working paper
department
of economics

IDENTIFICATION IN LINEAR SIMULTANEOUS EQUATIONS MODELS WITH COVARIANCE RESTRICTIONS: AN INSTRUMENTAL VARIABLES INTERPRETATION

Jerry A. Hausman
William E. Taylor*

Number 280 April 1981

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

IDENTIFICATION IN LINEAR SIMULTANEOUS EQUATIONS MODELS WITH COVARIANCE RESTRICTIONS: AN INSTRUMENTAL VARIABLES INTERPRETATION

Jerry A. Hausman
William E. Taylor*

Number 280 April 1981

* MIT and Bell Labs, respectively. Hausman thanks NSF for research support.

IDENTIFICATION IN LINEAR SIMULTANEOUS EQUATIONS MODELS WITH COVARIANCE RESTRICTIONS: AN INSTRUMENTAL VARIABLES INTERPRETATION

by

Jerry A. Hausman
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

and

William E. Taylor
Bell Telephone Laboratories, Inc.
Murray Hill, New Jersey 07974

ABSTRACT

Necessary and sufficient conditions for identifiability with linear coefficient and covariance restrictions are developed. In practical terms, covariance restrictions aid identification if and only if they imply either that (i) a set of endogenous variables is predetermined in a particular equation (generalizing the notion of recursiveness) or (ii) an identifiable residual is predetermined in a particular equation. In both cases, restrictions useful for identification yield instruments necessary for estimation; the implications for efficient estimation with covariance restrictions are worked out in a companion paper: Hausman-Taylor (1981).

IDENTIFICATION IN LINEAR SIMULTANEOUS EQUATIONS MODELS WITH COVARIANCE RESTRICTIONS: AN INSTRUMENTAL VARIABLES INTERPRETATION

by

Jerry A. Hausman
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

and

William E. Taylor
Bell Telephone Laboratories, Inc.
Murray Hill, New Jersey 07974

1. Introduction

The problems of identifiability and estimation of structural parameters in linear simultaneous equations models are closely related: see, e.g., Richmond (1974). When the necessary prior information consists of linear restrictions on the structural coefficients (henceforth "coefficient restrictions"), the early work at the Cowles Foundation determined necessary and sufficient conditions for identifiability and related these to maximum likelihood estimation, (Koopmans, Rubin, and Leipnik (1950), or Koopmans and Hood (1953)). For this case, (coefficient restrictions), the work was extended to show the relationship between identifiability and instrumental variables estimation: i.e., that the restrictions required for identification give rise to the instrumental variables required for estimation, (Fisher (1966), Theorem 2.7.2).

This picture is greatly complicated when restrictions on the structural disturbance variances and covariances, (henceforth "covariance restrictions") are allowed. Koopmans Rubin, and Leipnik (1950) recognized the usefulness of such restrictions for identification and demonstrated their equivalence to bilinear restrictions on the coefficients. This work was pursued by Wegge (1965), Rothenberg (1971), and especially Fisher (1963, 1965), surveyed in Fisher (1966, Chapters 3 and 4). Structural estimation is also complicated by covariance restrictions: as pointed out by Rothenberg and Leenders (1964), system instrumental variables estimators (3SLS) are asymptotically inefficient when covariance restrictions are present.

Two features of these results are (i) the absence of useful necessary and sufficient conditions for identifiability in the presence of covariance restrictions, and (ii) the disappearance of the link between restrictions required for identification and instrumental variables required for estimation. The problem of incorporating covariance restrictions into the theory of identification and estimation is thus incomplete. In this and a companion paper on estimation (Hausman-Taylor (1981)), we provide a simple, complete, and useful solution to the problem in terms of instrumental variables.

In the present paper, we derive necessary and sufficient conditions for identifiability in linear simultaneous equations models subject to linear restrictions on the coefficients and covariances. The result, in practical terms, shows that for covariance restrictions to aid identification, they must imply that either (i) a set of endogenous variables is predetermined in a particular equation or (ii) an identifiable residual is uncorrelated with the disturbance in a particular equation and may be treated as predetermined in that equation. Thus identifiability and the existence of instrumental variables are shown to be equivalent.

That this fact is unknown when covariance restrictions are present can be illustrated in the simplest supply and demand example. Suppose income is excluded from the supply curve, so that

$$y_1 = \beta_{12}y_2 + \gamma_{11}z_1 + u_1$$

$$y_2 = \beta_{21}y_1 \qquad\qquad + u_2,$$

where y_1 denotes equilibrium production and consumption, y_2 denotes the equilibrium price, and z_1 denotes exogenous income. Using only the prior information on the structural coefficients ($\gamma_{21} = 0$), the supply curve is identifiable and z_1 provides the necessary instrument for estimating β_{21}. Using only coefficient restrictions, the demand curve is not identifiable and there are not enough instruments to estimate β_{12} and γ_{11}.

Suppose, in addition, $\text{cov}(u_1,u_2) = \sigma_{12} = 0$. Now the reduced form parameters are given by

$$\Pi = \frac{1}{1-\beta_{12}\beta_{21}} \begin{bmatrix} \gamma_{11} \\ \\ \beta_{21}\gamma_{11} \end{bmatrix}$$

$$\Omega = \begin{bmatrix} \sigma_{11}+\beta_{12}^2\sigma_{22} & \beta_{21}\sigma_{11}+\beta_{12}\sigma_{22} \\ \\ \beta_{21}\sigma_{11}+\beta_{12}\sigma_{22} & \beta_{21}^2\sigma_{11}+\sigma_{22} \end{bmatrix} = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \\ \omega_{21} & \omega_{22} \end{bmatrix}$$

in standard notation, and we can recover all of the structural parameters from consistent estimates of the reduced form:

$$\hat{\beta}_{21} = \hat{\Pi}_2/\hat{\Pi}_1$$

$$\hat{\beta}_{12} = \frac{\hat{\beta}_{21}\hat{\omega}_{11}-\hat{\omega}_{12}}{\hat{\beta}_{21}\hat{\omega}_{12}-\hat{\omega}_{22}} \qquad \hat{\gamma}_{11} = \hat{\Pi}_1(1-\hat{\beta}_{12}\hat{\beta}_{21})$$

Thus the restriction $\sigma_{12} = 0$ makes the demand equation identifiable, yet there appears to be insufficient instruments to calculate the 2SLS estimator of (β_{12},γ_{11}).

To resolve this problem, it is important for expository purposes to proceed systematically. In Section 2, we set out the notation and the model. We begin with the limited information case in Section 3, permitting linear restrictions on the coefficients and disturbance covariances in the first equation. This reproduces the familiar results from the Cowles Commission analysis and from Fisher (1966), Chapters 2 and 3, albeit from a different point of view.

The main results are in Sections 4 and 5 in which, first, zero restrictions and then general linear restrictions on the system coefficients and covariances are appended. This completes and extends the results in Fisher (1966), Chapter 4, and, in our opinion, constitutes a satisfactory solution to the problem of identification in linear models with linear coefficient and covariance restrictions.

2. The Linear Model

Consider the classical simultaneous equations model

$$YB + Z\Gamma = U \tag{2.1}$$

in which Y is a $T\times G$ matrix of observations on the G jointly dependent random vectors y_i $(i=1,\ldots,G)$ and Z is the $T\times K$ matrix of observations on the K variables which are predetermined in every equation: i.e., $\operatorname{plim}_{T\to\infty} \frac{1}{T} Z'u_i = \underset{\sim}{0}$, where $U = [u_1,\ldots,u_G]$. The structural disturbances u_i are assumed independent and identically distributed, but contemporaneously correlated across equations. Thus $\operatorname{cov}(U_j) = \Sigma$ $(j=1,\ldots,T)$, where U_j is the j'th row of U and Σ is a positive semi-definite, symmetric, $G\times G$ matrix, which may be singular to accomodate identities among the structural equations.

Imposing exclusion restrictions and the normalization $\beta_{ii} = 1$ $(i=1,\ldots,G)$, equation (2.1) becomes

$$y_i = X_i\delta_i + u_i \quad (i=1,\ldots,G) \tag{2.2}$$

where, in the usual notation, $X_i = [Y_i \vdots Z_i]$ and $\delta_i' = [\beta_i' \vdots \gamma_i']$ so that δ_i represents the $(g_i-1) + k_i$ coefficients in the i'th structural equation which are unknown, a priori. Without loss of generality, we will be concerned with the identifiability of the unknown parameters $[\beta_1,\gamma_1,\sigma_{11}]$ in the first structural equation,

$$y_1 = Y_1\beta_1 + Z_1\gamma_1 + u_1 \quad \mathrm{var}(u_{1t}) = \sigma_{11}$$

which is, of course, equivalent to the identifiability of σ_{11} and the first column of B and Γ: i.e., of $[B_1, \Gamma_1, \sigma_{11}]$.

