Naturally constrained reduced form and stuctural parameters estimation

Pietro Balestra, Marie-Claude Pichery

I.M.E.

EQUIPE DE RECHERCHE ASSOCIEE AU C.N.R.S.

DOCUMENT DE TRAVAIL

INSTITUT DE MATHEMATIQUES ECONOMIQUES

UNIVERSITE DE DIJON

FACULTE DE SCIENCE ECONOMIQUE ET DE GESTION

4, BOULEVARD GABRIEL — 21000 DIJON

n° 101

NATURALLY CONSTRAINED REDUCED FORM
AND
STRUCTURAL PARAMETERS ESTIMATION

Pietro BALESTRA - Marie-Claude PICHERY

Février 1988

0. INTRODUCTION

This paper is concerned with the use of information. From the very beginning, every student of econometrics is taught that efficiency in estimation requires the full use of all available information. However, in his seminal paper on "Restricted and Unrestricted Reduced Forms", Dhrymes [1973] warned us that the relationship between the amount of information used and the relative efficiency of an estimator is not necessarily a monotonic one. To illustrate his point, consider the following simple two equation model of gas demand :

$$F_t = aN_t + bH_t + u_t$$

$$G_t = cF_t + dP_t + e\, G_{t-1} + v_t$$

Where F is total residential fuel demand, N is population, H is housing starts, G is residential gas demand and P is the relative price of gas. Suppose we are interested in the gas demand equation. In the likely event that the two errors u and v are contemporaneously correlated, consistency can be achieved by the 2SLS technique. The first stage of the procedure involves the estimation of F. How should this be done ? Since the first equation is already in reduced form, the obvious answer to this question is to apply OLS directly to it. Alternatively, we might think of estimating F from the unrestricted reduced form, i.e. an equation involving all predetermined variables (including P_t and G_{t-1}). Intuitively, the first choice appears more efficient since it uses more information (the fact that the coefficients of P_t and G_t are known a priori to be zero in the total fuel equation). However, as in Dhrymes' case, intuition is not a good guide and the problem deserves a careful analysis. That is exactly what we propose to do in this paper, in a more general framework and going beyond the 2SLS technique.

A situation like the one depicted in our example, in which one equation is already in reduced form but does not contain all the predetermined variables, is not uncommon in economic applications : it appears, for instance, in the estimation of rational expectation models (see Turkington [1985] and Pesaran [1986]). To characterize such a situation, we introduce the concept of a "natural constraint", which is presented and discussed is section 1 of the present paper. Section 2 is concerned with the efficient estimation of the naturally constrained reduced form. From it, we derive efficient estimators of the structural parameters (by indirect GLS) and propose a simple test of the a priori restrictions. In section 3 we discuss estimation by 2SLS. In addition to the results obtained by Turkington and Pesaran, we develop a 2SLS-GLS estimator and assess its asymptotic properties. Several full information estimators of the 3SLS type are presented and compared in section 4 and some efficiency aspects of the instrumental variable method are studied. Finally, the major findings of this paper are summarized in section 5.

1. THE STRUCTURAL FORM

In this section, we specify our model, define the concept of a natural constraint, point out the limitations of our approach and indicate some possible extensions.

1.1. Specification and Assumptions

We consider the following structural model composed of two sets of n_1, respectively n_2 equations :

$$(1.1)\quad \begin{cases} Y_1 = X_1B_1 + E_1 \\ Y_2 = Y_1A_1 + X_2A_2 + E_2 \end{cases}$$

where Y_1 and Y_2 are two matrices of endogenous variables of order $T \times n_1$ and $T \times n_2$; X_1 and X_2 are two matrices of exogenous variables of order $T \times m_1$ and $T \times m_2$; and E_1 and E_2 are two matrices of errors.

We shall adopt the following set of assumptions :

(i) The $T \times m$ matrix $X = [X_1 X_2]$, $m = m_1 + m_2$, is non-stochastic and of full rank $m < T$. Note that this assumption implies that X_1 and X_2 have no variables in common. Furthermore, the matrix X satisfies the following limits :

$$\lim \frac{1}{T} X'X = Q$$

$$\text{plim} \frac{1}{T} X'E_i = 0 \qquad i = 1, 2$$

where Q is a positive definite matrix.

(ii) Each row of the error matrix $E = [E_1\ E_2]$ is identically and independently distributed with zero mean and variance-covariance matrix Σ, a positive definite matrix.

(iii) $n_1 \leqslant m_1$. This condition is necessary for the identification of the structural parameters.

1.2. The Natural Constraints

It is immediately evident that the first set of equations are already in reduced form. But they do not contain all the exogenous variables. It is as if the coefficients of X_2 in the first set of equations are set to zero. When in the natural specification of a system of simultaneous equations one equation is already in reduced form but does not involve all of the exogenous variables, we speak of a natural constraint. How this information should be used in the estimation of the parameters of the model is the question we address in this paper.

Note that, according to Assumption (i), $X_1 \cap X_2 = \emptyset$. Our procedure, however, can easily be extended to the case where some exogenous variables are present in both sets of equations. The results are of the same kind as those given in the sequel, but slightly more complicated.

We finally observe that in our specification the coefficient matrices B_1, A_1 and A_2 are full matrices ; no restrictions are imposed on their elements. In principle, it is easy to introduce zero restrictions on some of the coefficients : the estimation methods would be similar in spirit to those developed here, but the simplicity of the results would be lost and consequently most of the interest in the comparison of the various estimators and in their interpretation would disappear. For these reasons, this extension is not pursued here.

2. THE REDUCED FORM

The (full) unconstrained reduced form of our model is :

$$(2.1)\quad \begin{cases} Y_1 = X_1\Pi_{11} + X_2\Pi_{12} + U_1 \\ Y_2 = X_1\Pi_{21} + X_2\Pi_{22} + U_2 \end{cases}$$

The constraints imposed by the a priori restrictions on the parameters of the reduced form are of two types :

. the natural constraint : $\Pi_{12} = 0$

. the other constraints : $\Pi_{21} = \Pi_{11}A_1$

When the natural constraint is explicitly accounted for in the reduced form, we speak of the "naturally constrained reduced form".

In this section we derive an efficient estimator of the naturally constrained reduced form, propose an indirect asymptotically efficient estimator of all structural parameters via the naturally constrained reduced form and develop a test of the a priori restrictions.

2.1. Efficient Estimation of the Naturally Constrained Reduced Form

The naturally constrained reduced form, i.e.

$$(2.2)\quad \begin{cases} Y_1 = X_1\Pi_{11} + U_1 \\ Y_2 = X_1\Pi_{21} + X_2\Pi_{22} + U_2 = X\Pi_2 + U_2 \end{cases}$$

is a Zellner-type model with the same regressors in the second set of equations and with only a subset of the regressors in the first set of equations. The GLS principle applied to this model leads to the following estimators :

$$(2.3)\quad \begin{cases} \hat{\Pi}_{11} = (X_1'X_1)^{-1}X_1'Y_1 \\ \hat{\Pi}_2 = (X'X)^{-1}X'Y_2 - (X'X)^{-1}X'(Y_1 - X_1\hat{\Pi}_{11})\Omega_{11}^{-1}\Omega_{12} \end{cases}$$

where Ω_{11} and Ω_{12} are the appropriate parts of the matrix Ω, the variance-covariance matrix associated with the reduced form residuals. Formulae (2.3) reveal that for the first set of equations, the (naturally constrained) OLS estimator is obtained, while for the second set of equations, the (full) OLS estimator is adjusted by a factor which depends on the estimated errors of the first set, $Y_1 - X_1\hat{\Pi}_{11}$, and the true covariance, Ω_{12}, of the errors in the two sets.

