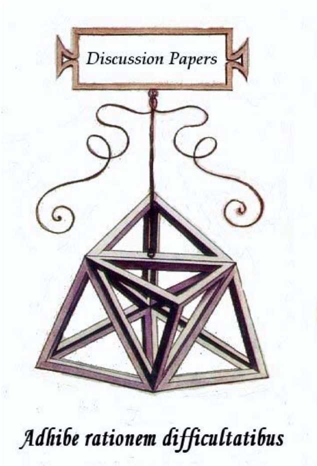
Discussion Papers
Collana di
E-papers del Dipartimento di Scienze Economiche – Università di Pisa



Annetta Maria Binotti - Enrico Ghiani

Interpreting reduced form cointegrating vectors of incomplete systems. A labour market application_

Discussion Paper n. 28

2004

Discussion Paper n. 28, presentato: **febbraio 2003**

Indirizzi dell'Autore:
Dipartimento di scienze economiche, via Ridolfi 10, 56100 PISA
fax: (39 +) 050 598040

Si prega di citare così:

Annetta Maria Binotti - Enrico Ghiani, "Interpreting reduced form cointegrating vectors of incomplete systems. A labour market application", Discussion Papers del Dipartimento di Scienze Economiche – Università di Pisa, n. 28 (http://www-dse.ec.unipi.it/ricerca/discussion-papers.htm).

Discussion Paper
n. 28

Annetta Maria Binotti - Enrico Ghiani

Interpreting reduced form cointegrating vectors of incomplete systems. A labour market application*

Abstract

This paper explores the issues arising when the reduced form cointegrating vectors are obtained from an incomplete VAR with omitted endogenous variables. Reconsidering some Wickens' (1996) results, we show that the specification of an incomplete VAR model, based on variables of the reduced form cointegrating vectors, produces an approximation of the long run coefficients. We also show that in certain circumstances this problem does not arise. An example concerning the estimation of the NAIRU is provided suggesting the empirical relevance of our criticism. We use a numerical example with simulated data to illustrate the potential pitfalls.

Keywords: Cointegration; Vector autoregression; Identification; Unemployment.
JEL classification: C32; C51; E24

*Sections II and III are a substantially revised version of Binotti and Ghiani (2001).

Contents

I. Introduction

This paper calls for the need to proceed cautiously in the interpretation of the estimation of reduced form cointegration relations using an incomplete VAR model in which the endogenous variables are omitted; a further problem is found in the use of these estimates to engage in conjectures on the long-term structure of the model.

Building on some of the problems outlined by Wickens (1996), we show that when the empirical specification of an incomplete VAR system is based on the variables that appear in the reduced form cointegrating vectors, this implies recourse to an approximation. The incomplete system cointegrating coefficients may differ considerably from coefficients of the corresponding cointegrating vectors of the complete system, and if the latter is not known there is no way to determine the extent of the approximation. In the framework of these problems, one important empirical application is constituted by NAIRU (Non-Accelerating Inflation Rate of Unemployment) estimations. Such estimations make reference to the models of R.Layard and S.Nickell (1985, 1991) and use cointegration analysis[1].

Overall, our analysis focuses on the circumstance, well-known in the literature, that the a priori of complete system structure[2] theory constitute a necessary basis for dynamic specification. This remains true even when, reduced form dynamic formulas are considered adequate for the particular research objectives and estimation of the structural model is therefore not the objective that is being pursued.

This paper is organised as follows. Section II, building on the model proposed by Wickens (1996), introduces some isomorphic forms of ECM (Error Correction Model), in order to provide the analytical bases of the work; we will show, furthermore, that the practice of estimating an incomplete VAR based on reduced form cointegrating vectors can lead to a false specification of the dynamics of the model and to an approximation of the cointegrating vector coefficients. In section III we discuss the empirical importance of the criticisms put forward in the work, examining the example of the NAIRU estimation; in section IV we use a small numerical example with simulated cointegrated data to illustrate the approximation. The concluding remarks then follow.

[1] An initial pioneering contribution was made by Jenkins (1988), while a recent application has been offered by the article of Brunello et al. (2000).

[2] The terms " structure" and " structural" are used here with the meaning indicated by the Cowles Commission on " relations or parameters that present a direct economic interpretation" .

II. Reduced form cointegrating vectors and incomplete models

Let x_t be an $N \times 1$ vector composed of $I(1)$ random variables with the following autoregressive representation of kth-order:

$$\Pi(L)x_t = v_t \quad , \quad v_t \sim IN(0, \Omega), \tag{1}$$

where $\Pi(L)$ is an $N \times N$ matrix of polynomials in the lag operator L, $\Pi(L) = \sum \Pi_j L^j$ and v_t is a vector of independently, identically distributed random variables with mean zero and covariance matrix Ω. Let n denote the number of linearly independent cointegration relations and m denote the number of linearly independent $I(1)$ processes in x_t. Partitioning $x_t = \ (y_t' \quad z_t')'$, where y_t is an $n \times 1$ vector of endogenous variables and z_t an $m \times 1$ vector of exogenous variables, we write (1) as

$$A(L)y_t + B(L)z_t = \tilde{u}_t \quad , \quad \tilde{u}_t \sim IN(0, \Sigma) \tag{2}$$

$$C(L)\Delta z_t = e_t \sim I(0). \tag{3}$$

Equation (2) can be rewritten as the conditional ECM

$$A_0 \Delta y_t + \tilde{A}^*(L)\Delta y_{t-1} + B^*(L)\Delta z_t + A(1)y_{t-1} + B(1)z_{t-1} = \tilde{u}_t, \tag{4}$$