Since B is non-singular, equation (2.1) can be solved to yield the reduced form equations

$$Y + Z\Gamma B^{-1} = UB^{-1}$$

or

(2.3) $$Y = Z\Pi + V$$

where the reduced form parameters Π and $\Omega = \mathrm{cov}(V)$ are related to the structural parameters by

(2.4) $$\Pi = -\Gamma B^{-1}$$

(2.5) $$\Omega = (B')^{-1}\Sigma B^{-1}.$$

Since $E(Y|Z) = Z\Pi$ and $\mathrm{cov}(Y|Z) = \Omega$, the reduced form parameters are identifiable from observations on Y and Z provided only that Z is of full rank, which we assume. Moreover, the Jacobian of the transformation from U to V is non-vanishing, so that the structural parameters (B, Γ, Σ) are identifiable from observations on Y and Z if and only if (B, Γ, Σ) are uniquely determined in equations (2.4) and (2.5), given (Π, Ω) and whatever prior information on $[B, \Gamma, \Sigma]$ that we are prepared to assume. In particular, for the parameters of the first structural equation to be identifiable, it is necessary and sufficient that equations (2.4) and (2.5) admit of a unique solution for $[B_1, \Gamma_1, \sigma_{11}]$, given (Π, Ω) and prior information on $[B, \Gamma, \Sigma]$.

Let us now isolate those parts of equations (2.4) and (2.5) which actually restrict the parameters of the first structural equation.

Lemma 1: Treating (B, Γ, Σ) as unknowns, the equation systems

$$\begin{aligned} \Pi B &= -\Gamma \\ \Omega B &= (B')^{-1}\Sigma \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} \Pi B_1 &= -\Gamma_1 \\ \Omega B_1 &= (B')^{-1}\Sigma_1 \end{aligned} \tag{2.7}$$

have the same solutions $(B_1^*, \Gamma_1^*, \Sigma_1^*)$.

Proof: Rewrite equations (2.6) as

$$\begin{aligned} \Pi B_i &= -\Gamma_i \\ B'\Omega B_i &= \Sigma_i \end{aligned} \qquad i = 1, \ldots, G$$

and observe that $[B_1^*, \Gamma_1^*, \Sigma_1^*]$ solves (2.6) if and only if it solves

$$\Pi B_1 = -\Gamma_1$$

$$\begin{bmatrix} B_1' \\ B_2' \\ \cdot \\ \cdot \\ \cdot \\ B_G' \end{bmatrix} \Omega B_1 = \Sigma_1$$

by the symmetry of Ω. This, in turn, is equivalent to equations (2.7).

The principal mathematical tool for examining the uniqueness of solutions to linear equation systems is the following.

Lemma 2: If $Ax = b$ is a consistent set of linear equations, all solutions x^* are of the form

$$x^* = A^+b + [I-AA^+]\xi$$

for any vector ξ, where $()^+$ denotes any generalized inverse.

It follows from this that x^* is unique if and only if the dimension of the null space of A is zero. The fact that the existence of a unique solution to a set of linear equations turns out to characterize the identification problem, (despite the fact that equation (2.7) is not linear in B_1), prompts the following definition.

Definition: A restriction on the structural parameters will be said to be useful for identification if and only if it reduces the dimension of the null space of A in the relevant system of linear equations.

For future reference, we denote the null space of a matrix A by $N(A)$, its column space by $C(A)$, the orthogonal projection operator onto $C(A)$ by P_A, and the projection operator onto $N(A)$ by $I - P_A \equiv Q_A$.

3. Limited Information

In this section, we confine our prior information to linear restrictions on the parameters of the first structural equation, including $\Sigma_1 = (\sigma_{11}, \sigma_{12}, \ldots, \sigma_{1G})'$ with these parameters in Section 3.2. The results correspond to Chapters 2 and 3 of Fisher (1966) but from a different point of view.

3.1 Linear Restrictions on (B_1, Γ_1)

This is the classical case, treated at length in Koopmans and Hood (1953) and in Chapter 2 of Fisher (1966). The prior information on (B_1, Γ_1) is written as

$$\Phi \begin{bmatrix} B_1 \\ \\ \Gamma_1 \end{bmatrix} = \phi \tag{3.1}$$

where Φ is an rx(G+K) matrix of known constants and ϕ is a known r vector. Since Σ is entirely unknown, the only restrictions placed on (B_1, Γ_1) in equations (2.4) and (2.5) are those from

$$\Pi B_1 = -\Gamma_1 \tag{3.2}$$

and the prior restrictions in equation (3.1). Thus (B_1, Γ_1) are identifiable in equations (3.1-3.2) if and only if

$$\begin{bmatrix} \Pi & \vdots & I \\ \\ & \Phi & \end{bmatrix} \begin{bmatrix} B_1 \\ \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{0} \\ \\ \phi \end{bmatrix}$$

has a unique solution for (B_1,Γ_1). A necessary and sufficient condition for this is that

$$\text{rank}\begin{bmatrix}\Pi & \vdots & I \\ & & \\ & \Phi & \end{bmatrix} \geq G+K$$

and a necessary condition is that $r \geq G$, so that the above $(K+r)x(K+G)$ matrix of restrictions has at least K+G rows. The familiar counting rules are specializations of this result for the case of exclusion restrictions and the normalization $\beta_{11} = 1$. Here, the unknown coefficients in the first equation are (β_1,γ_1) in equation (2.2), and these are identified if and only if

$$\Pi_{12}\beta_1 = \underset{\sim}{0}$$

has a unique solution for β_1, where Π_{12} is the submatrix of Π which relates the endogenous variables included in the first equation (Y_1) with the excluded predetermined variables.

For completeness, note that σ_{11} is identifiable provided B_1 is identifiable, since $\sigma_{11} = B_1\Omega B_1'$ from equation (2.7).

3.2 Prior Information on (B_1,Γ_1,Σ_1)

Since Σ_1 is at least potentially known, identification of the parameters of the first structural equation depends upon the number of solutions $(B_1, \Gamma_1, \Sigma_1)$ to equation (3.2) and

$$(3.3) \qquad \Omega B_1 = (B')^{-1}\Sigma_1$$

given the coefficient restrictions (3.1) and covariance restrictions on Σ_1. If elements of the vector $(B')^{-1}\Sigma_1$ are known, equation (3.3) places additional restrictions on the solution(s) to equations (3.2) and (3.3). Such information is thus useful for identifying B_1, and, through equation (3.2), Γ_1. In general, though, information on $(B_1, \Gamma_1, \Sigma_1)$ is not enough to determine any element of $(B')^{-1}\Sigma_1$ uniquely. Indeed, since every element of B^{-1} is a function of elements of B_i $(i=2,\ldots,G)$ - about which we know nothing - equation (3.3) is useful for identification of B_1 if and only if our prior information about Σ_1 is that $\Sigma_1 = \underset{\sim}{0}$. In particular, two interesting special cases yield no useful information for identification purposes: (i) $\Sigma_1 \neq \underset{\sim}{0}$ entirely known, and (ii) $\Sigma_1 = (\sigma_{11}\ 0\ 0\ \ldots\ 0)'$. In both cases, the right hand side of equation (3.3) is an unknown vector, so that the equation does not restrict the set of solutions to equations (3.1-3.3).

The only case of interest in this section is thus $\Sigma_1 = \underset{\sim}{0}$, which implies that $\sigma_{11} = 0$ so that the first structural disturbance is identically zero. Now $(B')^{-1}\Sigma_1$ is known and equation (3.3) must be added to the restrictions on B_1:

<u>Proposition 1</u>: (B_1,Γ_1) are identifiable, using only prior restrictions on (B_1,Γ_1) and $\sigma_{11} = 0$, if and only if

$$\operatorname{rank}\begin{bmatrix} \Pi & \vdots & I \\ & \Phi & \\ \Omega & \vdots & \underset{\sim}{0} \end{bmatrix} \geq G+K.$$

<u>Proof</u>: Denote the matrix in question by D. A necessary and sufficient condition for the equation

$$D\begin{bmatrix} B_1 \\ \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{0} \\ \phi \\ \underset{\sim}{0} \end{bmatrix}$$

to have a unique solution (B_1,Γ_1) is that $\operatorname{rank}(D) \geq G+K$, the number of unknown parameters.