To derive an explicit characterisation of the asymptotic distribution of the naturally constrained reduced form estimator (and also for future reference), we need some more notation. Calling L_1 the selection matrix which extracts X_1 from X, i.e. $X_1 = XL_1$, the vector representation of (2.2) is :

$$(2.4)\quad \begin{cases} \text{vec } Y_1 = (I_{n_1} \otimes X)\,(I_{n_1} \otimes L_1)\text{ vec } \Pi_{11} + \text{vec } U_1 \\ \text{vec } Y_2 = (I_{n_2} \otimes X)\text{ vec } \Pi_2 + \text{vec } U_2 \end{cases}$$

or,in compact form :

$$(2.5)\quad \text{vec } Y = (I_n \otimes X)\,L\,\gamma + \text{vec } U$$

where $\gamma' = [(\text{vec } \Pi_{11})', (\text{vec } \Pi_2)']$ is the vector of all non-zero reduced form parameters and L is the selection matrix

$$L = \begin{bmatrix} I_{n_1} \otimes L_1 & 0 \\ 0 & I_{n_2 m} \end{bmatrix}$$

With this notation, the estimator given in (2.3) takes the following form :

$$(2.6) \qquad \hat{\gamma} = [L'(\Omega^{-1} \otimes X'X)L]^{-1}L'(\Omega^{-1} \otimes X') \text{ vec } Y$$

and, given our assumptions, it can immediately be established that

$$(2.7) \qquad \sqrt{T}(\hat{\gamma} - \gamma) \to N(0, [L'(\Omega^{-1} \otimes Q)L]^{-1}).$$

We are now in position to draw a certain number of conclusions :

(i) The naturally constrained reduced form estimator is more efficient than the unconstrained estimator. Here, thus, the use of (partial) information leads to a gain in efficiency. This can be seen in the following way. Call γ_F the full set of unconstrained reduced form parameters, $\gamma_F' = (\text{vec } \Pi)'$, and $\hat{\gamma}_F$ the OLS estimator . We know that the limiting distribution of $\sqrt{T}(\hat{\gamma}_F - \gamma_F)$ is normal with zero mean and covariance matrix $\Omega \otimes Q^{-1}$. Noting that $\gamma = L'\gamma_F$, the part of the covariance matrix corresponding to the unconstrained estimator of γ is $L'(\Omega \otimes Q^{-1})L$, which is to be compared to the corresponding expression in (2.7). From Lemma 1 in Appendix 1, we immediately have

$$[L'(\Omega^{-1} \otimes Q)L]^{-1} = L'(\Omega \otimes Q^{-1})L - N$$

where N is a non-negative definite matrix. This proves the superiority of the constrained estimator.

(ii) The asymptotic distribution of the full maximum likelihood estimator of the naturally constrained reduced form (under normality) is the same as the one given in (2.7).

(iii) The expressions (2.3) and (2.6) do not define a proper estimator, since Ω is not usually known. It is clear, however, that the limiting distribution is the same when Ω is replaced by a consistent estimate. One such estimate can be obtained from the estimated residuals of the naturally constrained reduced form by OLS.

(iv) Iteration of the procedure outlined in (iii) gives in the limit the same numerical solution as the FIML of the naturally constrained reduced form (See Oberhofer and Kmenta [1974]).

2.2. Indirect Estimation of the Structural Parameters

Since identification of both B_1 and A_2 is immediate from the reduced form ($B_1 = \Pi_{11}$ and $A_2 = \Pi_{22}$), our only concern here is to provide an estimate of A_1. To do this, we apply the Asymptotic Least-Squares method (see Gourieroux, Monfort and Trognon [1985]).

As mentioned earlier, the true relationship involving A_1 and the reduced form parameters is $\Pi_{21} = \Pi_{11}A_1$. Its empirical counterpart is :

$$(2.8) \quad \hat{\Pi}_{21} = \hat{\Pi}_{11}A_1 + V$$

where $\hat{\Pi}_{21}$ and $\hat{\Pi}_{11}$ are the efficient estimator derived in the preceding paragraph and V is a matrix of errors to be shortly specified. (The relationship cannot hold exactly except in the just-identified case).

The OLS estimator of A_1, $\hat{A}_1(\mathrm{OLS}) = (\hat{\Pi}_{11}'\hat{\Pi}_{11})^{-1}\hat{\Pi}_{11}'\hat{\Pi}_{21}$, is consistent, but not in general efficient. Efficiency can be achieved by applying the GLS principle. This requires knowledge of the statistical properties of V.

Writing $\Pi^* = [\Pi_{11}\ \Pi_{21}]$ and $D' = [-A_1'\ I]$, the true relationship becomes $\Pi^*D = 0$ and, from (2.8), the matrix V can be expressed as

$$(2.9) \quad V = \hat{\Pi}^*D = (\hat{\Pi}^* - \Pi^*)D$$

or, in vector form :

$$\text{vec } V = (D' \otimes I) \text{ vec}(\hat{\Pi}^* - \Pi^*)$$

$$= (D' \otimes I) F'(\hat{\gamma} - \gamma)$$

where F' is the following selection matrix :

$$F' = \begin{bmatrix} I_{n_1 m_1} & 0 \\ 0 & I_{n_2} \otimes L_1' \end{bmatrix}$$

Consequently, the asymptotic distribution of vec V can be derived directly from that of $\hat{\gamma} - \gamma$, namely :

$$(2.10) \quad \sqrt{T} \text{ vec } V \rightarrow N(0, G)$$

with G given by :

$$(2.11) \quad G = (D' \otimes I_{m_1}) F' [L'(\Omega^{-1} \otimes Q)L]^{-1} F (D \otimes I_{m_1}).$$

Using again the results of Appendix 1, the matrix G can be shown to be equal to :

$$(2.12) \quad G = D'\Omega D \otimes Q_{11}^{-1} + (\Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}) \otimes Q_{11}^{-1}Q_{12}Q_o^{-1}Q_{21}Q_{11}^{-1}$$

where $Q_o = \text{plim} \frac{1}{T} X_2' M_1 X_2$, $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.

We are now in position to apply the GLS principle to Eq. (2.8). The resulting estimator is

$$(2.13) \quad \text{vec } \hat{A}_1(\text{GLS}) = [(I \otimes \hat{\Pi}_{11}')G^{-1}(I \otimes \hat{\Pi}_{11})]^{-1}(I \otimes \hat{\Pi}_{11}')G^{-1}\text{vec}\hat{\Pi}_{21}$$

Furthermore, we can prove that

$$(2.14)\quad \sqrt{T}\,\mathrm{vec}[\hat{A}_1(\mathrm{GLS}) - A_1] \rightarrow N(0,[(I \otimes \Pi_{11}')G^{-1}(I \otimes \Pi_{11})]^{-1})$$

The expression given in (2.13) involves the unknown matrix G. A feasible estimator is obtained by replacing G with a consistent estimator and its limiting distribution is again the same one as that shown in (2.14). The unknown components of G are Ω (which can be consistently estimated from the OLS reduced form residuals), Q (which can be replaced by the sample analog, $\frac{1}{T}X'X$) and A_1 (which can be obtained by applying OLS to (2.8)).