where Δ is the difference operator, $A(L) = A_0(1-L) + \tilde{A}^*(L)L(1-L) + A(1)L$ and $B(L) = B^*(L)(1-L) + B(1)L$; moreover, A_0 is a $n \times n$ non-singular matrix with $a_{ii,0} = 1$ for all $i = 1, ..., n$. The hypothesis of exogeneity of the variables z_t implies it is possible to consider the model (4) as complete[3]. Equations (4) and (3) can now be combined to give

$$\Pi^*(L)\Delta x_t + \Pi_S x_{t-1} = u_t, \quad u_t = [\tilde{u}_t \quad e_t]', \tag{5}$$

where $\Pi^*(L)\Delta x_t$ and $\Pi_S x_{t-1}$ may be viewed as the short run and long run relations between y_t and z_t. Premultiplying (4) by A_0^{-1} we obtain another conditional model, isomorphic to (4):

$$\Delta y_t + A_0^{-1}\tilde{A}^*(L)\Delta y_{t-1} + A_0^{-1}B^*(L)\Delta z_t +$$

$$+ A_0^{-1}\left[A(1)y_{t-1} + B(1)z_{t-1}\right] = A_0^{-1}\tilde{u}_t. \tag{6}$$

Equations (6) and (3) can now be combined to give

$$\Pi^{**}(L)\Delta x_t + \Pi_R x_{t-1} = v_t, \quad v_t = \left[A_0^{-1}\tilde{u}_t \quad e_t\right]'. \tag{7}$$

Equation (7) is the basis for Johansen (1988, 1991), Johansen and Juselius's (1990) complete system cointegration analysis. If the data generation process

[3]The following analysis requires only weak exogeneity in the sense of Engle et al. (1983); however, for ease of presentation of the arguments, in (3) we have also hypothesized strong exogeneity of the variables z_t

is (6) and (3) the parameters in $\Pi^{**}(L)$ and Π_R accept the following a priori restrictions:

$$\Pi^{**}(L) = \begin{bmatrix} I + A_0^{-1}\tilde{A}^*(L) & A_0^{-1}B^*(L) \\ 0 & C(L) \end{bmatrix},$$

$$\Pi_R = \begin{bmatrix} A_0^{-1}A(1) & A_0^{-1}B(1) \\ 0 & 0 \end{bmatrix},$$

then the rank of Π_R is $r = n$ and it can be factorised as $\Pi_R = \alpha\beta'$, where α is an $(n+m) \times n$ matrix of factor loadings and β is an $(n+m) \times n$ matrix of cointegrating vectors. We now wish to discuss briefly the following admissible formulations of matrix Π_R:

$$\Pi_R = \alpha\beta_S' = \begin{bmatrix} A_0^{-1}a_c \\ 0 \end{bmatrix} \begin{bmatrix} \beta_y' & \beta_z' \end{bmatrix}, \tag{8}$$

$$\Pi_R = \alpha^*\beta_R' = \begin{bmatrix} \alpha_c^* \\ 0 \end{bmatrix} \beta_R' = \begin{bmatrix} A_0^{-1}A(1) \\ 0 \end{bmatrix} \begin{bmatrix} I_n & \beta_y'^{-1}\beta_z' \end{bmatrix}, \tag{9}$$

where $\beta_S' = (\beta_y' \quad \beta_z')$, with β_y an $n \times n$ non-singular matrix, denotes the long-run structural parameters identified with Johansen's procedure[4]. Moreover, α is partitioned into a non-singular $n \times n$ matrix[5] $\alpha_c = A_0^{-1}a_c$ and a null $m \times n$ matrix of loading coefficients in the conditional and marginal model respectively; α_c^* may be viewed as the adjustment of n jointly dependent variables to the deviation from the equilibrium in the last period, where equilibrium is defined as: $y^e = A(1)^{-1}B(1)z$. The partitioning of Π_R proposed in (9) may be considered more appealing because it does not pose the problem of identifying multiple cointegrating vectors; moreover α_c^* can be obtained from the first $n \times n$ partition of Π_R estimated in equation (7). In contrast, in order to estimate (8) it is necessary to identify β_S' and to have a prior estimate of A_0 which involves identification of the short-run dynamics. We would also add the following comments concerning the latter issue of avoiding the 'problem' of identification.

[4]The minimum condition required is that the configuration given in β_S' satisfies the exact identification conditions defined by Johansen's theorem (1992). As specified in Hsiao (1997) and Davidson (1998), if one considers the long run structure there is no difference between the problem of identification as it is classically established by the Cowles Commission and the way in which it is configured in the cointegration analysis.

[5]If one hypothesizes that the matrix a_c is diagonal, the notion of structural ECM corresponds to that of Boswijk (1995). On the basis of such a definition an ECM is defined as structural if, within the specifications that regard variable $I(1)$, the following hypothesis are respected:

i)the variables z_t are weakly exogenous with respect to the parameters in question β;

ii)the endogenous number is equal to the rank of Π_R: $n = r$;

iii) a_c e β_y' are at full rank;

iv)β_S is identified by a set of restrictions and the loading matrix a_c is diagonal. The elements on the diagonal are not zero A_0 normalized with the unitary elements on the diagonal.

i) In the representation of Π_R given by (9), the cointegrating vectors are identified in such a way that each of the n variables enters one and only one cointegrating vector and, as it involves reduced form coefficients, it is more difficult to give an economic interpretation to the relations obtained.

ii) The traditional interpretation of α_c^*, as an adjustment matrix which contains the adjustment speeds will be misleading since it contains the long-run structural parameters[6]β_y'. Moreover, even if we consider a model with a triangular β_y' and with a diagonal loading matrix α_c, α_c^* will lose the diagonality characteristics that formed part of the original loading matrix; therefore, even in the specification of the incomplete system it is necessary to load the disequilibrium relations of all the other endogenous variables. Consequently it becomes indispensable to resort to the complete structural model, which is precisely what the adoption of specification (9) is usually designed to avoid.