This is equivalent to the Generalized Rank Condition (Theorem 3.8.1) of Fisher (1966). This will always hold in the case in which $\operatorname{rank}(\Omega) = \operatorname{rank}(\Sigma) = G-1$, and for every additional drop in the rank of Σ, the order condition will require an additional restriction on the coefficients of the first equation. This follows from the fact that the dimension of the row space of D is $K + r + \operatorname{rank}(\Omega)$ so that the Proposition reduces to $r \geq G - \operatorname{rank}(\Omega)$.

Some intuition is gained by specializing this result for exclusion restrictions and the usual normalization.

The first structural equation is thus

$$y_1 = X_1\delta_1 + u_1$$

where the only additional restriction is $\sigma_{11} = 0$, or $u_1 = \underset{\sim}{0}$. Without loss of generality, assume the first g_1 elements of B_1 are non-zero and that the reduced form covariance matrix corresponding to Y_1:

$$\Omega = \begin{bmatrix} \Omega_{11}\Omega_{12} \\ \\ \Omega_{21}\Omega_{22} \end{bmatrix} = [\Omega_1 \vdots \Omega_2]$$

is partitioned accordingly. Specializing Proposition 1, we immediately obtain

<u>Corollary 1.1</u>: When $\sigma_{11} = 0$, a necessary and sufficient condition for the identification of δ_1 above is that

$$\text{rank}\begin{bmatrix} \Pi_{12} \\ \\ \Omega_1 \end{bmatrix} \geq g_1 - 1.$$

As before, the equation

$$\begin{bmatrix} \Pi_{12} \\ \\ \Omega_1 \end{bmatrix} \beta_1 = \underset{\sim}{0}$$

provides $(K-k_1) + \text{rank}(\Omega_1)$ restrictions on β_1, so that a necessary condition for the identifiability of β_1 is that $(K-k_1) + \text{rank}(\Omega_1) \geq g_1 - 1$, or equivalently,

<u>Corollary 1.2</u>: A necessary condition for the identifiability of β_1 is that $\text{rank}(\Omega_1) \geq (g_1-1) - (K-k_1)$: i.e., that the rank of Ω_1 exceed the nominal degree of underidentification, based on exclusion restrictions.

Note that Ω_1 is a $Gx(g_1-1)$ matrix of rank less than $g_1 - 1$; one can show that if $\text{rank}(\Omega_1) = g_1 - 1$, $\beta_1 = \underset{\sim}{0}$ and the first equation must be a reduced form equation. If the rank of Σ and thus Ω is $G-1$, the above order condition is fulfilled, even in the absence of coefficient restrictions. Indeed, Fisher (Theorem 3.9.1, (1966)) shows that β_1 is identifiable in this peculiar non-stochastic case, provided only that X_1 is of full rank.

An alternative explanation of this last result is suggestive for the more serious cases to come. Since $\Sigma_1 = \underset{\sim}{0}$, $\underset{T\to\infty}{\text{plim}} \frac{1}{T} y_i' u_1 = 0$ $(i=1,\ldots,G)$, so that every endogenous variable is predetermined in the first equation and can be used as an instrument. The two stage least squares (2SLS) estimate of δ_1 thus corresponds with least squares (OLS); and we should expect δ_1 to be identifiable whenever regression coefficients are identifiable in ordinary single equation situations. Note that this will fail if any <u>other</u> structural disturbance corresponding to endogenous variables present in the first equation is identically zero, since the "instruments" are then formally uncorrelated with the included endogenous variables.

4. Exclusion Restrictions on B

We now consider prior information about structural coefficients from equations other than the first. To our linear restrictions on (B_1,Γ_1) and zero restrictions on Σ_1, we append knowledge that certain elements of B are zero.[2]

4.1 Relative Triangularity

To determine what type of exclusion restrictions on B and Σ_1 are useful for identifying (B_1,Γ_1), we must examine the structure of B. The following assumptions will be maintained throughout:

(i) B is non-singular, $B_{ii} = 1$;

(ii) certain elements of B are zero and the remainder are unrestricted; and

(iii) sums of products of non-zero elements of B are non-zero.

The last assumption rules out events of measure zero in the parameter space and ensures that minors of B can be zero only be appropriate zero restrictions in assumption (ii).

Definition: For a G×G matrix B, a chain product corresponding to the i'th row and the j'th column is a product of no more than G-1 elements of B of the form

$$\beta_{ia}\beta_{ab}\beta_{bc}\cdot\cdot\beta_{fj}$$

where all indices are distinct. The set of all such chain products is denoted $B_{[i,j]}$.

A chain product has a number of useful properties:
(i) If B_{ji} denotes the submatrix of B obtained by deleting the j'th row and the i'th column, all elements of B_{ji} and only elements of B_{ji} appear among the products in $B_{[i,j]}$.
(ii) No element of $B_{[i,j]}$ contains a shorter element of $B_{[i,j]}$, since no elements come from the j'th row of B.
(iii) In a given chain product, each index (except i and j) appears exactly once as a row index and once as a column index. For example, for a 4×4 matrix B,

$$B_{[2,4]} = \{\beta_{24},\beta_{21}\beta_{14},\beta_{23}\beta_{34},\beta_{21}\beta_{13}\beta_{34},\beta_{23}\beta_{31}\beta_{14}\}.$$

The notion of a chain product is central to the idea of relative triangularity, which turns out to be the feature of the B matrix which is relevant for identification.

<u>Definition</u>: Equations (i,j) are <u>relatively triangular</u> if and only if $B_{[i,j]} = \{0\}$. Equations (i,j) relatively triangular does not imply that equations (j,i) are relatively triangular.

The name is motivated by the observation that B is triangular if and only if equations (i,j) are relatively triangular for all i < j. Thus for the 4×4 case, B is triangular if and only if

$$\beta_{12} = \beta_{13} = \beta_{14} = \beta_{23} = \beta_{24} = \beta_{34} = 0$$

which is necessary and sufficient that equations (1,2), (1,3),(1,4),(2,3),(2,4), and (3,4) are relatively triangular.

Relative triangularity and the intuitive notion of feedback are closely connected. For example, consider

$$y_1 = \beta_{12}y_2 + \beta_{13}y_3 + \gamma_{11}Z_1 + u_1$$

$$y_2 = \beta_{23}y_3 + \gamma_{21}Z_1 + u_2$$

$$y_3 = \beta_{32}y_2 + \gamma_{31}Z_1 + u_3.$$

The first equation has a peculiar relationship to the rest of the system, in the sense that the variables (y_2,y_3) are simultaneously determined in the second and third equations with no direct feedback from the first. Indeed, one can readily check that $B_{[3,1]} = B_{[2,1]} = \{0\}$ so that equations (3,1) and (2,1) are relatively triangular.

The set $B_{[i,j]}$ details the paths by which a shock to y_j is transmitted to y_i. In the previous example, $B_{[1,3]} = \{\beta_{13},\beta_{12}\beta_{23}\}$, and since we have imposed the normalization $\beta_{ii} = 1$, a shock to u_3 is transmitted directly to y_3 in the third equation. The shock to y_3 perturbs y_1 in the first equation in two ways: directly, since $\beta_{13} \neq 0$ and indirectly through the second equation since both β_{12} and $\beta_{23} \neq 0$

Somewhat surprisingly, the notion of relative triangularity of equations (i,j) is precisely equivalent to a zero in the (i,j)'th position of $(B')^{-1}$, denoted $(B')^{-1}_{ij}$.

<u>Lemma 3</u>: $(B')^{-1}_{ij} = 0$ if and only if equations (i,j) are relatively triangular.

The proof appears in Appendix A. In the previous example,

$$B' = \begin{bmatrix} 1 & -\beta_{12} & -\beta_{13} \\ 0 & 1 & \beta_{23} \\ 0 & \beta_{32} & 1 \end{bmatrix}$$

so that

$$(B')^{-1} = \frac{1}{1-\beta_{32}\beta_{23}} \begin{bmatrix} 1-\beta_{32}\beta_{23} & \beta_{12}-\beta_{13}\beta_{32} & \beta_{13}-\beta_{12}\beta_{23} \\ 0 & 1 & -\beta_{23} \\ 0 & -\beta_{32} & 1 \end{bmatrix}$$

which has zeros in positions (2,1) and (3,1).

