}^n, \mathbb{B}_n, \mu)$ through the invertible transformation defined in proposition 2.2.

The next definition characterizes μ explicitly

Definition 4.1. *In the probability space $(\mathbb{R}^n, \mathbb{B}_n, \mu)$, μ is an absolutely continuous probability measure with density function*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \begin{cases} \frac{1}{S_p} & \textit{if } \|\boldsymbol{x}\| = 1 \\ 0 & \textit{otherwise} \end{cases}$$

where $S_p = \frac{2\pi^{p/2}}{\Gamma(p/2)}$ is the total surface area of the unit sphere centered at zero.

Suppose we have specified the theoretical priors and defined the implied set of structural orthogonal transformations O_P; there are then three possibilities: either the theoretical restrictions are such that any $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ satisfies them or there is no such $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ or there are some $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$ for which the restrictions hold and some others for which they don't. In terms of structural transformations we have that $O_P = O(p)$, $O_P = \emptyset$ and $\emptyset \neq O_P \subset O(p)$ respectively.

[15]The standard case of just-identifying restrictions puts all weight on one structure; in this sense uniform measure and just-identification are two opposite cases. The intermediate case would be to weight differently alternative structures specifying a non uniform measure.

[16]For any $\boldsymbol{O} \in O(p)$ it holds that $\boldsymbol{o}_i'\boldsymbol{o}_j = 1$ if $i = j$ and $\boldsymbol{o}_i'\boldsymbol{o}_j = 0$ if $i \neq j$; the geometric representation of $O(p)$ is thus the surface of the unitary sphere centered at zero in $\mathbb{R}^p$, $\mathbb{S}_p = \{\boldsymbol{x} \in \mathbb{R}^p : \|\boldsymbol{x}\| = 1\}$.

In the case $O_P = O(p)$ the bounds implied by the reduced form are stronger than the ones specified by the theoretical priors. In this case what we theoretically think as assumptions are actually not such when we confront the model with the data and this could possibly lead to a revision of our theoretical priors.

Suppose instead that the set of theoretical restrictions implies an empty O_P: in this case we may exclude the validity of those theoretical assumptions and have a complete falsification of the theory on empirical grounds[17].

In the intermediate case the set of structural transformations becomes a proper subset of the orthogonal group and its measure becomes smaller than one. If the theoretical restrictions are satisfied by a great number of equivalent structures, the set of structural transformations O_P is big and its measure close to one. On the contrary, if the priors select only a small number of equivalent structures the set O_P is small and its measure close to zero. The higher the number of structures that satisfy the a priori restrictions the greater the degree of their generality and for this reason the higher the probability that the *true* structure belongs to S_P. In this way we construct a method to compare the restrictiveness of the a priori assumptions expressed as inequalities which is exclusively based on the reduced form system.

Over identifying restrictions change the value of the likelihood function and allow for a statistical comparison of alternative structures; on the contrary, just and under-identified systems are observationally equivalent and no statistical test is possible.

The next table summarizes the previous arguments:

Systems	**Statistical Tests**	**Measure-theoretical Comparisons**
Over-identified	YES	NO
Just-identified	NO	NO
Under-identified	NO	YES

Table 4: The two ways of comparing structures.

4.1 Example 2 continued

Suppose that there exists an alternative theory that defines the economic innovation in section 3.2.1 in terms of its contemporaneous non positive effect on the first and on the third variables in the system; that is we impose that $r_1^0 \leq 0$ and $r_3^0 \leq 0$ and we leave r_2^s unrestricted at every s. We want to compare these restrictions with those defined in section 3.2.1, that is with $r_2^0 \leq 0$ and $r_3^0 \leq 0$ and r_1^s unrestricted at every s.

Notice that the two sets of restrictions imply the same response of the third variable ($r_3^0 \leq 0$) but they differ with respect to the other assumption: in the first set of priors we require that the contemporaneous response of the second variable is non positive, that is $r_2^0 \leq 0$; in the second case we impose that $r_1^0 \leq 0$.

According to the first set of priors the set of economically admissible impulse response schemes R_{P_1} is defined as $R_{P_1} = \{r \in R : r_2^0 \leq 0 \text{ and } r_3^0 \leq 0\}$ and the associated set of structural transformations O_{P_1} is depicted in fig.3.

According to the second set of priors the set R_{P_2} is defined as $R_{P_2} = \{r \in R : r_1^0 \leq 0 \text{ and } r_3^0 \leq 0\}$ and the implied impulse response scheme is depicted in the next figure in the plots on the right hand side; on the left hand side plots we report the impulse propagation scheme implied by R_{P_1}:

[17]This statement is actually true only if we knew the population moments. With sample quantities we should introduce some statistical criterions; in the present article we avoid this issue completely.