The latter observation allows us now to turn to the central question of this work. In some cases the goal of empirical research is limited to obtaining an estimation of reduced form cointegrating vectors. However, if it is true that this practice facilitates avoidance of identification problems[7], it is also true that we will be unable to produce inferences on the structure by starting out from the cointegrating vectors obtained. In practice, the VECM founded on the subsets of variables contained in the reduced form cointegrating vectors will be an approximation of the dynamic specification and of the long-run cointegration relations that could be obtained from the complete system.

Let y_t be partitioned as $(y_{1t}' \quad y_{2t}')'$, respectively of order $n_1 \times 1$ and $n_2 \times 1$, moreover by partitioning the structural model conformably with y_t, we will have

$$\begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} B_1(L) \\ B_2(L) \end{bmatrix} z_t = \begin{bmatrix} \tilde{u}_{1t} \\ \tilde{u}_{2t} \end{bmatrix}. \tag{10}$$

given the matrices $A_{22}^{-1}(L)$, y_{2t} can be solved in terms of y_{1t} and and z_t; then, substituting into the first block, the model can be written as

$$\begin{aligned} &\left[A_{11}(L) - A_{12}(L)A_{22}^{-1}(L)A_{21}(L)\right] y_{1t} + \\ &+ \left[B_1(L) - A_{12}(L)A_{22}^{-1}(L)B_2(L)\right] z_t = \tilde{u}_{1t} - A_{12}(L)A_{22}^{-1}(L)\tilde{u}_{2t}. \end{aligned} \tag{11}$$

Lags appearing in equation (11) are of infinite order, the errors are an infinite-order moving average $MA(\infty)$ and the steady state is the reduced form cointegrating vectors for y_{1t}. However, in practice the empirical specification of an incomplete VECM in y_{1t} and z_t will be of finite lag order and will therefore be an approximation of (11). There clearly exists a difficulty in controlling the adequacy

[6]This can be considered as an extension of what Rossana (1998) demonstrated on the problem of normalisation.

[7]See, for example, the empirical analysis inspired by Layard and Nickell's models that are concerned with the labour market. In particular the works by Jenkinson (1988) and Brunello et al.. (2000)

of this approximation, because it is not possible to discover the long-run structural form of the complete model from the estimates of a finite approximation of a VECM reparametrization of (11).

Nevertheless, a particular case exists in which the representation of an incomplete VECM is not an approximation. If β_y' is a triangular matrix and a_c is a diagonal loading matrix, then the resulting loading matrix α_c^* will be triangular. For this reason the first n_1 equations do not contain the remaining n_2 endogenous variables, neither current nor lagged. Therefore, it is possible to obtain a correct dynamic specification for Δy_{1t} through an incomplete VECM specified for y_{1t} and z_t without falling into approximation.

III. The estimates of the NAIRU in wage setting and price setting models

We can now exemplify the problem put forward in the previous pages by examining an important case addressed by the macroeconomic literature: the estimation of NAIRU in imperfect competition models that also have wage bargaining. In their most recent developments, such models are linked to the names of R.Layard and S.Nickell[8].

A formulation of the model that is quite representative of this literature is transcribed here below[9]:

$$\begin{aligned}
y_t &= c_1(m_t - p_t) + c_2 g_t, && (12)\\
u_t &= -c_3 y_t, && (13)\\
p_t - w_t &= -h(u_t, u_{t-1}...) - a_1\,(p_t - p_t^e) + Z_{1,t}, && (14)\\
w_t - p_t^e &= -f'(u_t, u_{t-1}...) + d_1'\,(w_{t-1} - p_{t-1}) + Z_{2,t}, && (15)\\
m_t &= m_{t-1} + \varepsilon_{m,t}, && (16)\\
g_t &= \overline{g} + \varepsilon_{g,t}, && (17)\\
Z_{1,t} &= \overline{Z}_1 + \varepsilon_{Z_1,t}, && (18)\\
Z_{2,t} &= \overline{Z}_2 + \varepsilon_{Z_2,t}, && (19)
\end{aligned}$$

where: y_t = real GDP, m_t = nominal money supply, g_t =exogenous component of real demand, w_t = money wages, p_t = price index, p_t^e = expected price index, $Z_{1,t}$ = all exogenous factors influencing the mark up of prices over wages, u_t = unemployment rate, $Z_{2,t}$ = all exogenous factors influencing wages (unemployment benefits, fiscal variables, exchange rates etc). All the variables are expressed in logarithms. Functions $h(.)$ and $f'(.)$ describe the influence of the level of economic activity in the price setting and the wage setting equation respectively; these functions contemplate the possibility of persistence mechanisms or hysteresis through lags in unemployment rate values. Finally, $\varepsilon_{m,t}, \varepsilon_{g,t}, \varepsilon_{Z_1,t}, \varepsilon_{Z_2,t}$ are

[8] Layard et al. (1985), Layard et al. (1991).