What are now the efficiency properties of the proposed estimator ? It is remarkable that the asymptotic covariance matrix of $\hat{A}_1$(GLS) is the same as that of the FIML estimator of the structural form. This important result (proved directly in Appendix 2) is a special case of a more general theorem established by Gourieroux and Monfort [1985, Property 3, Corollary 4].

2.3. Testing the A Priori Restrictions

We extend a procedure originally suggested by Gourieroux, Monfort and Renault [1985] in order to construct a test of the a priori restriction $\Pi_{21} = \Pi_{11}A$.The extension is twofold : first, in our case, the restrictions are expressed in terms of matrices of coefficients and not of vectors of coefficients ; second, we propose here a new characterisation of the null hypothesis which is conceptually simpler and does not require the use of generalized inverses.

For the null hypothesis to be true, the case must be that the columns of Π_{21} lie in the space spanned by the columns of Π_{11}. We can express this fact in the following way :

$$(2.15)\quad P'\Pi_{21} = 0$$

where P' is any $(m_1 - n_1) \times m_1$ matrix of full rank such that $P'\Pi_{11} = 0$. The hypothesis to be tested, therefore, is $\Gamma = 0$, with $\Gamma = P'\Pi_{21}$. Define the following estimator of Γ :

$$(2.16) \quad \hat{\Gamma} = \hat{P}'\hat{\Pi}_{21}$$

where $\hat{\Pi}_{21}$ is the efficient naturally constrained reduced form estimator and $\hat{P}'$ is a consistent estimator of P such that $\hat{P}'\hat{\Pi}_{11} = 0$ (this matrix, as we shall see shortly, need not be computed). Developing, we get successively :

$$\begin{aligned}
\hat{\Gamma} &= \hat{P}'\hat{\Pi}_{21} \\
&= \hat{P}'(\hat{\Pi}_{21} - \hat{\Pi}_{11}A_1) && \text{(since } \hat{P}'\hat{\Pi}_{11} = 0\text{)} \\
&= \hat{P}'\hat{\Pi}^*D \\
&= \hat{P}'(\hat{\Pi}^* - \Pi^*)D && \text{(since } \Pi^*D = 0\text{)} \\
&= \hat{P}'V
\end{aligned}$$

where V is the matrix of errors defined in (2.9).

The asymptotic distribution of $\hat{\Gamma}$ may now be obtained directly from that of V, namely :

$$(2.17) \quad \sqrt{T}\ \text{vec}\ \hat{\Gamma} \rightarrow N(0, (I \otimes P')\ G\ (I \otimes P))$$

Consequently, the quantity

$$(2.18) \quad q = T(\text{vec}\ \hat{\Gamma})'[(I \otimes P')\ G\ (I \otimes P)]^{-1}\text{vec}\ \hat{\Gamma}$$

is asymptotically distributed as a Chi-square variate with $(m_1 - n_1)n_2$ degrees of freedom. It is this quantity (with $\hat{P}$ and $\hat{G}$ replacing P and G) that is used for the test.

What does this quantity represent ? Let us rewrite it in the following way :

$$\frac{1}{T} q = (\text{vec } \hat{\Pi}_{21})'(I \otimes \hat{P}) \, [(I \otimes \hat{P}')G(I \otimes \hat{P})]^{-1}(I \otimes \hat{P}') \text{ vec } \hat{\Pi}_{21}$$

$$= (\text{vec } \hat{\Pi}_{21})'\{ G^{-1} - G^{-1}(I \otimes \hat{\Pi}_{11})[(I \otimes \hat{\Pi}_{11})'G^{-1}(I \otimes \hat{\Pi}_{11})]^{-1}(I \otimes \hat{\Pi}_{11})'G^{-1}\} \text{ vec } \hat{\Pi}_{21}$$

where, again, we have used Lemma 1 of Appendix 1 for the inverse of $(I \otimes \hat{P}')G(I \otimes \hat{P})$. The last form of the above expression reveals immediately that the test statistics q is T times the generalized sum of squared residuals of the regression model defined in (2.8). In conclusion, the indirect method provides an asymptotically efficient estimator of the structural parameters and, at the same time, gives an easily computed statistics for the test of the a priori restrictions.

3. 2SLS ESTIMATORS

We discuss in this section the estimation of the structural parameters by the 2SLS method (a limited information approach). Given that the first set of equations is already in reduced form, the only equations of interest in this connection are the ones in the second set, namely :

$$(3.1) \qquad Y_2 = Y_1A_1 + X_2A_2 + E_2$$

$$= HA + E_2$$

where $H = [Y_1 \; X_2]$ and $A' = [A_1' \; A_2']$. Direct estimation of (3.1) by OLS (or by GLS, which, in this context, produce the same results) is inconsistent, unless $\Sigma_{12} = 0$. When $\Sigma_{12} \neq 0$, in the spirit of the 2SLS method, consistency is achieved by replacing Y_1 by an appropriate estimate $\hat{Y}_1$. The corresponding estimating equation is :

$$(3.2) \qquad Y_2 = \hat{H}A + \tilde{E}_2$$

where $\hat{H} = [\hat{Y}_1 \; X_2]$ and

$$(3.3) \qquad \tilde{E}_2 = E_2 + (Y_1 - \hat{Y}_1)A_1.$$

Equation (3.2) raises two different types of problems. The first one concerns the choice of $\hat{Y}_1$. According to whether the information contained in the natural constraints is used or not, alternative forms of the 2SLS estimator are obtained. Their statistical properties should be assessed and compared. The second one focuses on the choice of the estimation method : should one use OLS or GLS ? These questions are studied in the following two paragraphs.

3.1. Choices for $\hat{Y}_1$

Two obvious choices for $\hat{Y}_1$ (Turkington [1985] and Pesaran [1986]) are :

(i) $\hat{Y}_1 = X_1(X_1'X_1)^{-1}X_1'Y_1$

(ii) $\hat{Y}_1 = X(X'X)^{-1}X'Y_1$

corresponding to, respectively, the naturally constrained and unconstrained estimation of the reduced form. The first choice has an "error in variable" interpretation, while the second represents the usual instrumental variables method. In both cases, when OLS is applied to (3.2), the resulting estimators -called, respectively, 2SLS-OLS1 and 2SLS-OLS2- are consistent. Their limiting distribution can be shown to be normal with zero mean and variance-covariance matrix :

(3.4) $\quad$ 2SLS-OLS1 : $\Sigma_{22} \otimes (S'QS)^{-1} + \Sigma_* \otimes (S'QS)^{-1}S'\tilde{Q}S(S'QS)^{-1}$

(3.5) $\quad$ 2SLS-OLS2 : $\Sigma_{22} \otimes (S'QS)^{-1}$

where :

$$\Sigma_* = A_1'\Sigma_{11}A_1 + A_1'\Sigma_{12} + \Sigma_{21}A_1$$

$$S = \begin{bmatrix} \Pi_{11} & 0 \\ 0 & I_{m_2} \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} 0 & 0 \\ 0 & Q_o \end{bmatrix}, \; Q_o = \operatorname{plim} \frac{1}{T} X_2'M_1X_2 .$$

Since Σ_* is not, in general, a definite matrix, the relative efficiency of the two estimators cannot be assessed unambiguously ; it depends (as Turkington and Pesaran have shown) on the parameter values. Obviously, when $\Sigma_{12} = 0$, the matrix Σ_* is non-negative definite, and in that case, 2SLS-OLS2 is more efficient than 2SLS-OLS1. But, then, when $\Sigma_{12} = 0$, direct OLS estimation of the structural equation is more efficient than both.