The relative triangularity of equations (i,j) is obviously a necessary condition for y_j to be uncorrelated with u_i and thus to be predetermined in the i'th equation. However, as in the case of full triangularity, something further must be said about the disturbance covariances before this information is useful for identification.

4.2 <u>Relative Recursiveness</u>

We wish to determine when zero restrictions on (B,Σ_1) restrict solutions (B_1,Γ_1,Σ_1) to

$$\Pi B_1 = -\Gamma_1 \tag{4.1}$$

$$\Omega B_1 = (B')^{-1}\Sigma_1 \tag{4.2}$$

given the coefficient restrictions

$$\Phi\begin{bmatrix} B_1 \\ \\ \Gamma_1 \end{bmatrix} = \phi. \tag{4.3}$$

If all elements of the G vector $(B')^{-1}\Sigma_1$ are unknown, despite prior restrictions on (B,Σ_1), then only equations (4.1) and (4.3) are relevant for identification, and we are back in the world of Section 3.1. Under our current assumptions, if elements of $(B')^{-1}\Sigma_1$ are known, they must be zero.[3] Formalizing this, let the selection matrix Ψ be the appropriate $s\times G$ submatrix of the identity I_G.

Lemma 4: Zero restrictions on (B,Σ_1) are useful for identification if and only if they imply

$$\Psi(B')^{-1}\Sigma_1 = \underset{\sim}{0}.$$

In order for the j'th element of $(B')^{-1}\Sigma_1$ - which we denote by $((B')^{-1}\Sigma_1)_j$ - to be zero, we need zeros in the appropriate places of $(B')^{-1}$ and Σ_1. Assuming $\sigma_{11} \neq 0$,

Proposition 2: $((B')^{-1}\Sigma_1)_j = 0$ if and only if

(i) equations $(j,1)$ are relatively triangular, and

(ii) u_1 is uncorrelated with every u_k for which equations $(k,1)$ are <u>not</u> relatively triangular.

<u>Proof</u>: In general, these conditions ensure that every term in the summation that determines $((B')^{-1}\Sigma_1)_j$ is zero. Specifically, let $B_{[j,k]} \otimes \sigma_{k1}$ denote the set of products of each element of $B_{[j,k]}$ with σ_{k1} $(k=1,\ldots,G)$; the Proposition can then be restated as $((B')^{-1}\Sigma_1)_j = 0$ if and only if $B_{[j,k]} \otimes \sigma_{k1} = \{0\}$. Now $((B')^{-1}\Sigma_1)_j = 0$ is the inner product of the j'th row of $(B')^{-1}$ with Σ_1. Since sums of non-zero terms are assumed to be non-zero, $((B')^{-1}\Sigma_1)_j = 0$ if and only if $((B')^{-1})_{jk} \neq 0$ implies $\sigma_{k1} = 0$ for all $k = 1,\ldots,G$; and this is equivalent to $B_{[j,k]} \otimes \sigma_{k1} = \{0\}$ by Lemma 3.

Note that the condition on Σ_1 is less restrictive than $\sigma_{1i} = 0$ $(i=2,\ldots,G)$, which, for the entire system, would imply diagonality of Σ_1. In particular, σ_{1k} need not be zero, provided equations $(k,1)$ are relatively triangular. Since $\beta_{jj} = 1$, σ_{j1} must equal zero for $((B')^{-1}\Sigma_1)_j = 0$. If equations (i,j) are such that $((B')^{-1}\Sigma_i)_j = 0$ for all $i < j = 2,\ldots,G$, the system of equations is fully recursive, which motivates the following

<u>Definition</u>: Equations (i,j) are <u>relatively recursive</u> if and only if $((B')^{-1}\Sigma_i)_j = 0$.

Intuitively, equations (i,j) are relatively recursive if and only if there are no paths by which a shock to u_j

can be transmitted to y_i. In our previous example, equations (3,1) and (2,1) were relatively triangular. If $\sigma_{12} = 0$ but $\sigma_{13} \neq 0$, shocks to u_1 cannot reach y_2, but affect y_3 through the correlation with u_3. Thus equations (2,1) are relatively recursive and equations (3,1) are not.

Thus equations (j,1) relatively recursive suggests that y_j is uncorrelated with u_1 and can be considered to be predetermined in the first equation. This is precisely correct, as noted in Fisher (1966, Chapter 4).

<u>Proposition 3</u>: Equations (j,1) are relatively recursive if and only if y_j is predetermined in the first equation.

<u>Proof</u>: Since $\underset{T\to\infty}{\text{plim}} \frac{1}{T} Y'U = \underset{T\to\infty}{\text{plim}} \frac{1}{T} \Pi'Z'U + \underset{T\to\infty}{\text{plim}} \frac{1}{T} (B'^{-1})U'U = (B')^{-1}\Sigma$, it follows that $((B')^{-1}\Sigma_1)_j = ((B')^{-1}\Sigma)_{j1} = 0$ if and only if $\underset{T\to\infty}{\text{plim}} \frac{1}{T} y_j'u_1 = 0$.

Summarizing to this point, we have shown that zero restrictions on (B,Σ_1) are useful for identification if and only if they are equivalent to a set of endogenous variables being predetermined in the first equation. Before formally analyzing identifiability in these circumstances, we might well ask where this type of prior information is to come from.

In a sense, the information that $\Psi(B')^{-1}\Sigma_1 = \underset{\sim}{0}$ is no more arcane than the information that $\underset{T\to\infty}{\text{plim}} \frac{1}{T} Z'U = \underset{\sim}{0}$. As pictured by Fisher (1966, p. 101), the list of exogenous variables (Z) in our model is derived from the block-recursive structure of the "universe-embracing" equation system that determines all economic variables. If it is reasonable to possess the knowledge that Z is predetermined in <u>every</u> equation in the model, it is surely no less reasonable to know that an endogenous variable is predetermined in a particular equation in the model. The prior information in this section simply extends the block-recursive structure used to justify the exogeneity of Z to a relatively recursive structure which justifies treating y_j as predetermined in the first equation. A priori, there is no reason why the first type of prior information should be reasonable but the second arcane.

4.3 Identification with Exclusion Restrictions

Having zero restrictions on (B, Σ_1), identifiability of the parameters of the first equation reduces to the uniqueness of solutions $(B_1, \Gamma_1, \Sigma_1)$ to

$$\Pi B_1 = -\Gamma_1 \tag{4.4}$$

$$\Omega B_1 = (B')^{-1}\Sigma_1 \tag{4.5}$$

given

$$\Phi \begin{pmatrix} B_1 \\ \\ \Gamma_1 \end{pmatrix} = \phi, \text{ and} \tag{4.6}$$

$$\Psi\Omega B_1 = \Psi(B')^{-1}\Sigma_1 = \underset{\sim}{0}. \tag{4.7}$$

We wish to treat this as a system of linear equations in $(B_1, \Gamma_1, \Sigma_1)$ subject to linear constraints; however, the presence of $(B')^{-1}$ in equations (4.5) and (4.7) appears to rule this out. On the contrary, we shall show that the solution(s) B_1 to

$$\begin{aligned} &\Psi\Omega B_1 = \underset{\sim}{0} \\ \text{subject to } &\Psi(B')^{-1}\Sigma_1 = \underset{\sim}{0} \end{aligned}$$

are identical to the unrestricted solution(s) B_1 to

$$\Psi\Omega B_1 = \underset{\sim}{0}.$$

Intuitively, this works because the coefficients of the first equation (B_1) are irrelevant in determining if equations $(j,1)$ are relatively recursive. Formally,

<u>Lemma 5</u>: For any Σ_1, $(B')^{-1}\Sigma_1 = 0$ is consistent with any value of B_1.

<u>Proof</u>: $((B')^{-1}\Sigma_1)_j = 0 \Longleftrightarrow B_{[j,k]} \otimes \sigma_{k1} = \{0\}$ by Proposition 2. Now $B_{[j,1]}$ contains no elements of B_1, so $B_{[j,1]} = \{0\}$ does not restrict B_1. Similarly, for $k \neq 1$, $B_{[j,k]} \otimes \sigma_{k1} = \{0\}$ does not restrict any element of B_1 since any element of B_1 in a $B_{[j,k]}$ chain product must be preceded by a $B_{[j,1]}$ chain product which we know to be zero.

In our example, equations (2,1) are relatively recursive since $\sigma_{12} = 0$ and $B_{[2,1]} = \{\beta_{21}, \beta_{23}\beta_{31}\} = \{0\}$. Neither of these conditions involve $B_1 = (1\ \beta_{12}\ \beta_{13})'$. We may thus treat $\Psi\Omega B_1 = \underset{\sim}{0}$ as a linear function of B_1 and ask if the coefficient restrictions Φ and the covariance restrictions Ψ are adequate to determine (B_1,Γ_1) uniquely.