[9] We make reference to, for example, the work of S.Nickell (1988), (1998).

serially uncorrelated shocks and $\overline{g}, \overline{Z}_1, \overline{Z}_2$ represent the anticipated component of the corresponding variables. As we are dealing with a well-known model we will omit the description of its function. A set of constraints usually imposed in literature [10] is the following:

$$h(u_t, u_{t-1}...) = 0, \tag{20}$$

$$f'(u_t, u_{t-1}...) = d_2 u_t + d_{22} \Delta u_t, \tag{21}$$

$$(w_{t-1} - p_{t-1}) = (w_t - p_t) - \zeta, \tag{22}$$

where $\zeta = \ln(1 + \zeta_{w-p})$ and ζ_{w-p} is the uniform growth-rate of real wages. (20) expresses the hypothesis, supported by the results from empirical work[11], that the real wage of price setting is independent of the labour market conditions. (22), for simplicity's sake, imposes uniform growth conditions.

Based on these hypotheses, equations (14) and (15) can be written as follows:

$$p_t - w_t = -a_1 \left(p_t - p_t^e\right) + Z_{1,t}, \tag{23}$$

$$w_t - p_t = -d_1 \left(p_t - p_t^e\right) - d_2 u_t - d_{22} \Delta u_t + Z_{2,t}, \tag{24}$$

where we have eliminated the constant and posited $d_1 = (1 - d_1')^{-1}$.

The importance of this model for an empirical analysis in the framework of cointegration is evident: it presents a set of theoretically founded structural, dynamic equations and with steady state equilibrium solutions that have an immediate economic interpretation. The main path to follow therefore consists in specifying an unconstrained 'closed' VAR in the variables that enter into the long-run equilibrium relations. This path has been followed in many works[12]; nevertheless the objective of identifying the complete system of the equilibrium relations has not been convincingly reached[13]. An alternative pathway followed in the literature is represented by estimation of the reduced form of an incomplete system. In the example considered here, we directly searched for the cointegration relation suggested by the solution of the system for the unemployment rate, eliminating the real wage from the variables that are important for equilibrium analysis (but also in specification of the dynamics).

Therefore, the following structural system is considered to be compatible with the steady state (23)-(24):

[10] Nickell (1998), Brunello et al. (2000), Carlin et al. (1990), Blanchard et al. (1997).

[11] See the literature cited in the previous note

[12] Clements et al. (1991), Durby et al. (1993), Mizon (1995), Marcellino et al. (1995), Johansen et al. (1997), Jacobson et al. (1998), Chiarini et al. (1998), Binotti et al. (1999).

[13] The reason for this is generally found in the problems of identification of wage setting and price setting equations notoriously present in the labour market model of Layard and Nickell. The problem arises because $Z_{1,t}$ is a sub set of $Z_{2,t}$. See Manning (1993).

$$\begin{aligned} (w-p)_t &= a_{11,1}(w-p)_{t-1} + b_1 Z_{1,t-1} + \xi_{w-p,t}, \\ a_{21,0}(w-p)_t + u_t &= a_{21,1}(w-p)_{t-1} + a_{22,1}u_{t-1} + \\ &+ b_2 Z_{1,t-1} + b_3 Z_{2,t-1} + \xi_{u,t}, \\ Z_{1,t} &= Z_{1,t-1} + \xi_{Z_1,t}, \\ Z_{2,t} &= Z_{2,t-1} + \xi_{Z_2,t}, \end{aligned} \tag{25}$$

Cointegration analysis is generally conducted in a non-constrained VECM sphere, where, by limiting the analysis to a single lag and imposing the necessary constraints on the matrix β'[14] to identify the reduced form of the complete system, we have

$$\Delta y_t = -\alpha_c^* \beta_R' x_{t-1} + \varepsilon_t, \tag{26}$$

$$\begin{bmatrix} \Delta(w-p)_t \\ \Delta u_t \end{bmatrix} = -\begin{bmatrix} 1-a_{11,1} & 0 \\ -a_{21,0}\left(1-a_{11,1}\right) + a_{21,0} - a_{21,1} & 1-a_{22,1} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & -b_1(1-a_{11,1})^{-1} & 0 \\ 0 & 1 & \left[(a_{21,0}-a_{21,1})b_1(1-a_{11,1})^{-1} - b_2\right](1-a_{22,1})^{-1} & -b_3(1-a_{22,1})^{-1} \end{bmatrix}$$

$$\begin{bmatrix} (w-p)_{t-1} \\ u_{t-1} \\ Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{w-p,t} \\ \varepsilon_{u,t} \end{bmatrix}.$$

In contrast[15], in order to estimate the NAIRU, we estimate a VECM system with a finite order of lags, specified solely in the variables that appear in the cointegrating vector normalized for unemployment and expressed in reduced form:

$$\widetilde{\Pi}^{**}(L)\Delta\widetilde{x}_t + \widetilde{\Pi}_R \widetilde{x}_{t-1} = \epsilon_t, \quad \text{where } \widetilde{x}_t = \begin{bmatrix} u_t & Z_{1,t} & Z_{2,t} \end{bmatrix}'.$$

[14] The restriction imposed is that of an identity matrix in the first block of β'.