It appears thus that there is no clear answer to the question whether one should use or not the natural constraint in 2SLS estimation.

3.2. A Generalized 2SLS Estimator

As mentioned earlier, the selection of an estimating technique for Eq. (3.2) should be guided by the statistical properties of the matrix of errors, $\tilde{E}_2$. Writing $M_Z = I - Z(Z'Z)^{-1}Z'$, for Z either equal to X or equal to X_1, (3.3) can be put in the following form :

$$(3.6) \quad \tilde{E}_2 = E_2 + M_Z E_1 A_1$$

Its vector representation is :

$$(3.7) \quad \text{vec } \tilde{E}_2 = [A_1' \otimes M_Z]\text{vec } E_1 + \text{vec } E_2$$

$$= [A_1' \otimes M_Z \; , \; I] \text{ vec } E$$

from which it can easily be established that vec $\tilde{E}_2$ has zero expectation and variance-covariance matrix equal to :

$$(3.8) \quad V(\text{vec } \tilde{E}_2) = \Sigma_* \otimes M_Z + \Sigma_{22} \otimes I$$

$$= \Sigma_{22} \otimes (I - M_Z) + \Sigma_{**} \otimes M_Z$$

$$\stackrel{d}{=} K$$

where $\Sigma_{**} = \Sigma_* + \Sigma_{22}$, a positive definite matrix.

The above variance-covariance structure has some important implications :

(i) The estimating set of equations (3.2) is indeed a Zellner type model with the same regressors in each equation and non-zero contemporaneous covariances, but the present covariance structure is a sum of two Kronecker products and not simply a single Kronecker product.

(ii) The particular form of K is similar to the one which appears in error components models. Its inverse (see Magnus [1982]) is readily given by :

$$K^{-1} = \Sigma_{22}^{-1} \otimes (I - M_Z) + \Sigma_{**}^{-1} \otimes M_Z$$

(iii) It is not true, in general, that OLS and GLS are equal, although the regressors are the same in each equation. For OLS to be equal to GLS (see MILLIKEN and ALBOHALI [1984]) the case must be that :

$$(I \otimes M_{\hat{H}})K^{-1}(I \otimes \hat{H}) = 0$$

with $M_{\hat{H}} = I - \hat{H}(\hat{H}'\hat{H})^{-1}\hat{H}'$. For the above expression to hold, it is necessary and sufficient that

(3.9) $$M_{\hat{H}} M_Z \hat{H} = 0$$

We can prove that condition (3.9) is satisfied in two cases.

- First case : $Z = X$ (the usual 2SLS method).
  When $Z = X$, $\hat{H} = (I - M_Z)H$ and therefore $M_Z\hat{H} = 0$

- Second case : $Z = X_1$, with $m_1 = n_1$ (just identification).

In this case, we can write

$$\hat{H} = [X_1(X_1'X_1)^{-1}X_1'Y_1 \quad X_2] = [X_1 \quad X_2] \begin{bmatrix} (X_1'X_1)^{-1}X_1'Y_1 & 0 \\ 0 & I \end{bmatrix} = X\mathbf{B}$$

where $\mathbf{B}$ is non-singular. It then follows that

$M_{\hat{H}} = M_X$. Furthermore $M_X M_Z = M_X M_{X_1} = M_X$. Finally :

$M_{\hat{H}} M_Z \hat{H} = M_X M_Z X\mathbf{B} = M_X X\mathbf{B} = 0.$

However, in the case of overidentification, the use of the natural constraint ($Z = X_1$) does not in general guarantee the equality of OLS and GLS (condition (3.9) is not satisfied). For this case, a gain in efficiency can be achieved by using the GLS principle. The corresponding estimator, which we shall call 2SLS-GLS, is

$$\text{(3.10)} \quad \text{vec } \hat{A}_{2SLS-GLS} = [(I \otimes \hat{H})'K^{-1}(I \otimes \hat{H})]^{-1}(I \otimes \hat{H})'K^{-1}\text{vec } Y_2$$

and its limiting distribution is normal with zero mean and variance-covariance matrix given by :

$$\text{(3.11)} \quad [\Sigma_{22}^{-1} \otimes S'QS + (\Sigma_{**}^{-1} - \Sigma_{22}^{-1}) \otimes \tilde{Q}]^{-1}$$

The 2SLS-GLS estimator has the following properties :

- it is (asymptotically) more efficient than the 2SLS-OLS1 estimator (unless $n_1 = m_1$, in which case they are equal) ;
- its relation with the 2SLS-OLS2 estimator depends on the parameter values ;
- it is the only 2SLS estimator that can be totally efficient for one set of parameter values, namely $A_1 = - \Sigma_{11}^{-1}\Sigma_{12}$.

The last of the above properties can be established as follows. For the corresponding parameters, the variance-covariance matrix of the limiting distribution of the FIML estimator is :

$$(3.12) \qquad [\Sigma_{22}^{-1} \otimes S'QS + (\Sigma^{22} - \Sigma_{22}^{-1}) \otimes \tilde{Q}]^{-1}$$

where Σ^{22} is the appropriate block of Σ^{-1}. The comparison of (3.12) with (3.11) reveals that the FIML estimator is more efficient than the 2SLS-GLS estimator whenever $\Sigma^{22} - \Sigma_{**}^{-1}$ is non negative definite (or, equivalently $\Sigma_{**} - (\Sigma^{22})^{-1}$ is non negative definite). Now :

$$\begin{aligned}
\Sigma_{**} - (\Sigma^{22})^{-1} &= \Sigma_{**} - \Sigma_{22} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \\
&= A_1'\Sigma_{11}A_1 + A_1'\Sigma_{12} + \Sigma_{21}A_1 + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \\
&= (A_1 + \Sigma_{11}^{-1}\Sigma_{12})' \, \Sigma_{11}(A_1 + \Sigma_{11}^{-1}\Sigma_{12})
\end{aligned}$$

which is clearly non negative definite. Furthermore, the equality holds iff $A_1 = - \Sigma_{11}^{-1}\Sigma_{12}$.

The feasible counterpart of the 2SLS-GLS estimator involves the estimation of the unknown matrices Σ_{22} and Σ_{**}. Noting that $\Sigma_{**} = \Omega_{22}$, this matrix can be consistently estimated from the reduced form residuals. As for Σ_{22}, it can be estimated consistently from the 2SLS-OLS1 residuals.