Rewriting the system as

$$\begin{bmatrix} \Pi & \vdots & I \\ & \Phi & \\ \Psi\Omega & \vdots & \underset{\sim}{0} \end{bmatrix} \begin{bmatrix} B_1 \\ \\ \Gamma_1 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{0} \\ \phi \\ \underset{\sim}{0} \end{bmatrix}$$

we obtain necessary and sufficient conditions for the identifiability of the parameters of the first equation, given linear coefficient restrictions on (B_1,Γ_1) and zero restrictions on (B,Σ_1). This is a new result, not found in Fisher (1966), and in our opinion, it represents the solution to the problem of identification with covariance restrictions:

<u>Proposition 4 (rank)</u>: A necessary and sufficient condition for the identifiability of (B_1,Γ_1) using linear restrictions on (B_1,Γ_1) and zero restrictions on (B,Σ_1) is

$$\text{rank} \begin{bmatrix} \Pi & \vdots & I \\ & \Phi & \\ \Psi\Omega & \vdots & \underset{\sim}{0} \end{bmatrix} \geq G + K.$$

The three block rows of the matrix in question have dimensions $K\times(G+K)$, $r\times(G+K)$ and $s\times(G+K)$ respectively. For the matrix to have rank $G+K$, $K+r+s$ must be greater than or equal to $G+K$.

Corollary 4.1 (order): A necessary condition for identifiability in these circumstances is that the number of endogenous variables not exceed the number of restrictions: $r+s \geq G$.

Specializing these results to the familiar case of exclusion restrictions, we obtain

$$\begin{bmatrix} \Pi_{12} \\ \\ \Psi\Omega_1 \end{bmatrix} \beta_1 = \underset{\sim}{0}$$

where Ω_1 is defined above Corollary 1.1. Thus for exclusion restrictions,

Corollary 4.2 (rank): The parameters of the first structural equation are identifiable if and only if

$$\text{rank} \begin{bmatrix} \Pi_{12} \\ \\ \Psi\Omega_1 \end{bmatrix} \geq g_1 - 1;$$

and

Corollary 4.3 (order): A necessary condition for identifiability is $(K-k_1) + s \geq g_1 - 1$: that the number of explanatory endogenous variables not exceed the number of excluded exogenous variables plus the number of covariance restrictions.

A similar interpretation in terms of instruments for the case of general linear coefficient restrictions is easily obtained. The coefficient restrictions Φ effectively reduce by r the dimension of the space spanned by the explanatory variables corresponding to unconstrained coefficients, (see Appendix B). The restrictions Ψ add s columns of Y to the K columns of Z which are predetermined in the first equation. In order to have as many instruments as explanatory variables in the first equation, we thus need r and s restrictions such that $K + G - r \leq K + S$. Using Corollary 4.1, we obtain

<u>Proposition 5 (order)</u>: A necessary condition for the identifiability of $(B_1, \Gamma_1, \Sigma_1)$ given Φ and Ψ is that the number of unconstrained coefficients in the first equation not exceed the number of instrumental variables for the first equation.

For the corresponding rank condition, we show that identification is equivalent to the existence of the 2SLS estimator, using exogenous and predetermined endogenous variables as instruments. By the argument in Appendix B, we can limit ourselves to exclusion restrictions Φ without loss of generality. Let the matrix of instruments for the first equation be

$$W = [Y\Psi' \vdots Z]$$

and recall that the 2SLS estimator must invert the matrix $X'[\hat{Y}_1 \vdots Z_1]$ where $\hat{Y}_1 = P_W Y_1$ and $Z_1 = P_W Z_1$.

Proposition 6 (rank): Rank $[\hat{Y}_1 \vdots Z_1] = g_1 - 1 + k_1$ if and only if

$$\text{rank} \begin{bmatrix} \Pi_{12} \\ \\ \Psi\Omega_1 \end{bmatrix} = g_1 - 1.$$

Proof: Note first that

$$\underset{T\to\infty}{\text{plim}} \frac{1}{T} W'V_1 = \begin{bmatrix} \underset{T\to\infty}{\text{plim}} \frac{1}{T} \Psi Y'V_1 \\ \\ \underset{T\to\infty}{\text{plim}} \frac{1}{T} Z'V_1 \end{bmatrix} = \begin{bmatrix} \Psi\Omega_1 \\ \\ \underset{\sim}{0} \end{bmatrix}$$

where V_1 is the reduced form disturbance corresponding to Y_1:

$$Y_1 = Z_1\Pi_{11} + Z_2\Pi_{12} + V_1.$$

Projecting this equation onto $\mathcal{C}(W)$, we obtain

$$P_W Y_1 = Z_1\Pi_{11} + Z_2\Pi_{12} + P_W V. \tag{4.8}$$

Now suppose the rank condition fails, so that there exists a $g_1 - 1$ vector $\lambda \neq \underset{\sim}{0}$ such that

$$\begin{bmatrix} \Pi_{12} \\ \\ \Psi\Omega_1 \end{bmatrix} \lambda = \underset{\sim}{0}.$$

Postmultiplying equation (4.8) by λ, we see that

$$P_W Y_1 \lambda = Z_1 \Pi_{11} \lambda + \underset{\sim}{0}$$

so that $\hat{Y}_1$ $(=P_W Y_1)$ and Z_1 are linearly dependent. In the other direction, suppose there exist (g_1-1) and k_1 vectors ξ_1, ξ_2 such that $\hat{Y}_1 \xi_1 + Z_1 \xi_2 = 0$. Since $P_W = I - Q_W$,

$$Y_1 \xi_1 + Z_1 \xi_2 - Q_W V_1 \xi_1 = 0$$

where $Q_W Y_1 = Q_W V_1$ since Z_1 and Z_2 are elements of $C(W)$. Adding this to the first structural equation, we obtain

$$y_1 = Y_1(\xi_1 + \beta_1) + Z_1(\xi_2 + \gamma_1) + \mu_1 - Q_W V_1 \xi_1$$

which is observationally equivalent to the first structural equation. Thus the parameters $(\beta_1, \gamma_1, \sigma_{11})$ are not identifiable and by Corollary 4.2,

$$\operatorname{rank} \begin{bmatrix} \Pi_{12} \\ \\ \Psi\Omega_1 \end{bmatrix} < g_1 - 1,$$

which completes the proof.

The intuitive correspondence between restrictions for identification and instruments for estimation is stated precisely in Proposition 6. Fisher (1966, Theorem 2.7.2) showed that for coefficient restrictions, identification was equivalent to the existence of 2SLS, using all exogenous variables as instruments. Our Lemma 4 shows that zero restrictions on (B,Σ_1) are useful for identification if and only if they are equivalent to particular endogenous variables being predetermined in the first structural equation. Proposition 6 then shows that identification is equivalent to the existence of the 2SLS estimator which uses all exogenous variables plus endogenous variables predetermined in the first equation as instruments. In this sense, then, identification given the prior information assumed in this section, is equivalent to the existence of 2SLS.

4.4 An Example

Consider the example of Section 4.1, with the added restriction $\beta_{23} = \beta_{13} = 0$

$$y_1 = \beta_{12}y_2 + \gamma_{11}z_1 + u_1$$

$$y_2 = \gamma_{21}z_1 + u_2$$

$$y_3 = \beta_{32}y_2 + \gamma_{31}z_1 + u_3,$$

where only $\sigma_{12} = 0$. Here

$$(B')^{-1}\Sigma_1 = \begin{bmatrix} \sigma_{11} \\ 0 \\ \sigma_{13} \end{bmatrix}$$

so that equations (2,1) are relatively recursive and y_2 is predetermined in the first equation. The order condition (Corollary 4.3) is fulfilled for the first equation, since there is one explanatory endogenous variable and one restriction. To check the rank condition, there are no excluded exogenous variables, so $\Pi_{12} = 0$. The reduced form covariance matrix is given by

$$\Omega = \begin{bmatrix} \sigma_{11}+\beta_{12}^2\sigma_{22} & x & x \\ \beta_{12}\sigma_{22} & \sigma_{22} & x \\ \sigma_{13}-\beta_{12}\beta_{32}\sigma_{22}+\beta_{12}\sigma_{23} & \sigma_{23}-\beta_{32}\sigma_{22} & \beta_{32}^2\sigma_{22}-2\beta_{32}\sigma_{23}+\sigma_{33} \end{bmatrix};$$

thus

$$\Omega_1 = \begin{bmatrix} \sigma_{22} \\ \beta_{12}\sigma_{22} \\ \sigma_{13}-\beta_{12}\beta_{32}\sigma_{22}+\beta_{12}\sigma_{23} \end{bmatrix} \text{ and } \Psi = (0\ 1\ 0).$$