[15] It is easier to understand the origin of the approximation if we derive the solution for u_t with successive substitutions of $(w-p)_{t-j}$ $(j = 0, 1, 2, ...)$ in the second equation of the complete system (25), obtaining

$$\begin{aligned} u_t = {} & a_{22,1}u_{t-1} + (b_2 - a_{21,0}b_1)Z_{1t-1} + (a_{21,1}b_1 - a_{21,0}a_{11,1}b_1)Z_{1t-2} + \\ & + (a_{21,1}a_{11,1}b_1 - a_{21,0}a_{11,1}^2 b_1)Z_{1t-3} + ... + b_3 Z_{2,t-1} + \\ & + MA(\xi_{w-p,t}) + \xi_{u,t} \ , \end{aligned} \tag{1}$$

where $MA(\xi_{w-p,t})$ is an infinite-order moving average of $\xi_{w-p,t}$. The steady state of (1) will have the following configuration:

$$\begin{aligned} u = {} & [b_2 + (1 + a_{11,1} + a_{11,1}^2 + ...)a_{21,1}b_1 - \\ & - (1 + a_{11,1} + a_{11,1}^2 + ...)a_{21,0}b_1](1 - a_{22,1})^{-1}Z_1 + \\ & + b_3(1 - a_{22,1})^{-1}Z_2 \ , \end{aligned} \tag{2}$$

In the example given here, the approximation consists in neglecting the infinite sum of the complete multiplication process linked to the inertia of the system. In addition, the two effects of the procedure adopted are quite evident: an approximation of the reduced form cointegrating vector is determined on the basis of the incomplete VECM and an erroneous specification of the short-run dynamics is produced.

Since the extent of the approximation is not controllable, the example given clearly illustrates the possible consequences in terms of erroneous specification of the dynamics and approximation of the cointegrating vector coefficients. Such consequences spring from the estimation of reduced forms of incomplete systems. The erroneousness of the procedure assumes particular importance when the estimated cointegrating vectors form the basis for inferences concerning the coefficients of the corresponding structural forms.

The analysis that has been conducted also gives rise to another interesting suggestion, namely that in empirical practice it may be useful to resort to a high number of lags in order to obtain an acceptable approximation of the VECM dynamic specification and coefficients of the reduced form cointegrating vectors of the complete system. The implicit limits of this procedure should however be borne in mind[16].

As we have shown, there are cases where, due to the a priori of the theory, the circumstances that bring about the results illustrated here are considered as irrelevant. If for example we expect the model to be endowed with recursiveness, it will be possible to make use of formulations of incomplete systems for the endogenous variables placed at the beginning of the recursive chain. In the example considered, as can easily be verified, this is the case for the real wage. Nevertheless, bearing in mind the criticisms levelled by Wickens (1996), it must not be forgotten that cointegration analysis should not be affected by this type of a priori in the specification phase of the initial VAR.

IV. A numerical example with simulated cointegrated data

We use a numerical example to highlight the main point being made in this paper that in general, when the empirical specification of an incomplete VAR system is based on the variables that appear in the reduced form cointegrating vectors, this implies recourse to an approximation. We show that a) exists a

---

from which

$$u = \left[b_2 + b_1(1 - a_{11,1})^{-1}(a_{21,1} - a_{21,0})\right](1 - a_{22,1})^{-1} Z_1 + \\ + b_3(1 - a_{22,1})^{-1} Z_2 \ . \tag{3}$$

[16] This precaution is not easy to apply due to the loss of degrees of freedom that affects a VAR specification with a high number of lags. It is in fact not applied in the work by Brunello et al. (2000), which fixes a maximum number of lags in one period.

difficulty in controlling the adequacy of this approximation if we do not know the estimated complete system, b) we incur in misspecification of the long run dynamic matrix. 100 observations were generated from the following data generating process (DGP):

$$\begin{bmatrix} 1 & 0 \\ -0.2 & 1 \end{bmatrix} \begin{bmatrix} (w-p)_t \\ u_t \end{bmatrix} = \begin{bmatrix} 0.5 & 0 \\ -0.9 & 0.9 \end{bmatrix} \begin{bmatrix} (w-p)_{t-1} \\ u_{t-1} \end{bmatrix} + \\ + \begin{bmatrix} 0.5 & 0.2 \\ 0.8 & 0 \end{bmatrix} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{w-p,t} \\ u_{u,t} \end{bmatrix},$$

$$\begin{bmatrix} u_{w-p,t} \\ u_{u,t} \end{bmatrix} = \begin{bmatrix} \epsilon_{(w-p),t} \\ \epsilon_{u,t} \end{bmatrix} + \begin{bmatrix} -0.2 & -0.4 \\ -0.4 & 0.2 \end{bmatrix} \begin{bmatrix} \epsilon_{Z_1,t} \\ \epsilon_{Z_2,t} \end{bmatrix}.$$

In the first equation the real wage depends on real wage observed in the previous period and on both lagged strongly exogenous variables $Z_{1,t-1}$ and $Z_{2,t-1}$; in the second equation unemployment depends on both lagged endogenous variables and on one strongly exogenous process $Z_{1,t-1}$. The system is closed by the random walks:

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} = \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{Z_1,t} \\ \epsilon_{Z_2,t} \end{bmatrix},$$

Where $\epsilon_{.,t} \sim n.i.i.d.$ The equilibrium coefficients of DGP are given by the following cointegrating vectors:

$$\beta_S' = \begin{bmatrix} 1 & 0.14 & -1.14 & 0 \\ 1 & 0 & -1 & -0.4 \end{bmatrix},$$

$$\beta_R' = \begin{bmatrix} 1 & 0 & -1 & -0.4 \\ 0 & 1 & -1 & 2.86 \end{bmatrix}.$$