4. 3SLS ESTIMATORS

The simplicity of our model provides a perfect ground for discussing some of the issues that arise in connection with the use of the instrumental variable (IV) method. We present, at first, three alternative 3SLS estimators and compare their asymptotic efficiency ; then, we examine in a systematic way the conditions under which the IV method produces asymptotically efficient results.

4.1. Three Alternative 3SLS Estimators

The presence of natural constraints suggests the construction of various types of 3SLS estimators in addition to the usual one. Two obvious candidates are proposed below, after having given, for ready-made reference, the usual estimator.

(i) The usual estimator : 3SLS-1

Each equation of both sets of equations of the structural form is pre-multiplied by the full matrix of instruments, X', and then estimation is performed by GLS. Since there are no restrictions on the variance-covariance matrix, the resulting estimator is asymptotically efficient (its limiting distribution being equal to that of the FIML estimator).

(ii) A mixed estimator : 3SLS-2

Since the equations of the first set are already in reduced form, we may leave them as they are and use the full set of instruments only for the equations of the second set. We then apply GLS to the whole system. As shown in Appendix 3, the 3SLS-2 estimator is numerically identical to the 3SLS-1 estimator (and hence fully efficient).

(iii) A constrained IV estimator : 3SLS-3

The first set of equations (already in reduced form) involve only a sub-set of the exogenous variables, namely X_1. The matrix X_1, being exogenous, is its own instrument. Therefore, we may think of pre-multiplying each equation of the first set by X_1' and each equation of the second set by X', before applying the GLS principle. Surprisingly, this estimator is consistent but not efficient. It can be shown to be numerically identical, for the parameters of the second set of equations, to the 2SLS-OLS2 estimator (see again Appendix 3).

The reason why the 3SLS-3 estimator is not fully efficient deserves some consideration.

4.2. Efficient IV estimation

When the variance-covariance structure of the residuals is not spherical, several options are available for the definition of IV estimators. Bowden and Turkington [1984] distinguish two such estimators, which they call, respectively, OLS-ANALOG and GLS-ANALOG. In order to present them, we need some notation.

Our structural system may be written in the following vector form :

$$(4.1)\qquad \begin{cases} y_1 = W_1\alpha_1 + \varepsilon_1 \\ y_2 = W_2\alpha_2 + \varepsilon_2 \end{cases} \Leftrightarrow \quad y = W\alpha + \varepsilon$$

where :

$$y_i = \text{vec } Y_i \;, \quad \varepsilon_i = \text{vec } E_i \qquad i = 1, 2$$

$$\alpha_1 = \text{vec } B_1 \;, \quad \alpha_2 = \text{vec } A$$

$$W_1 = I \otimes X_1, \quad W_2 = I \otimes H$$

$$y' = [y_1' \; y_2'] \;, \quad \alpha' = [\alpha_1' \; \alpha_2'] \qquad \varepsilon' = [\varepsilon_1' \; \varepsilon_2']$$

$$W = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}$$

and $\quad \mathrm{Var}(\varepsilon) = \Sigma \otimes I \overset{d}{=} \mathbf{V}$.

Let Z be a matrix of instruments. Then, the two IV estimators are :

(4.2) OLS-ANALOG $\quad \hat{\alpha}_{OLS} = [W'Z(Z'\mathbf{V}Z)^{-1}Z'W]^{-1}W'Z(Z'\mathbf{V}Z)^{-1}Z'y$

(4.3) GLS-ANALOG $\quad \hat{\alpha}_{GLS} = [W'\mathbf{V}^{-1}Z(Z'\mathbf{V}^{-1}Z)^{-1}Z'\mathbf{V}^{-1}W]^{-1}W'\mathbf{V}^{-1}Z(Z'\mathbf{V}^{-1}Z)^{-1}Z'\mathbf{V}^{-1}y$

Although the superiority of the GLS-ANALOG has been established for some particular models (as in the case of 2SLS estimation with non spherical disturbances), a general comparison in terms of efficiency of the two estimators is not possible, as indicated by Bowden and Turkington. They also give examples in which the OLS-ANALOG is consistent, while the GLS-ANALOG is not (1). However, a general result concerning the two estimators has been established by Balestra [1988]. He shows that the OLS-ANALOG and the GLS-ANALOG are equal if and only if the following conditions hold :

(4.4) $\quad Z'\mathbf{V}M_Z = 0$

where $\quad M_Z = I - Z(Z'Z)^{-1}Z'$.

---

(1) Our 3SLS-2 estimator can also abusively be put in the form of a IV estimator with

$$Z = \begin{bmatrix} I \otimes I & 0 \\ 0 & I \otimes X \end{bmatrix}$$

The above matrix is not a proper instrumental variable matrix since the number of columns of Z increases when the number of observations grows. If we use the matrix Z so defined, the OLS-ANALOG is consistent but the GLS-ANALOG is not.

Now, in the case of our usual 3SLS estimator, $Z = I \otimes X$ and $M_Z = I \otimes M_X$. Therefore

$$Z'\mathbf{V}M_Z = (I \otimes X')(\Sigma \otimes I)(I \otimes M_X) = 0$$

and the two estimators are equal (a result that has been often pointed out). But in the case of our constrained estimator (3SLS-3), Z and M_Z take the following form :

$$Z = \begin{bmatrix} I \otimes X_1 & 0 \\ 0 & I \otimes X \end{bmatrix} ; \quad M_Z = \begin{bmatrix} I \otimes M_1 & 0 \\ 0 & I \otimes M_X \end{bmatrix}$$

and consequently

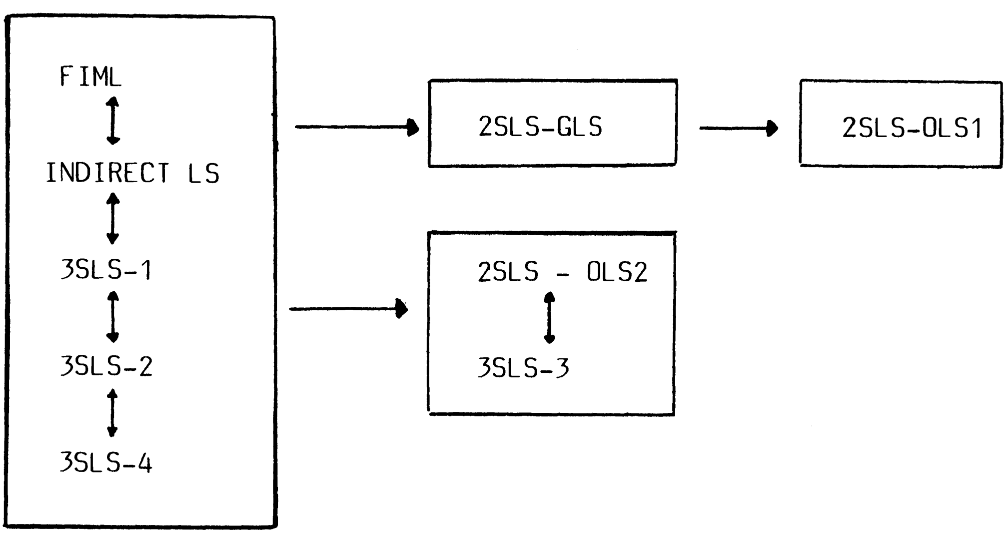
$$Z'\mathbf{V}M_Z = \begin{bmatrix} 0 & 0 \\ \Sigma_{21} \otimes X'M_1 & 0 \end{bmatrix} \neq 0$$

The two estimators are not equal. What happens now if, instead of the OLS-ANALOG, we use the GLS-ANALOG ? The answer is that the resulting estimator (see Appendix 3), which we shall call 3SLS-4, is identical to the usual 3SLS estimator and therefore fully efficient.