The rank condition is satisfied, provided $\sigma_{22} \neq 0$, since

$$\text{rank}\begin{bmatrix} \Pi_{12} \\ \Psi\Omega_1 \end{bmatrix} = \text{rank}\begin{bmatrix} 0 \\ \beta_{12}\sigma_{22} \end{bmatrix} = 1 = g_1 - 1.$$

If we treat y_2 as predetermined in the first equation, the 2SLS estimate of β_{12} corresponds to the least squares estimate:

$$\hat{\beta}_{12} = \frac{y_2' Q_{Z_1} y_1}{y_2' Q_{Z_1} y_2} \, .$$

Since $Q_{Z_1} y_i = Q_{Z_1} v_i = v_i$, this estimator is equivalent to

$$\hat{\beta}_{12} = \frac{v_2' v_1}{v_2' v_2}$$

which converges in probability to $\frac{\omega_{21}}{\omega_{22}} = \beta_{12}$. Thus 2SLS using y_2 as an instrument yields the estimator obtained from solving the reduced form.

5. Linear Restrictions on (B,Γ,Σ)

In the previous section, we showed that for zero restrictions on (B,Σ_1) to be useful for identifying (B_1,Γ_1), they must restrict elements of $(B')^{-1}\Sigma_1$ to be zero. Of course, general homogeneous restrictions on the vector $(B')^{-1}\Sigma_1$ are equally useful for identification but are ruled out by the assumption that sums of products of non-zero parameters are non-zero. However, if we have enough prior information to identify the parameters of equations other than the first, we can often derive a homogeneous restriction on $(B')^{-1}\Sigma_1$. In this section, we assume sufficient information that any equation might potentially be identified.

5.1 Residuals as Instruments

There is one simple case in which homogeneous restrictions on $(B')^{-1}\Sigma_1$ are available, and, paradoxically, this case is much more relevant in practice than the relatively recursive structure in Section 4.

Proposition 7: If $\sigma_{1j} = 0$ and B_j is known or estimable, then YB_j (or $Y\hat{B}_j$) is predetermined in the first structural equation.

Proof: Suppose the j'th equation is identifiable and $\hat{B}_j$ is consistent for B_j. Then $\operatorname*{plim}_{T\to\infty} \frac{1}{T} \hat{B}_j' Y'\mu_1 = B_j'(B')^{-1}\Sigma_1 = \sigma_{1j} = 0$.

To be precise, by B_j or $\hat{B}_j$, we mean the coefficients of the endogenous variables in the first equation, after imposing normalization and other coefficient restrictions: i.e., $\hat{B}_j \equiv (1\ \hat{\beta}_j)'$. Note that $Y\hat{B}_j$ predetermined in the first equation says nothing about the status of y_j in that equation. If equations (j,1) are relative recursive, both $Y\hat{B}_j$ and y_j will be predetermined in the first equation; if not, the former is predetermined and the latter is not. If the j'th equation is not identifiable, it is possible that y_j is predetermined but that $Y\hat{B}_j$ is not calculable, and thus not available as an instrument. Formally, if d_j is a G vector of zeros with a 1 in the j'th position, the statement that both $Y\hat{B}_j'$ and y_j are predetermined in the first equation can be written as $\Psi(B')^{-1}\Sigma_1 = \underset{\sim}{0}$ where $\Psi = (d_j'\ \hat{B}_j')$.

Finally, it should be clear that one can use $Y\hat{B}_j$ or the residual $\hat{u}_j$ interchangeably as an instrumental variable in the first equation. Since $\hat{u}_j$ is a linear combination of the columns of $[Y\hat{B}_j \vdots Z]$ and orthogonal to Z, $C([Y\hat{B}_j \vdots Z]) = C([\hat{u}_j \vdots Z]) = C([\hat{u}_j \vdots Y\hat{B}_j \vdots Z])$ and either $Y\hat{B}_j$ or $\hat{u}_j$ (but not both) can be used as an instrument.

This case often appears in practice. Indeed, if every equation (other than the first) is identifiable, then every covariance restriction $\sigma_{1j} = 0$ results in an additional instrument, $\hat{u}_j$, for the first equation. In the limit,

<u>Corollary 7.1</u>: If every equation other than the first is identifiable and Σ is diagonal, then the first equation is identifiable.

5.2 Rank Considerations

Given the unlimited prior information assumed in this section, it may be the case that the j'th equation is identifiable using both coefficient and covariance restrictions. While such restrictions aid in identifying the first equation (by providing $\hat{u}_j$'s to serve as instruments), they also impose orthogonality conditions which may undermine the rank condition.

Proposition 8: If y_1 is predetermined in the j'th equation ($j \neq 1$), $\hat{u}_j$ cannot be used as an instrument in the first equation.

Proof: That y_1 is predetermined in the j'th equation implies $\text{plim} \frac{1}{T} y_1' u_j = 0$ and $\sigma_{1j} = 0$. From the first structural equation,

$$u_j' y_1 = u_j' X_1 \delta_1 + u_j' u_1$$

so that $\text{plim} \frac{1}{T} \hat{u}_j' X_1 = \text{plim} \frac{1}{T} u_j' X_1 = \underset{\sim}{0}$. As an instrument, $\hat{u}_j$ is orthogonal to the explanatory variables in the first equation and is therefore useless. Specifically,

$$\underset{T\to\infty}{\text{plim}} \frac{1}{T} y_1' \hat{u}_j = 0 \Rightarrow \underset{T\to\infty}{\text{plim}} \frac{1}{T} \hat{u}_j' Y_1 = \underset{\sim}{0}$$

$$\Rightarrow B_j' \Omega_1 = \underset{\sim}{0}.$$

Thus knowledge of B_j does not contribute to the rank condition in Corollary 4.2.

<u>Corollary 8.1</u>: If y_k is predetermined in the j'th equation, $\hat{u}_j$ cannot be used as an instrument for y_k in the first (or any other) structural equation.

The proof follows from the fact that plim $\frac{1}{T} y_k' u_j = 0$.

5.3 <u>Examples</u>

The simplest example of this case is the supply and demand model from the Introduction, which we reproduce below

$$y_1 = \beta_{12} y_2 + \gamma_{11} Z_1 + u_1$$

$$y_2 = \beta_{21} y_1 \qquad\qquad + u_2$$

where $\sigma_{12} = 0$. Here the second equation is identifiable by coefficient restrictions alone, but equations (2,1) are not relatively recursive ($\beta_{21} \neq 0$) so that y_2 is not predetermined in the first equation. However, $\hat{u}_2$ is a legitimate instrument in the first equation; given our consistent estimate of β_{21},

$$\frac{\hat{\beta}_{21}\omega_{11} - \omega_{12}}{\hat{\beta}_{21}\omega_{12} - \omega_{22}} = \frac{\hat{\beta}_{21}(\sigma_{11} + \beta_{12}^2\sigma_{22}) - (\beta_{21}\sigma_{11} + \beta_{12}\sigma_{22})}{\hat{\beta}_{21}(\beta_{21}\sigma_{11} + \beta_{12}\sigma_{22}) - (\beta_{21}^2\sigma_{11} + \sigma_{22})}$$

$$\frac{\sigma_{22}(\beta_{21}\beta_{12}^2 - \beta_{12})}{\sigma_{22}(\beta_{21}\beta_{12} - 1)} = \beta_{12}.$$

Thus the restriction $\sigma_{12} = 0$ identifies β_{12}, and the above consistent estimator is easily seen to be equivalent to

$$\hat{\beta}_{12} = \frac{(y_2-\hat{\beta}_{21}y_1)'Q_{Z_1}y_1}{(y_2-\hat{\beta}_{21}y_1)'Q_{Z_1}y_2},$$

the instrumental variables estimator for β_{12} in the demand equation, using $(y_2-\hat{\beta}_{21}y_1)$ as an instrument.