Moreover, the corresponding loadings and long run matrix are:

$$\alpha = \begin{bmatrix} 0 & -0.5 \\ -0.8 & 0 \end{bmatrix}, \alpha_c^* = \begin{bmatrix} -0.5 & 0 \\ -0.8 & -0.11 \end{bmatrix},$$

$$\Pi_R = \begin{bmatrix} -0.5 & 0 & 0.5 & 0.2 \\ -0.8 & -0.11 & 0.91 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The Johansen cointegration procedure[17] confirms the presence of two cointegrating vectors. Next we impose the identifying restrictions on the steady state by exclusion of $Z_{2,t}$ from wage setting, and exclusion of u_t from price setting, we also

[17] We estimate a congruent $VAR(1)$ without identification of adjustment or steady state. First we use the OLS estimation with the following diagnostics of the system: the $F-test$ of no serial correlations against second order, the χ^2 test for normality, the test for no autoregressive conditional heteroscedasticity and the test of the 'functional form' which is again a tests for

impose an over-identifying restriction on the coefficient of $Z_{1,t}$, and the restriction is not rejected ($\chi^2(1) = 1.5649[0.2109]$). Finally, we impose the restrictions on the loading matrix obtaining the following estimates (standard errors are in parentheses):

$$\widehat{\alpha}\widehat{\beta}'_S x_{t-1} = \begin{bmatrix} 0 & \underset{(0.06)}{-0.51} \\ \underset{(0.05)}{-0.79} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & \underset{(0.01)}{0.14} & \underset{(0.01)}{-1.14} & 0 \\ 1 & 0 & -1 & \underset{(0.03)}{-0.39} \end{bmatrix} \begin{bmatrix} (w-p)_{t-1} \\ u_{t-1} \\ Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix},$$

which corresponds to (8) in section II, and it is obtained after imposing non-rejected over-identifying restrictions that are tested using a LR statistic ($\chi^2(7) = 2.242[0.9452]$). We also obtain the representation of Π_R given by (9), where the cointegrating vectors are identified in such a way that each of the two variables enters one and only one cointegrating vector involving reduced form coefficients:

$$\widehat{\alpha}^*\widehat{\beta}'_R x_{t-1} = \begin{bmatrix} \widehat{\alpha}^*_c \\ 0 \end{bmatrix} \widehat{\beta}'_R x_{t-1} =$$

$$= \begin{bmatrix} \underset{(0.06)}{-0.51} & 0 \\ \underset{(0.07)}{-0.81} & \underset{(0.01)}{-0.11} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & \underset{(0.03)}{0.39} \\ 0 & 1 & \underset{(0.03)}{-0.99} & \underset{(0.27)}{2.79} \end{bmatrix} \begin{bmatrix} (w-p)_{t-1} \\ u_{t-1} \\ Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix}.$$

The over-identifying restrictions are tested obtaining $\chi^2(6) = 1.8277[0.9348]$. With reference to Δu_t, we note that $\widehat{\alpha}^*_c$ loses the diagonality characteristics that formed part of the original loading matrix; therefore, even in the specification of the incomplete system it is necessary to load the disequilibrium relation of the excluded endogenous variables $(w-p)_t$.

The estimated reduced form long-run matrix is:

$$\widehat{\Pi}_R = \begin{bmatrix} -0.51 & 0 & 0.51 & 0.20 \\ -0.81 & -0.11 & 0.92 & 0.00 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{27}$$

---

the null of no heteroscedasticity (for details see J. Doornik and D. Hendry (1997).

$$Vec\ AR1-2F(32,289) = 0.93466[0.5726]$$
$$Vec\ normality\ \chi^2(8) = 3.1311[0.9259]$$
$$Vec\ x_i^2 F(80,458) = 1.2436[0.0896]$$
$$Vec\ X_iX_jF(140,550) = 1.1262[0.1778].$$

All of these calculations were performed with PcFIML 9.0 by J. Doornik and D. Hendry (1997).

Incomplete not congruent (Vector $AR1-2F(18,232) = 2.3113[0.0024]^{**}$) $VAR(1)$ system, based on the variables that appear in the reduced form cointegrating vectors $\widetilde{x}_t = \begin{bmatrix} u_t & Z_{1,t} & Z_{2,t} \end{bmatrix}$, produces the following results:

$$\begin{bmatrix} \underset{(0.013)}{-0.074} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \underset{(0.06)}{-1.18} & \underset{(0.37)}{2.54} \end{bmatrix} \begin{bmatrix} u_{t-1} \\ Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix}.$$

To eliminate the approximation we must estimate an incomplete congruent $VAR(6)$ system[18] (Vector $AR1-2F(18,189) = 0.53917[0.9364]$), which produces the following results:

$$\begin{bmatrix} \underset{(0.023)}{-0.11} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \underset{(0.046)}{-1.06} & \underset{(0.30)}{2.79} \end{bmatrix} \begin{bmatrix} u_{t-1} \\ Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix}.$$

In the absence of the approximation, one stray also into erroneous specification of the short-run dynamics; in fact, if we compare with the correct estimates given in (27), the first row of the estimated long run matrix is misleading in terms of the dynamic reaction of Δu with respect toZ_1 and Z_2:

$$\begin{bmatrix} -0.11 & 0.11 & -0.30 \end{bmatrix}.$$

On the contrary it is possible to make use of an incomplete system for the endogenous variables placed at the beginning of the recursive chain; in fact an incomplete congruent (Vector $AR1-2F(18,232) = 0.81635[0.6802]$) $VAR(1)$ system with $\widetilde{x}_t = \begin{bmatrix} (w-p)_t & Z_{1,t} & Z_{2,t} \end{bmatrix}$ produces the following results:

$$\begin{bmatrix} \underset{(0.06)}{-0.51} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -1 & \underset{(0.03)}{0.39} \end{bmatrix} \begin{bmatrix} (w-p)_{t-1} \\ Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix},$$

which is exactly the first row of the estimated long run matrix in the complete model:

$$\begin{bmatrix} -0.51 & 0.51 & 0.20 \end{bmatrix}.$$

Repeated simulation shows that these results are typical and are not due to the specific sample drawn nor are they dependent on the particular choice of residual covariance matrix.