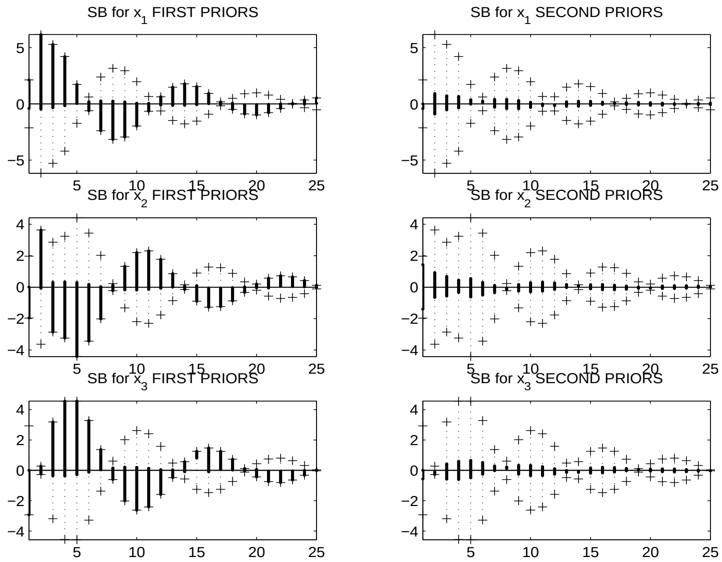

Figure 5: The structural bounds implied by the first and second set of priors.

On the left we visualize the impulse propagation scheme already discussed in section 3.2.1: the imposition of a negative *contemporaneous* reaction of variable 2 and 3 is able to determine the sign of the impulse propagation scheme at every period for the three variables. On the right hand side we visualize the impulse propagation scheme of an innovation that satisfies the second set of priors; even in this case the impulse response structure is clearly determined: an innovation that satisfies the second set of priors, that is the imposition of a negative *contemporaneous* reaction of variable 1 and 3, will have no effect on the variables. The structural bounds are in fact very short and concentrated around the zero line.

Now we ask the following: which of the two propagation schemes is to be trusted?

The set of structural transformations O_{P_2} associated to the second set of priors is depicted in the next figure:

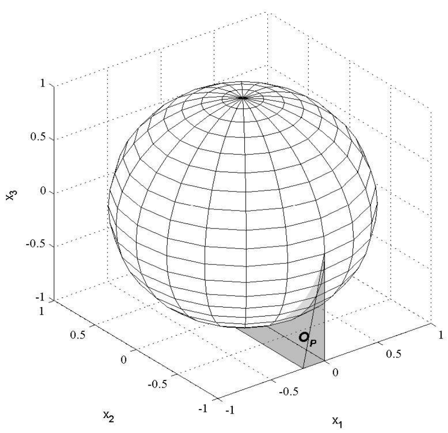

Figure 6: The geometric representation of the set of economically plausible transformations according to the second set of priors.

As it is clear from a visual comparison of fig.3 and fig.6 the set of structural transformations O_{P_1} is much bigger than O_{P_2}, meaning that there are many more structures that satisfy the first set of impositions rather than the second.

More precisely, the measure of the set of structural representations according to the first set of priors is $\mu(S_{P_1}) = 0.37$, meaning that roughly 40% of the mathematically admissible structures is selected by that specific theoretical priors. According to the second set of priors the measure of the set of structural representations is $\mu(S_{P_2}) = 0.02$: only the 2% of equivalent structures satisfies the second set of priors.

For this reason we conclude that the assumptions that define the first of priors are much more general than the others and thus the impulse propagation scheme implied by the first set of priors is the one to be trusted.

5 Conclusion

The main concern of this paper is to show that in the VAR framework the transformation of a reduced form system into a structural representation can rely on a priori restrictions expressed an inequalities.

Inequality restrictions are much more general than linear restrictions and it is often the case that competing paradigms agree on signs and bounds of some coefficients while they don't agree on exclusions. In this sense a system of inequality restrictions can provide a core of well established a priori impositions on which deriving a structural interpretation of the reduced form system. As showed in the examples the selection of a small structural set is a sufficient but not a necessary condition to trace out a clear impulse-propagation scheme for the system. Just identified systems trace out a perfectly defined propagation scheme; but this is completely wrong if the identifying restrictions are wrong.

There is in this sense a trade off between generality of assumptions and certainty of results; it is our opinion that under-identified systems provide a good compromise between the two needs. Whether we are able or not to trace out a clear propagation scheme for an innovation it depends *both* on the characteristics of the system and on the theoretical priors we have imposed on it. Once again results are a mix of observational information and theoretical impositions.