5. CONCLUSIONS

In this paper we have analyzed in a systematic way the role played by the "natural constraints" in the estimation of a system of simultaneous equations. Our findings are summarized in Fig. 1, where the dominance relationship (in terms of efficiency) among the different estimators is portrayed.

Fig. 1 : Dominance Relationship Between the Different Estimators of the Structural Parameters



Although our model is quite simple, we believe that it might have some important applications in economics ; furthermore, it sheds some light on the problem of the proper use of information in econometric analysis.

Appendix 1 : THE INVERSE OF A'QA

Let Q be a positive definite matrix of order n and let A' be an $m \times n$ matrix of rank r. Define the $m \times m$ matrix $B = A'QA$, obviously non-negative definite. In many statistical applications (as in the comparison of variance-covariance matrices of different estimators) it is useful to evaluate the inverse of B (the proper inverse or the generalized inverse) in terms of Q^{-1}.

The following Lemmas address this question.

Lemma 1 The generalized inverse of B is given by :

$$B^{+} = A^{+}(Q^{-1} - Q^{-1}S(S'Q^{-1}S)^{-1}S'Q^{-1})(A^{+})'$$

where S' is an $(n - r) \times n$ matrix of full rank such that $S'A = 0$.

Corollary 1

(i) If $r = m$, B is positive definite and therefore $B^{+} = B^{-1}$. In this case, $A^{+} = (A'A)^{-1}A'$.

(ii) If $r = n$, S is empty and therefore $B^{+} = A^{+}Q^{-1}(A^{+})'$.

Proof of Lemma 1

The homogenous system $A'h = 0$ has exactly $n - r$ linearly independent solutions. Let S be the $n \times (n - r)$ matrix whose columns are $n - r$ linearly independent solutions, $A'S = 0$. The matrix AA^{+} is idempotent and symmetric of order n and rank r. Therefore $AA^{+}S = (A^{+})'A'S = 0$ and we have :

$$AA^{+} + S(S'S)^{-1}S' = I$$

Let us compute BB^+ and B^+B :

$$BB^+ = A'QAA^+Q^{-1}(A^+)' - A'QAA^+Q^{-1}S(S'Q^{-1}S)^{-1}S'Q^{-1}(A^+)'$$

$$= A'QAA^+Q^{-1}(A^+)' - A'Q(I - S(S'S)^{-1}S')Q^{-1}S(S'Q^{-1}S)^{-1}S'Q^{-1}(A^+)'$$

$$= A'QAA^+Q^{-1}(A^+)' + A'QS(S'S)^{-1}S'Q^{-1}(A^+)'$$

$$= A'Q[AA^+ + S(S'S)^{-1}S']Q^{-1}(A^+)'$$

$$= A'(A^+)'$$

$$= A^+A \quad \text{symmetric and idempotent}$$

$$B^+B = A^+A \quad \text{(in a similar fashion)}$$

Therefore :

$$B^+BB^+ = A^+AB^+ = B^+$$

$$BB^+B = BA^+A = BA'(A^+)' = B$$

and B^+ is indeed the Penrose-generalized inverse of B.

Proof of Corollary 1

(i) We note that when $m = r$, $A^+A = I$ and therefore $B^+B = I$ which implies $B^+ = B^{-1}$.

(ii) Obvious.
Let us just note that if $m > n$, we have
$B^+ = A'(AA')^{-1}Q^{-1}(AA')^{-1}A$.

It is useful to find the conditions under which the generalized inverse of B is given by

$$B^{+} = A^{+} Q^{-1}(A^{+})'$$

These conditions are given in Lemma 2.

Lemma 2

The generalized inverse of B is equal to $B^{+} = A^{+}Q^{-1}(A^{+})'$ if and only if the two matrices AA^{+} and Q commute.

Corollary 2

When $r = n$, $AA^{+} = I$ and the condition of lemma 2 is always satisfied.

Proof of Lemma 2

From the definition of B^{+} and the fact that $S'Q^{-1}S$ is positive definite we conclude that $B^{+} = A^{+}Q^{-1}(A^{+})'$ if and only if

$$A^{+}Q^{-1}S = 0$$

We therefore have to prove that the condition stated in Lemma 2 is equivalent to $A^{+}Q^{-1}S = 0$.

a) Sufficiency. Let us suppose that AA^{+} and Q commute :

$$AA^{+}Q = QAA^{+}$$

$$Q^{-1}AA^{+} = AA^{+}Q^{-1}$$

$$A^{+}Q^{-1}AA^{+} = A^{+}Q^{-1}$$

Post-multiplication by S gives $A^{+}Q^{-1}S = 0$

b) Necessity. let us suppose that $A^+Q^{-1}S = 0$:

$$A^+Q^{-1}S = 0$$

$$A^+Q^{-1}S(S'S)^{-1}S' = 0$$

$$AA^+Q^{-1}(I - AA^+) = 0$$

$$AA^+Q^{-1} = AA^+Q^{-1}AA^+$$

$$AA^+Q^{-1} = AA^+Q^{-1}AA^+ + S(S'S)^{-1}S'Q^{-1}(A^+)'A'$$

$$AA^+Q^{-1} = Q^{-1}AA^+$$

$$QAA^+ = AA^+Q$$

and the two matrices commute.

Appendix 2 : COVARIANCE MATRIX FOR THE ESTIMATORS OF THE FIML

For the constrained model (with the symmetry constraint), the likelihood function has the following form :

$$\mathcal{L}(Y/B_1,\ A_1,\ A_2,\ \Sigma) = \text{cte} - \frac{T}{2}\ \ln\ |\Sigma| - \frac{1}{2}\ \text{tr}\ \{\Sigma^{-1}E'E\} + \lambda'C'\ \text{Vec}\ \Sigma$$

where λ is the $(\frac{n(n-1)}{2} \times 1)$ vector of the Lagrange multiplier and C the $(n^2 \times \frac{n/n-1)}{2})$ selection matrix verifying

(A2.1) $C'\ C = I$

$$C\ C' = \frac{1}{2}\ (I - P)$$

with P the $(n^2 \times n^2)$ permutation matrix.