[18] $VAR(6)$ is estimated to eliminate the approximation. In order to obtain the congruency is sufficient a $VAR(2)$.

V. Conclusions

The specification of an incomplete VAR based on reduced forms, excluding some of the endogenous variables which are part of the complete theoretical model, represents a manner of abstaining from use of the entire set of a priori theories. It may be supposed that in some cases[19] what leads to this procedure is the importance of modelling a subset of the model's endogenous variables. When, however, the coefficients of the cointegration relations thereby obtained are utilized to produce inferences on the structural form, one must deduce that rejection of full use of the theory's a priori derives from a different origin and springs from problems such as incomplete identification of the structure. In this work we highlight the fact that this (possibly unconscious) return to practices of measurement without theory is not devoid of drawbacks and, in particular, can lead to unfounded inferences on the model's structure. What we argue is that even when estimation of the structural system is not a directly pursued objective – because reduced form formulations suffice for the purposes of the study - the long run structure must guide the dynamic specification. In the absence of such guidance, one may stray into approximations when estimating the cointegration coefficients, and also into erroneous specification of the short-run dynamics. The empirical relevance of the problem has been shown with reference to cointegration analysis in NAIRU estimation. On this point, our findings suggest that the estimated coefficients of cointegration relations should be considered with due caution. They also suggests that the use of a suitable number of lags in the specification of the initial VAR could lead to an acceptable approximation of such coefficients.

Authors' affiliation
Dipartimento di Scienze Economiche, University of Pisa
Dipartimento di Scienze Economiche, University of Pisa

[19]For example, Jenkinson (1988).

Appendix

A Data

The computer generated data-set utilized in this paper can be downloaded from the Discussion Papers' web site. The files are stored in the file *http://www-dse.ec.unipi.it/discussion-papers/lavori/dataGhianiBinotti.zip* in both ASCII and PcGive format. The ASCII files are human-readable when listed on a computer screen, and each is for a single variable, named $w - p.DAT$, $u.DAT$, $Z1.DAT$ and $Z2.DAT$. Readers possessing PcGive will find it more convenient to use the files called $SimulSys$ (a GiveWin file) and $SimulSys.bn7$ (a binary file)

References

Binotti, A. M., and E. Ghiani. (1999). "Analisi Empirica della Relazione tra Salari Reali, Produttività e Tasso di Disoccupazione in Italia", *Rivista Internazionale di Scienze Sociali* 4, 543-555.

Binotti, A. M.,and E. Ghiani. (2001). "Forme Ridotte e Modelli incompleti nell'analisi di cointegrazione. Il caso del NAIRU", *Statistica* 4, 678-689.

Blanchard, O., and L. F. Katz. (1997). "What We Know and Do Not Know About the Natural Rate of Unemployment", *Journal of Economic Perspectives* 11, 51-72.

Boswijk, H. P. (1995). "Efficient Inference on Cointegration Parameters in Structural Error Correction Models", *Journal of Economatrics* 69, 133-158.

Brunello, G., Lupi, C., and P. Ordine. (2000). "Regional Disparities and the Italian NAIRU", *Oxford Economic Papers*, 146-177.

Carlin, W., and D. Soskice. (1990). *Macroeconomics and the Wage Bargain: a Modern Approach to Employment, Inflation and the Excange Rate*. Oxford: Oxford University Press.

Chiarini, B., and P.Piselli. (1997). "Wage Setting, Wage Curve and Phillips Curve: the Italian Evidence", *Scottish Journal of Political Economy* 5, Vol. 44.

Clements, M.P., and G. E. Mizon. (1991). "Empirical Analysis of Macroeconomic time series. VAR and Structural Models", *European Economic Review*, vol. 35, 887-932.

Darby, J., and Wren-Lewis S. (1993). "Is There a Cointegrating Vector for UK Wages?", *Journal of Economic Studies*, vol. 20, 87-115.

Davidson, J. (1998). "Structural Relations, Cointegration and Identification: some Simple Results and their Application", *Journal of Economatrics*, *87*, 87-113.

Engle, R.F., Hendry, D.F., and J.F. Richard. (1983). "Exogeneity", *Econometrica*, 51, 227-304.

Hsiao, C. (1997). "Cointegration and Dynamic Simultaneous Equations Model", *Econometrica*, 3, Vol. 65, 647-669.

Jacobson, T., Vredin, A., and A. Warne. (1998). "Are Real Wages and Unemployment Related", *Economica*, 65, 69-96.

Jenkinson, T.J. (1988). "The NAIRU: Statistical Fact or Theoretical Straitjacket?", pp. 365-377 in Rod Cross (ed.), *Unemployment, Hysteresis and the Natural Rate Hypothesis*. Basil Blackwell, .

Johansen, K., and B. Strom. (1997). "Wages, Prices and Politics: Evidence from Norway", *Oxford Bulletin of Economics and Statistics*, 4, vol. 59,.

Johansen, S. (1988). "Statistical Analysis of Cointegrating Vectors", *Journal of Economic Dynamics and Control*, 231-244.