A Proofs

Under the assumption of normality of $\boldsymbol{u}_t$, the conditional joint distribution of the endogenous variables is completely specified by the first two moments; thus it holds that $P(\boldsymbol{x}_t|a,\boldsymbol{x}_{t-1}) = P(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1}) \Leftrightarrow E(\boldsymbol{x}_t|a,\boldsymbol{x}_{t-1}) = E(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1})$ and $Var(\boldsymbol{x}_t|a,\boldsymbol{x}_{t-1}) = Var(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1})$.

According to the structures $a = (\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{\Sigma})$ and $b = (\boldsymbol{B}_0, \boldsymbol{B}_1, \boldsymbol{\Sigma})$ the conditional moments are

$$E(\boldsymbol{x}_t|a,\boldsymbol{x}_{t-1}) = \boldsymbol{A}_0^{-1}\boldsymbol{A}_1\boldsymbol{x}_{t-1} \qquad E(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1}) = \boldsymbol{B}_0^{-1}\boldsymbol{B}_1\boldsymbol{x}_{t-1} \tag{A.1}$$

and

$$Var(\boldsymbol{x}_t|a,\boldsymbol{x}_{t-1}) = \boldsymbol{A}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{A}_0'^{-1} \qquad Var(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1}) = \boldsymbol{B}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{B}_0'^{-1} \tag{A.2}$$

Proposition A.1. *A necessary and sufficient condition for the equivalence of two structures $a = (\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{\Sigma})$ and $b = (\boldsymbol{B}_0, \boldsymbol{B}_1, \boldsymbol{\Sigma})$ is that they are connected through a transformation which belongs to $E_\Sigma = \{\boldsymbol{L} : \boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L} = \boldsymbol{\Sigma}\}$. That is*

$$a \sim b \Leftrightarrow \boldsymbol{B}_i = \boldsymbol{L}'\boldsymbol{A}_i \quad \text{for } i = 0, 1$$

and $\boldsymbol{L} \in E_\Sigma$.

Proof. ($\Rightarrow$) Let $a \sim b$; then there exists $\boldsymbol{L} \in E_\Sigma$ such that $\boldsymbol{B}_i = \boldsymbol{L}'\boldsymbol{A}_i$ for $i = 0, 1$. In fact $\boldsymbol{\Omega} = \boldsymbol{B}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{B}_0'^{-1} = \boldsymbol{B}_0^{-1}\boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L}\boldsymbol{B}_0'^{-1} = \boldsymbol{A}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{A}_0'^{-1}$ holds if and only if $\boldsymbol{B}_0 = \boldsymbol{L}'\boldsymbol{A}_0$. Moreover $E(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1}) = \boldsymbol{A}_0^{-1}\boldsymbol{L}'^{-1}\boldsymbol{B}_1\boldsymbol{x}_{t-1} = \boldsymbol{A}_0^{-1}\boldsymbol{A}_1\boldsymbol{x}_{t-1}$ if and only if $\boldsymbol{B}_1 = \boldsymbol{L}'\boldsymbol{A}_1$.

($\Leftarrow$) Let $\boldsymbol{B}_0 = \boldsymbol{L}'\boldsymbol{A}_0$ and $\boldsymbol{B}_1 = \boldsymbol{L}'\boldsymbol{A}_1$ with $\boldsymbol{L} \in E_\Sigma$; then $E(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1}) = E(\boldsymbol{x}_t|a,\boldsymbol{x}_{t-1})$ and $Var(\boldsymbol{x}_t|b,\boldsymbol{x}_{t-1}) = Var(\boldsymbol{x}_t|a,\boldsymbol{x}_{t-1})$. □

Proposition A.2. *Let $E_\Sigma = \{\boldsymbol{L} : \boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L} = \boldsymbol{\Sigma}\}$. Then*

$$\boldsymbol{L} \in E_\Sigma \Leftrightarrow \boldsymbol{L} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{O}\boldsymbol{\Sigma}^{1/2}$$

and $\boldsymbol{O} \in O(p)$.

Proof. ($\Rightarrow$) Let $\boldsymbol{L} \in E_\Sigma$; from $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{L}'\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{L}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{O}'\boldsymbol{O}$ it follows that $\boldsymbol{L} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{O}\boldsymbol{\Sigma}^{1/2}$ and $\boldsymbol{O} \in O(p)$.

($\Leftarrow$) Let $\boldsymbol{L} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{O}\boldsymbol{\Sigma}^{1/2}$ and $\boldsymbol{O} \in O(p)$; then

$$\boldsymbol{L}'\boldsymbol{\Sigma}\boldsymbol{L} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{O}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{O}\boldsymbol{\Sigma}^{1/2} = \boldsymbol{\Sigma}$$

and thus $\boldsymbol{L} \in E_\Sigma$. □

Proposition 2.2 in the text is derived combining propositions A.1 and A.2 for the particular case $\boldsymbol{\Sigma} = \boldsymbol{I}$.