The bordered information matrix is

(A2.2)

$$\left[\begin{array}{ccc:cc}
\Sigma^{11} \otimes Q_{11} & \Sigma^{12} \otimes Q_{11}B_1 & \Sigma^{12} \otimes Q_{12} & 0 & 0 \\
\Sigma^{21} \otimes B_1'Q_{11} & \Sigma^{22} \otimes (B_1'Q_{11}B_1 + \Sigma_{11}) & \Sigma^{22} \otimes B_1'Q_{12} & L_2'\Sigma^{-1} \otimes L_1' & 0 \\
\Sigma^{21} \otimes Q_{21} & \Sigma^{22} \otimes Q_{21}B_1 & \Sigma^{22} \otimes Q_{22} & 0 & 0 \\
\hdashline
0 & \Sigma^{-1}L_2 \otimes L_1 & 0 & \frac{1}{2}(\Sigma^{-1} \otimes \Sigma^{-1}) & C \\
0 & 0 & 0 & C' & 0
\end{array}\right]$$

where $L_1^{'} = [I_{n1} \quad 0]$ and $L_2^{'} = [0 \quad I_{n2}]$ are two selection matrices, and Σ^{ij} are the appropriate block of the inverse matrix Σ^{-1}.

From the inverse of the bordered matrix, the covariance matrix of the parameters are obtained as :

$$\text{(A2.3)} \quad \mathrm{Var}\begin{pmatrix} \mathrm{Vec}\ \hat{B}_1 \\ \mathrm{Vec}\ \hat{A}_1 \\ \mathrm{Vec}\ \hat{A}_2 \end{pmatrix} = \begin{bmatrix} \Sigma^{11} \otimes Q_{11} & \Sigma^{12} \otimes Q_{11}B_1 & \Sigma^{12} \otimes Q_{12} \\ \Sigma^{21} \otimes B_1^{'}Q_{11} & \Sigma^{22} \otimes B_1^{'}Q_{11}B_1 & \Sigma^{22} \otimes B_1^{'}Q_{12} \\ \Sigma^{21} \otimes Q_{21} & \Sigma^{22} \otimes Q_{21}B_1 & \Sigma^{22} \otimes Q_{22} \end{bmatrix}$$

A new inversion in partitioned form gives :

$$\text{(A2.4)} \quad \mathrm{Var}(\mathrm{Vec}\ \hat{B}_1) = [\Sigma^{11} \otimes Q^{11} + \Sigma_{11}^{-1} \otimes Q_{12}Q_{22}^{-1}Q_{21} + $$
$$+ (\Sigma_{11}^{-1} - \Sigma^{11}) \otimes Q^{11}B_1[B_1^{'}(Q^{11})^{-1}B_1]^{-1}B_1^{'}Q_{11}]^{-1}$$

with $Q^{11} = [Q_{11} - Q_{12}\ Q_{22}^{-1}\ Q_{21}]^{-1}$

$$\text{(A2.5)} \quad \mathrm{Var}(\mathrm{Vec}\ \hat{A}_1) = [\Sigma_{22}^{-1} \otimes B_1^{'}Q_{11}B_1 - (\Sigma_{22}^{-1} \otimes B_1^{'}Q_{12})K^{-1}(\Sigma_{22}^{-1} \otimes Q_{21}B_1)]^{-1}$$

with $K = \Sigma^{22} \otimes Q_0 + \Sigma_{22}^{-1} \otimes Q_{21}Q_{11}^{-1}Q_{12}$

$$\text{(A2.6)} \quad \mathrm{Var}(\mathrm{Vec}\ \hat{A}_2) = [\Sigma^{22} \otimes Q_0 + \Sigma_{22}^{-1} \otimes Q_{21}[Q_{11}^{-1} - B_1(B_1^{'}Q_{11}B_1)^{-1}B_1^{'}]Q_{12}]^{-1}$$

We can now compare the results for different estimators, in particular the indirect asymptotic one.

The expression (A2.5) has to be compared to the covariance matrix given by (2.14) with $\Pi_{11} = B_1$. Noting that $D'\Omega D = \Sigma_{22}$ and $\Omega^{22} = \Sigma^{22}$, G takes the following form :

(A2.7) $$G = \Sigma_{22} \otimes Q_{11}^{-1} + (\Sigma^{22})^{-1} \otimes Q_{11}^{-1} Q_{12} Q_0^{-1} Q_{21} Q_{11}^{-1}$$

Its inverse is :

(A2.8) $$G^{-1} = \Sigma_{22}^{-1} \otimes Q_{11} - (\Sigma_{22}^{-1} \otimes Q_{12}) \{\Sigma_{22}^{-1} \otimes Q_{21} Q_{11}^{-1} Q_{12} + \Sigma^{22} \otimes Q_0\}^{-1} (\Sigma_{22}^{-1} \otimes Q_{21})$$

Finally, we obtain :

(A2.9) $$\mathrm{Var}(\mathrm{Vec}\ \hat{A}_1\ \mathrm{GLS}) = [\Sigma_{22}^{-1} \otimes B_1' Q_{11} B_1 - (\Sigma_{22}^{-1} \otimes B_1' Q_{12}) K^{-1} (\Sigma_{22}^{-1} \otimes Q_{21} B_1)]^{-1}$$

and this expression is identical to (A2.5).

Appendix 3 : RELATIONSHIPS AMONG THE VARIOUS 3SLS ESTIMATORS

The three 3SLS estimators presented in section (4.1) can be put in the following form :

$$\hat{\alpha}_{3SLS-i} = [W'Z_i(Z_i'\mathbf{V}Z_i)^{-1}Z_i'W]^{-1}W'Z_i(Z_i'\mathbf{V}Z_i)^{-1}Z_i'y \qquad i = 1,2,3$$

where W, **V** and y are the matrices and vector defined in section (4.2). The matrices Z_i are, respectively :

$$Z_1 = I_n \otimes X\ , \quad Z_2 = \begin{bmatrix} I_{n_1} \otimes I_T & 0 \\ 0 & I_{n_2} \otimes I_T \end{bmatrix}, \quad Z_3 = \begin{bmatrix} I_{n_1} \otimes X_1 & 0 \\ 0 & I_{n_2} \otimes X \end{bmatrix}$$

It is convenient to recall the usual 3SLS estimator, i.e. :

$$\hat{\alpha}_{3SLS-1} = [W'(\Sigma^{-1} \otimes X(X'X)^{-1}X')W]^{-1}\ W'(\Sigma^{-1} \otimes X(X'X)^{-1}X')y$$

and also to observe that for all Z_i

$$W'Z_i(Z_i'Z_i)^{-1}Z_i' = W'(I \otimes X(X'X)^{-1}X').$$

We are now in position to study the relationships among the various estimators.

(i) Equality between 3SLS-1 and 3SLS-2

From Lemma 1, Appendix 1, we can write :

$$(Z_2'\mathbf{V}Z_2)^{-1} = (Z_2'Z_2)^{-1}Z_2'\mathbf{V}^{-1}Z_2(Z_2'Z_2)^{-1} - (Z_2'Z_2)^{-1}Z_2'\mathbf{V}^{-1}S_2(S_2'\mathbf{V}^{-1}S_2)^{-1}S_2'\mathbf{V}^{-1}Z_2(Z_2'Z_2)^{-1}$$

where $S_2' = [0 \quad I_{n_2} \otimes C']$ with C such that $X'C = 0$.