Johansen, S. (1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, *59, 1551-1580.*

Johansen, S. (1992). *Identifying Restrictions of Linear Equations*, Institute of Mathematics, University of Copenhagen. Preprint 1992, 4.

Johansen, S., and K. Juselius. (1990). "Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money", *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

Johansen, S., and K. Juselius. (1994). "Identification of the Long-run and the Short-run Structure. An Application to the IS-LM Model", *Journal of Econometrics*, 63, 7-36.

Layard, R., and S. Nikell. (1985). "The Causes of British Unemployment", *National Institute Economic Review*, 111, 62-85.

Layard, R., Nikell, S., and R. Jackman. (1991). *Unemployment: Macroeconomics Performance and the Labour Market*. Oxford: Oxford University Press.

Manning, A. (1993). "Wage Bargaining and the Phillips Curve: the Identification and Specification of Aggregate Wage Equations", *The Economic Journal*, 103, 98-118.

Marcellino, M., and G. E. Mizon. (1995). *Modelling Shifts in the Wage-Price and Unemployment Relationships in Italy, Poland, and the UK.*. Mimeo, 1-26.

Mizon, G. E. (1995). *Progressive Modelling of Macroeconomic Time Series: The LSE Methodology*. Mimeo. Economics Department, European University Institute, Florence.

Nickell, S. (1988). "The NAIRU: some Theory and Statistical Facts", in R. Cross (ed.) *Unemployment, Hysteresis and the Natural Rate Hypothesis*. Oxford: Basil Blackwell.

Nickell, S. (1998). "Unemployment: Questions and some Answers", *The Economic Journal*, 108, 802-816.

Rossana, R. J. (1998). "On the Adjustment Matrix in error Correction Models", *Journal of Monetary Economics*, 42, 427-444.

Wickens, M.R. (1996). "Interpreting Cointegrating Vectors and Common Stocastic Trends", *Journal of Econometrics*, 74, 255-271.

1. Luca Spataro, Social Security And Retirement Decisions In Italy, (luglio 2003)
2. Andrea Mario Lavezzi, Complex Dynamics in a Simple Model of Economic Specialization, (luglio2003)
3. Nicola Meccheri, Performance-related-pay nel pubblico impiego: un'analisi economica, (luglio 2003)
4. Paolo Mariti, The BC and AC Economics of the Firm, (luglio 2003)
5. Pompeo Della Posta, Vecchie e nuove teorie delle aree monetarie ottimali, (luglio 2003)
6. Giuseppe Conti, Institutions locales et banques dans la formation et le développement des districts industriels en Italie, (luglio 2003)
7. F. Bulckaen - A. Pench - M. Stampini, Evaluating Tax Reforms through Revenue Potentialities: the performance of a utility-independent indicator, (settembre 2003)
8. Luciano Fanti - Piero Manfredi, The Solow's model with endogenous population: a neoclassical growth cycle model (settembre 2003)
9. Piero Manfredi - Luciano Fanti, Cycles in dynamic economic modelling (settembre 2003)
10. Gaetano Alfredo Minerva, Location and Horizontal Differentiation under Duopoly with Marshallian Externalities (settembre 2003)
11. Luciano Fanti - Piero Manfredi, Progressive Income Taxation and Economic Cycles: a Multiplier-Accelerator Model (settembre 2003)
12. Pompeo Della Posta, Optimal Monetary Instruments and Policy Games Reconsidered (settembre 2003)
13. Davide Fiaschi - Pier Mario Pacini, Growth and coalition formation (settembre 2003)
14. Davide Fiaschi - Andre Mario Lavezzi, Nonlinear economic growth; some theory and cross-country evidence (settembre 2003)
15. Luciano Fanti , Fiscal policy and tax collection lags: stability, cycles and chaos (settembre 2003)
16. Rodolfo Signorino- Davide Fiaschi, Come scrivere un saggio scientifico:regole formali e consigli pratici (settembre 2003)
17. Luciano Fanti, The growth cycle and labour contract lenght (settembre 2003)
18. Davide Fiaschi , Fiscal Policy and Welfare in an Endogenous Growth Model with Heterogeneous Endowments (ottobre 2003)

19. Luciano Fanti, Notes on Keynesian models of recession and depression (ottobre 2003)
20. Luciano Fanti, Technological Diffusion and Cyclical Growth (ottobre 2003)
21. Luciano Fanti - Piero Manfredi, Neo-classical labour market dynamics, chaos and the Phillips Curve (ottobre 2003)
22. Luciano Fanti - Luca Spataro, Endogenous labour supply and Diamond's (1965) model: a reconsideration of the debt role (ottobre 2003)
23. Giuseppe Conti (2003), Strategie di speculazione, di sopravvivenza e frodi bancarie prima della grande crisi (novembre 2003)
24. Alga D. Foschi (2003), The maritime container transport structure in the Mediterranean and Italy (dicembre 2003)
25. Davide Fiaschi - Andrea Mario Lavezzi, On the Determinants of Growth Volatility: a Nonparametric Approach (dicembre 2003)
26. Alga D. Foschi, Industria portuale marittima e sviluppo economico negli Stati Uniti (dicembre 2003)

27 Giuseppe Conti - Alessandro Polsi, *Elites* bancarie durante il fascismo tra economia regolata ed autonomia (gennaio 2004)

28 Annetta Maria Binotti - Enrico Ghiani, Interpreting reduced form cointegrating vectors of incomplete systems. A labour market application)(febbraio 2004)