In addition to the above, Fisher uses two examples in (1966), Chapter 4. To illustrate the sufficiency condition, Theorem 4.7.1, consider (Fisher, 1966, pp. 111-112):

$$\begin{array}{llllll} y_1+\beta_{12}y_2+\beta_{13}y_3+ & & \gamma_{12}Z_2+\gamma_{13}Z_3 & = u_1 \\ \beta_{21}y_1+ & y_2+\beta_{23}y_3+\beta_{21}Z_1+ & \gamma_{23}Z_3 & = u_2 \\ \beta_{31}y_1+\beta_{32}y_2+ & y_3+ & \gamma_{32}Z_2 & = u_3 \end{array}$$

where $\sigma_{12} = \sigma_{13} = 0$. Since there are no restrictions at all on B (except $\beta_{ii} = 1$), y_i is not predetermined in equation j (i,j = 1,2,3) by Proposition 2. Equation 3 is identifiable (with probability one) by coefficient restrictions alone, but equations 1 and 2 are not: Z_1 and Z_3 provide the necessary instruments for the third equation and the first two equations are each one instrument short. Using the covariance restriction $\sigma_{13} = 0$, $\hat{u}_3$ (or $\hat{\beta}_{31}y_1+\hat{\beta}_{32}y_2+y_3$) is predetermined in equation 1; since y_1 is not predetermined in equation 3, $\hat{u}_3$ is a legitimate instrument in the first equation, identifying it by Corollary 4.2. Finally, $\sigma_{12} = 0$ implies that $\hat{u}_1$ is predetermined in the second equation, which again is identifiable by Corollary 4.2.

A similar example illustrates Theorem 4.8.1 - a necessary and sufficient condition for identifiability, given that there are no coefficient restrictions on the first equation: (Fisher, 1966, pp. 116-117).

$$\begin{aligned} y_1+\beta_{12}y_2+\beta_{13}y_3+\gamma_{11}Z_1+\gamma_{12}Z_2 &= u_1 \\ \beta_{21}+y_1+\ y_2+\qquad\qquad \gamma_{22}Z_2 &= u_2 \\ \beta_{31}y_1+\beta_{32}y_2+\ y_3+\gamma_{31}Z_1 \qquad &= u_3 \end{aligned}$$

where $\sigma_{12} = \sigma_{13} = \sigma_{23} = 0$. Here, equation 2 is identifiable by coefficient restrictions alone, equation 3 is identifiable, using $\hat{u}_2$ as an instrument, and equation 1 is identifiable using $\hat{u}_2$ and $\hat{u}_3$ as instruments. This follows from Corollary 4.2, noting that since there are no coefficient restrictions on the first equation, the rank condition becomes $\text{rank}[\Psi\Omega_1] \geq g_1 - 1$, or

$$\text{rank} \begin{bmatrix} \hat{\beta}_{21} & 1 & 0 \\ \hat{\beta}_{31} & \hat{\beta}_{32} & 1 \end{bmatrix} \begin{bmatrix} \omega_{11} \\ \omega_{12} \\ \omega_{13} \end{bmatrix} = 2 = g_1 - 1,$$

provided $\underset{T\to\infty}{\text{plim}} \frac{1}{T} \hat{u}_2'Y_1 \neq \underset{\sim}{0}$ and $\underset{T\to\infty}{\text{plim}} \frac{1}{T} \hat{u}_3'Y_1 \neq \underset{\sim}{0}$.

A final example illustrates failure of the rank condition when the order condition holds:

$$\begin{aligned} y_1 + \qquad\qquad & \beta_{13}y_3 + Z\Gamma_1 = u_1 \\ \beta_{21}y_1 + \quad y_2 \quad & + \beta_{23}y_3 + Z\Gamma_2 = u_2 \\ \beta_{31}y_1 + \qquad\qquad & \quad\; y_3 + Z\Gamma_3 = u_3 \end{aligned}$$

and $\sigma_{12} = \sigma_{23} = \sigma_{13} = 0$. Using Proposition 2, it is easily shown that y_1 and y_3 are predetermined in the second equation and that no other y_i is predetermined in any equation. If their respective equations are identifiable, $\hat{u}_2$ is predetermined in equations 1 and 3, and $(\hat{u}_1,\hat{u}_3)$ are predetermined in equation 2. However, since (y_1,y_3) are predetermined in the second equation, u_2 cannot be used as an instrument in equations (1,3) by Proposition 8. Thus the parameters of the second equation are the only identifiable parameters of the system, despite the fact that the order condition (Corollary 4.3) holds in equations (1,3)

To see this in terms of the mapping between the structure and the reduced form, observe that the instrumental variables estimator of β_{13} is given by

$$\hat{\beta}_{13} = \frac{\hat{u}_2' Q_Z y_1}{\hat{u}_2' Q_Z y_3} = \frac{\hat{\beta}_{21}\omega_{11}+\omega_{12}+\hat{\beta}_{23}\omega_{13}}{\hat{\beta}_{21}\omega_{13}+\omega_{23}+\hat{\beta}_{23}\omega_{33}}.$$

It is straight-forward but tedious to check that both the numerator and denominator of $\hat{\beta}_{13}$ is zero. This verifies the observation in the proof of Proposition 8:

$$\operatorname{rank}\begin{bmatrix}\Pi_{12}\\ \Psi\Omega_1\end{bmatrix} = \operatorname{rank}\,(\hat{B}_2\Omega_1) = \operatorname{rank}\,(\hat{\beta}_{21}\omega_{11}+\omega_{12}+\hat{\beta}_{23}\omega_{13})$$

$$= 0 < g_1 - 1 = 1,$$

so that the failure of $\hat{u}_2$ as an instrument in equation 1 implies failure of the rank condition in Corollary 4.2. Despite the fact that the order condition is fulfilled, the parameters of equation 1 are not identifiable.

6. Summary

Identifiability of the first structural equation is equivalent to the existence of a unique solution (B_1, Γ_1) to the equations which link the structural parameters and the reduced form parameters

$$\begin{aligned} \Pi B_i &= -\Gamma_i \\ \Omega B_i &= (B')^{-1}\Sigma_i \end{aligned} \qquad (i=1,\ldots,G).$$

We saw that coefficient or disturbance covariance restrictions are useful for identification if they restrict the set of solutions (B_1^*, Γ_1^*) above. It follows (Lemma 4) that covariance restrictions on (B, Σ_1) are useful for identification if and only if they identify elements (or linear combinations of elements) of $(B')^{-1}\Sigma_1$; this, in turn, is equivalent to the existence of linear combinations of the endogeneous variables which are predetermined in the first equation. Thus all restrictions which are useful for identification correspond to instruments which can be used for estimation. This extends the intuition for coefficient restrictions - developed in Fisher (1966), Theorem 2.7.2 - to the totality of useful restrictions.

This approach also yields necessary and sufficient conditions for identifiability in linear models with linear coefficient and covariance restrictions (Proposition 4) - a result which eluded Fisher, except in special cases. In practical terms, the result boils down to two useful cases:

(i) equations $(j,1)$ are relatively recursive and y_j is predetermined in equation (1), and

(ii) equation (j) is identifiable, $\sigma_{1j} = 0$, and $\hat{u}_j$ is predetermined in equation (1).

The prior information embodied in (i) is qualitatively no different than that required to assert that Z_1 is exogenous in the model; that involved in (ii) is even less. In addition, if the parameters of the first equation are over-identified, the restrictions implied by (i) and (ii) can be tested by a specification test described in Hausman-Taylor (1980). Briefly, the test compares 2SLS estimates of the structural parameters of the first equation using and omitting the instruments described in (i) or (ii). If the point estimates differ by more than sampling error, we reject relative recursiveness in (i) or the hypothesis $\sigma_{1j} = 0$ in (ii), while the maintained hypothesis is the remaining prior information.

Finally, the approach suggests a way of incorporating covariance restrictions into 3SLS or FIML estimation. Since useful covariance restrictions imply that endogenous variables may be treated as predetermined, it might be conjectured that using those instruments imposes the required restrictions in estimation. This is indeed correct and the full solution to the estimation problem with covariance restrictions is worked out in a companion paper, Hausman-Taylor (1981).

Appendix A

Lemma: $(B')^{-1}_{ij} = 0$ if and only if equations (i,j) are relatively triangular.

Proof: $(B')^{-1}_{ij}$ is proportional to the minor $|B'_{ji}|$ which is the determinant of the matrix B'_{ji} formed by striking out the j'th row and i'th column of B'. $|B'_{ji}|$ is defined as the sum of the (G-1)! distinct terms formed by multiplying together G-1 elements of B'_{ji}, exactly one from each row and column, with a rule for determining the sign.