We now compute the following two quantities :

- $W'Z_2(Z_2'Z_2)^{-1}Z_2'\mathbf{V}^{-1}S_2 = W'(\Sigma^{-1} \otimes X(X'X)^{-1}X')S_2 = 0$
- $(I \otimes X(X'X)^{-1}X')Z_2(Z_2'Z_2)^{-1}Z_2 = I \otimes X(X'X)^{-1}X'$

and conclude that :

$$W'Z_2(Z_2'\mathbf{V}Z_2)^{-1}Z_2' = W'[\Sigma^{-1} \otimes X(X'X)^{-1}X']Z_2(Z_2'Z_2)^{-1}Z_2'$$

$$= W'[\Sigma^{-1} \otimes X(X'X)^{-1}X']$$

Therefore, the two estimators are equal.

(ii) Equality between 3SLS-4 and 3SLS-1

The 3SLS-4 estimator is given in section (4.2), formula (4.3), with $Z = Z_3$.

From $W'Z_3(Z_3'Z_3)^{-1}Z_3' = W'(I \otimes X(X'X)^{-1}X')$ we obtain successively :

$$W'Z_3(Z_3'Z_3)^{-1}Z_3'\mathbf{V}^{-1} = W'(\Sigma^{-1} \otimes X(X'X)^{-1}X')$$

$$W'Z_3(Z_3'Z_3)^{-1}Z_3'\mathbf{V}^{-1}Z_3 = W'(\Sigma^{-1} \otimes X(X'X)^{-1}X')Z_3$$

$$W'Z_3(Z_3'Z_3)^{-1}Z_3'\mathbf{V}^{-1}Z_3 = W'\mathbf{V}^{-1}(I \otimes X(X'X)^{-1}X')Z_3$$

$$W'Z_3(Z_3'Z_3)^{-1}Z_3'\mathbf{V}^{-1}Z_3 = W'\mathbf{V}^{-1}Z_3$$

$$W'\mathbf{V}^{-1}Z_3(Z_3'\mathbf{V}^{-1}Z_3)^{-1} = W'Z_3(Z_3'Z_3)^{-1}$$

$$W'\mathbf{V}^{-1}Z_3(Z_3'\mathbf{V}^{-1}Z_3)^{-1}Z_3' = W'(I \otimes X(X'X)^{-1}X')$$

$$W'\mathbf{V}^{-1}Z_3(Z_3'\mathbf{V}^{-1}Z_3)^{-1}Z_3'\mathbf{V}^{-1} = W'(\Sigma^{-1} \otimes X(X'X)^{-1}X')$$

This last expression shows that the two estimators are equal.

(iii) <u>Equality between 3SLS-3 and the Usual 2SLS Estimator For the Coefficients of the Second Set of Equations</u>

We first note that

$$W'Z_3 = \begin{bmatrix} I \otimes X_1'X_1 & 0 \\ 0 & I \otimes H'X \end{bmatrix} = \begin{bmatrix} I \otimes X_1'X_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \otimes H'X \end{bmatrix} \overset{d}{=} F_1'F_2'$$

The non-singularity of F_1 enables us to write the 3SLS-3 estimator in the following form :

$$\hat{\alpha}_{3SLS-3} = (F_1)^{-1}[F_2'(Z_3'\mathbf{V}Z_3)^{-1}F_2]^{-1}F_2'(Z_3'\mathbf{V}Z_3)^{-1}Z_3'y$$

The coefficients pertaining to the second set of equations, which we shall denote by $\hat{\alpha}^*_{3SLS-3}$, can be obtained from $\hat{\alpha}_{3SLS-3}$ by applying the selection matrix $L_*' = [0 \quad I]$. Since $L_*'(F_1)^{-1} = L_*'$, we have :

$$\hat{\alpha}^*_{3SLS-3} = L_*'[F_2'(Z_3'\mathbf{V}Z_3)^{-1}F_2]^{-1}F_2'(Z_3'\mathbf{V}Z_3)^{-1}Z_3'y$$

Let us call $G = Z_3'\mathbf{V}Z_3$ and G^{ij} the blocs of G^{-1}. We then can write :

$$F_2'(Z_3'\mathbf{V}Z_3)^{-1}F_2 = \begin{bmatrix} G^{11} & G^{12}(I \otimes X'H) \\ (I \otimes H'X)G^{21} & (I \otimes H'X)G^{22}(I \otimes X'H) \end{bmatrix}$$

We shall need only the second row of the inverse of the above matrix, which is given by

$$R' = [R_1' \quad R_2']$$

with

$$R_2' = \Sigma_{22} \otimes (H'X(X'X)^{-1}X'H)^{-1}$$

$$R_1' = -R_2'(I \otimes H'X)G^{21}(G^{11})^{-1}$$

Consequently :

$$\hat{\alpha}^*_{3SLS-3} = R'F_2'(Z_3'\mathbf{V}Z_3)^{-1}Z_3y$$

$$= \begin{bmatrix} 0 & R_2'(I \otimes H'X)[G^{22} - G^{21}(G^{11})^{-1}G^{12}](I \otimes X') \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= R_2'(I \otimes H'X)\, G_{22}^{-1}\, (I \otimes X')y_2$$

$$= [I \otimes (H'X(X'X)^{-1}X'H)^{-1}H'X(X'X)^{-1}X']y_2$$

The equality is established. The interpretation of this result is the following. Although the equations of the first set are over-identified (they do not contain all the exogenous variables), when each equation is multiplied by X_1' (their own instruments) the resulting system of equations corresponding to the first set is square in the sense that the number of parameters is equal to the number of equations. These equations, then, play the same role as the just-identified equations in the usual 3SLS procedure, i.e. the estimation of the other equations (the second set) is not affected by them.

REFERENCES

BALESTRA, P. (1988) Ordinary and Generalized Procedures in Econometrics, Papers of the Econometric Department, University of Geneva.

BOWDEN, R.J. and TURKINGTON, D.A. (1984) Instrumental Variables, Econometric Society Monograph in Quantitative Economics, Cambridge University Press.

DHRYMES, P.J. (1973) Restricted and Unrestricted Reduced Forms. Asymptotic Distribution and Relative Efficiency, Econometrica, Vol. 41, n° 1 (January), pp. 119-134.

GOURIEROUX, C. and MONFORT A. (1985) A General Theory For Asymptotic Tests, INSEE, Document de Travail, n° 8509.

GOURIEROUX, C., MONFORT, A. and RENAULT E. (1985) Testing Unknown Linear Restrictions on Parameter Functions, CEPREMAP-D.P.

GOURIEROUX, C., MONFORT, A. and TROGNON A. (1985) Moindres carrés asymptotiques, Annales de l'INSEE, 58.

MAGNUS, J.R. (1982) Multivariate Error Components Analysis of Linear and Nonlinear Regression Models by Maximum Likelihood, Journal of Econometrics, 19, 239-285.

MILLIKEN, G.A. and ALBOHALI, M. (1984) On Necessary and Sufficient Conditions for Ordinary Least Squares Estimators to be Best Linear Unbiased, The American Statistician, 38, 298-9.

OBERHOFER, W. and KMENTA, J. (1974) A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models, Econometrica, 42, 579-590.

PESARAN, M.H., Two Step, Instrumental Variable and Maximum Likelihood Estimation of Multivariate Rational Expectations Models, presented at the Econometrics European Congress, Budapest, August 1986.

TURKINGTON, D.A. (1985) A Note on Two-stage Least Squares - three-stage Least Squares and Maximum Likelihood Estimation in an Expectations Model, International Economic Review, 26, 507-510.