Proposition A.3. *Let $f : S_\Sigma \to \mathbb{R}$ be a continuous function; then*

$$|f(\boldsymbol{A}_0, \boldsymbol{A}_1)| \leq c$$

for any $(\boldsymbol{A}_0, \boldsymbol{A}_1) \in S_\Sigma$.

Proof. f is a continuous function defined on a compact domain; for Weierstrass Theorem it has a maximum and a minimum. Since $\boldsymbol{O} \in O(p) \Leftrightarrow -\boldsymbol{O} \in O(p)$ we have that $\min_{(\boldsymbol{A}_0,\boldsymbol{A}_1)\in S_\Sigma} f(\boldsymbol{A}_0, \boldsymbol{A}_1) = -\max_{(\boldsymbol{A}_0,\boldsymbol{A}_1)\in S_\Sigma} f(\boldsymbol{A}_0, \boldsymbol{A}_1) = c$ and thus that $|f(\boldsymbol{A}_0, \boldsymbol{A}_1)| \leq c$. □

Propositions 2.4 and 2.5 in the text are a particular case of A.3.

B Numerical values of the examples

Cointegration analysis:

$$\boldsymbol{\alpha} = \begin{bmatrix} 0.21 \\ -0.63 \end{bmatrix} \qquad \boldsymbol{\Omega} = \begin{bmatrix} 0.79 & 0.33 \\ 0.33 & 0.43 \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 0.98 & 0 \\ 0 & 0.23 \end{bmatrix} \tag{B.1}$$

Stationary system:

$$\boldsymbol{\Pi} = \begin{bmatrix} 1.153 & -0.487 & -1.030 \\ -0.179 & -0.464 & -1.132 \\ 0.561 & -0.081 & 0.429 \end{bmatrix} \qquad \boldsymbol{\Omega} = \begin{bmatrix} 4.546 & -2.936 & -6.130 \\ -2.936 & 3.878 & 3.981 \\ -6.130 & 3.981 & 8.590 \end{bmatrix} \tag{B.2}$$

The eigenvalues of $\boldsymbol{\Pi}$ are $0.77 \pm 0.47i$ and -0.43 which are in modulus less than one.

References

Blanchard, O., and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–673.

Canova, F., and G. De Nicoló (2000): "Monetary disturbances matter for the business fluctuations in the G-7," International Finance Discussion Paper 660, Board of Governors of the Federal Reserve System.

Chesher, A. (2003): "Identification in Non Separable Models," *Econometrica*, 71, 1405–1441.

Engle, R., and C. Granger (1987): "Co-integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251–276.

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

Favero, C. (2001): *Applied Macroeconometrics*. Oxford University Press.

Fisher, F. (1959): "Generalization of the Rank and Order Conditions for Identifiability," *Econometrica*, 27, 431–447.

——— (1963): "Uncorrelated Disturbances and Identifiability Criteria," *International Economic Review*, 4, 134–152.

——— (1966): *The identification problem in econometrics*. McGraw-Hill.

Frisch, R. (1934): "Statistical Confluence Analysis by Means of Complete Regression Systems," Publication no.5, Oslo: Universitetes Økonomiske Institutt.

Haavelmo, T. (1943): "The Statistical Implications of a system of Simultaneous equations," *Econometrica*, 11, 1–12.

——— (1944): "The Probability Approach in Econometrics," *Econometrica*, 12, supplement, 1–115.

Hurwicz, L. (1950): "Generalization of the Concept of Identification," *in Statistical Inference in Dynamic Economic models*, Cowles Commission Monograph 10, New York, John Wiley.

Johansen, S. (1995): "Identifying restrictions of linear equations with applications to simultaneous equations and cointegration," *Journal of Econometrics*, 69, 111–132.

——— (1996): *Likelihood-based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford University Press.

KOOPMANS, T. (1949): "Identification Probelms in Economic Model Construction," *Econometrica*, 17, 125–144.

KOOPMANS, T., AND O. REIESOL (1950): "The Identification of Structural Characteristics," *The Annals of Mathematical Statistics*, 21, 165–181.

KOOPMANS, T., H. RUBIN, AND R. LEIPNIK (1950): "Measuring the Equation Systems of Dynamic Economics," *in Statistical Inference in Dynamic Economic models*, Cowles Commission Monograph 10, New York, John Wiley.

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

SIMS, C. (1980): "Macroeconomics and reality," *Econometrica*, 48, 1–48.

UHLIG, H. (1999): "What are the effects of monetary policy on output? Results from an agnostic identification procedure," Discussion paper 2137, CEPR.

WALD, A. (1950): "A Note on Identification of Economic Relations," *in Statistical Inference in Dynamic Economic models*, Cowles Commission Monograph 10, New York, John Wiley.