Let $B^{(G-p)}_{[i,j]}$ denote the set of all chain products containing exactly (G-p) elements of B'_{ji}. In each chain product, exactly p-1 indices do not appear; call them $(m_1, m_2, \ldots, m_{p-1})$. Let $B_{\{m_1,\ldots,m_{p-1}\}}$ represent the set of all products of (p-1) elements of B'_{ji} - with exactly one from each row and column indexed by $(m_1,\ldots,m_{p-1})$. Finally, multiply each chain product in $B^{(G-p)}_{[i,j]}$ by one of its associated members of $B_{\{m_1,\ldots,m_{p-1}\}}$ and denote the sum of all such terms by $\Sigma[B^{(G-p)}_{[i,j]} \otimes B_{\{m_1,\ldots,m_{p-1}\}}]$.

Two characteristics of the series

$$\Sigma[B^{(G-p)}_{[i,j]} \otimes B_{\{m_1,\ldots,m_{p-1}\}}]$$

are important:

(i) <u>It contains (G-2)! distinct terms</u>.

An element of $B^{(G-p)}_{[i,j]}$ is uniquely specified by listing the G-p-1 unfixed column indices of the elements, since the chain property then determines the row indices. The first row index is i, the final column index is j and (i,j) appear in no other index. The first column index can be chosen (G-2) different ways, the second, (G-3) different ways, etc., so that there are $(G-2)\times(G-3)\times\ldots\times(G-(G-p+1)) = (G-2)!/(p-1)!$ different elements in $B^{(G-p)}_{[i,j]}$. Similarly, reordering the elements of $B_{\{m_1,\ldots,m_{p-1}\}}$ so that the row indices increase from left to right, the first column index can be chosen (p-1) different ways, etc. The number of distinct elements of $B_{\{m_1,\ldots,m_{p-1}\}}$ is thus $(p-1)!$, so that the series $\Sigma[B^{(G-p)}_{[i,j]} \otimes B_{\{m_1,\ldots,m_{p-1}\}}]$ contains $(G-2)!$ distinct terms.

(ii) <u>Each term is a product of (G-1) elements of B'_{ji} with exactly one element from each row and column</u>.

This is derived from the definition of a chain product and the construction of the residual product $B_{\{m_1,\ldots,m_{p-1}\}}$.

Now consider the following expansion for the minor $|B'_{ji}|$:

$$
\begin{aligned}
|B'_{ji}| = \; & \Sigma\, B^{(G-1)}_{[i,j]} \\
& \pm \Sigma\, B^{(G-2)}_{[i,j]} \otimes B_{\{m_1\}} \pm \ldots \\
& \pm \Sigma\, B^{(G-p)}_{[i,j]} \otimes B_{\{m_1,\ldots,m_{p-1}\}} \pm \ldots \\
& \pm \Sigma\, B^{(1)}_{[i,j]} \otimes B_{\{m_1,\ldots,m_{G-2}\}} .
\end{aligned}
$$

That is, we expand $|B'_{ji}|$, organizing terms by the length of the [i,j] chain product they contain. To verify that this equals $|B'_{ji}|$, observe that

(i) terms in different rows in the expansion are distinct since they contain chain products of different lengths, and

(ii) there are (G-1) rows, each containing (G-2)! distinct terms.

The expansion thus represents the sum of (G-1)! distinct terms, consisting of products of G-1 elements of B'_{ji} with one element from each row and column in each term. Up to the sign rule for each term, then, equation (A.1) is a valid expansion for the minor $|B'_{ji}|$. Since we are only interested in $|B'_{ji}| = 0$ and have ruled out the possibility that $\Sigma[B^{(G-p)}_{[i,j]} \otimes B_{\{m_1,\ldots,m_{p-1}\}}]$ can equal zero unless each term is zero, the sign rule for each term is irrelevant for our purposes.

The Lemma follows immediately from the expansion (A.1). Every term in the expansion contains an element of $B_{[i,j]}$, so that $B_{[i,j]} = \{0\}$ is obviously sufficient for $|B'_{ji}| = 0$. For necessity, every distinct chain product $B^{(G-p)}_{[i,j]}$ appears once in a term in (A.1) with residual elements of the form $\prod_i(\beta_{m_i m_i})$, and in the usual normalization for simultaneous equations systems $\Pi(\beta_{m_i m_i}) = 1$. Thus if any chain product differs from zero, at least one term in (A.1) differs from zero and - by assumption - $|B'_{ji}| \neq 0$.

Appendix B

Let

$$Y = X\beta + e$$
$$(B.1) \qquad R\beta = \underset{\sim}{0}$$

be a canonical linear regression model with parameters subject to r homogeneous restrictions. (The extension to nonhomogeneous restrictions is straightforward.) Let $\mathcal{R}$ denote the r dimensional subspace of $\mathcal{V}$ - the column space of X - spanned by the columns of $X(X'X)^{-1}R'$. Similarly, $\mathcal{R}\perp\mathcal{V}$ denotes the orthocomplement of $\mathcal{R}$ in $\mathcal{V}$, and P_A denotes the orthogonal projection operator onto the column space of A. Using the identity $X = P_{\mathcal{R}\perp\mathcal{V}}X + P_{\mathcal{R}}X$ and the restrictions, equation (B.1) becomes

$$(B.2) \qquad Y = P_{\mathcal{R}\perp\mathcal{V}}X\beta + e$$

since $P_{\mathcal{R}}X\beta = \underset{\sim}{0}$. The least squares estimates of β in equation (B.2) are the restricted least squares estimates of β in equation (B.1). Since the design matrix in equation (B.2) is singular (of rank p-r), we re-parameterize the equation as

$$(B.3) \qquad Y = Z\delta + e$$

where Z is a Tx(p-r) matrix whose columns span $\mathcal{R}\perp\mathcal{V}$. Least squares estimates of δ are related to restricted least squares estimates of β by

$$\hat{\delta} = (Z'Z)^{-1}Z'X\hat{\beta}_{RLS}; \ \hat{\beta}_{RLS} = (X'X)^{-1}X'Z\hat{\delta}.$$

To estimate β in equation (B.1) by 2SLS subject to the restrictions, we need only (p-r) instruments for the columns of Z, not p instruments for the columns of X. Thus if W denotes the matrix of instruments, a necessary and sufficient condition for the identifiability of β is that $Z'P_W Z$ be non-singular, so that 2SLS estimates of δ are well-defined.

Footnotes

1. We wish to acknowledge useful comments from F. Fisher and R. Radner.

2. The following results are easily extended to the case of general linear restrictions on both coefficients and covariances.

3. Since we have ruled out the possibility that sums of non-zero terms are zero.

References

Fisher, F.: "Uncorrelated Disturbances and Identifiability Criteria," International Economic Review, 4 (1963).

Fisher, F.: "Near-identifiability and the Variances of the Disturbance Terms," Econometrica, 33 (1965).

Fisher, F.: The Identification Problem in Econometrics, New York: McGraw-Hill, 1966.

Hausman, J. A., and W. E. Taylor: "Comparing Specification Tests and Classical Tests," Bell Laboratories Economic Discussion Paper #178, 1980.

Hausman, J. A., and W. E. Taylor: "Estimation of Linear Simultaneous Equations Models with Covariance Restrictions," unpublished working paper, 1981.

Koopmans, T. C.: "Identification Problems in Economic Model Construction," in W. C. Hood and T. C. Koopmans, eds., Studies in Econometric Method, New York: John Wiley & Sons, 1953, (Cowles Commission Monograph 14).pp. 27-4

Koopmans, T. C., and W. C. Hood: "The Estimation of Simultaneous Linear Economic Relationships," in W. C. Hood and T. C. Koopmans, op. cit., 1953, pp. 112-199.

Koopmans, T. C., H. Rubin, and R. B. Leipnik: "Measuring the Equation Systems of Dynamic Economics," in T. C. Koopmans, ed., Statistical Inference in Dynamic Economic Models, New York: John Wiley & Sons, 1950, (Cowles Commission Monograph 10), pp. 53-237.

Richmond, J.: "Identifiability in Linear Models," Econometrica, 42 (1974), pp. 731-736.

Rothenberg, T. J.: "Identification in Parametric Models," Econometrica, 39 (1971), pp. 577-592.

Rothenberg, T. J., and C. T. Leenders: "Efficient Estimation of Simultaneous Equation Systems," Econometrica, 32 (1964), pp. 57-76.

Wegge, L.L.: "Identifiability Criteria for a System of Equations as a Whole," Australian Journal of Statistics, 3 (1965), pp. 67-77.