Notes On The Simultaneous Equations Model

B. J. Christensen and N. M. Kiefer

March 29, 2007

Abstract

The simultaneous equations model (SEM) is the classical econometric setting for the study of identification and estimation of economic relationships. Virtually all new approaches to inference in econometrics are routinely tested and demonstrated using the SEM. In this paper we analyze the SEM in a framework which clearly illustrates the key issues without becoming mired in coordinate-specific inessential details.

1 Economies, Data and Identification

Economies are classically specified by the system of equations

$$\beta^a_r y^r + \gamma^a_k X^k + u = 0, \qquad a = 1, ..., G$$

where $\{\omega\} = \{\beta, \gamma\}$ is a $G \times (G + K)$ matrix of coefficients, repeated indices in multiplicative expressions indicate summation over the range of the index and $\{u\}$ is $G \times 1$ gaussian random vector. The space of economies is thus Ω, the space of $G \times (G+K)$ matrices. In this section we are concerned solely with specification and identification of economic models, so we abstract from considerations of sample size.

We now turn our attention to the space of observables. The sample space itself is not quite what is needed - instead we ask "what can be learned?" The answer, of course, is the reduced form $\{\pi\} \in \Psi = F(G, K)$, G-frames in R^K where $\{\pi\}$ is the matrix of coefficients in the reduced-form regressions

$$y^a = \pi_{ar} X^r + \varepsilon^a, \quad a = 1, ..., G$$

The data gives π and the identification problem is basically the sorting out of β and γ from π.

Consider now the observation window S: $\Omega \to \Psi$, mapping economies to observables. The "identification problem" arises since $\dim \Omega > \dim \Psi$ in general. The inverse map S^{-1}: $\Psi \to \Omega$ maps observations to "observationally equivalent economies" in Ω. Generally, there is only one observable associated with each economy (i.e., multiple equilibria are ruled out by the linear structure). Then the collection of sets

$$\{S^{-1}(\pi) \colon \pi \in \Psi\}$$

is a foliation of Ω. Define equivalence relation $\tilde{}$ by $\omega \;\tilde{}\; \omega'$ if $\omega,\ \omega' \in S^{-1}(\pi)$. The situation is that the data determine the leaf of the foliation. The economist is reduced to theory, intuition, and the like to select a particular economy from the observationally equivalent set along the leaf. Typically this is accomplished by restricting attention to subsets Ω' of Ω. We say that Ω' is an "economic model".

Definition 1 *The model Ω' is weakly identified if there exists an injection g: $\Omega' \to \Omega/\tilde{}$. Ω' is strongly identified if g can be chosen to be continuous.*
The model is just-identified if g can be chosen to be onto.

Let us turn now to the classical "rank and order" conditions. The order condition is
OC: $\dim \Omega' \leq \dim \Psi$
and the rank condition
RC: Define S' as the restriction of S to Ω'. The rank condition is S': $\Omega' \to \Psi$ is an injection.
Classically, the order condition is considered necessary and the rank condition sufficient for identification. Precisely:

Theorem 2 *OC is necessary for strong identifiction.*

Proof: There is no $1 - 1$ continuous map from a higher to a lower dimension manifold (Brouwer theorem on invariance of domain).

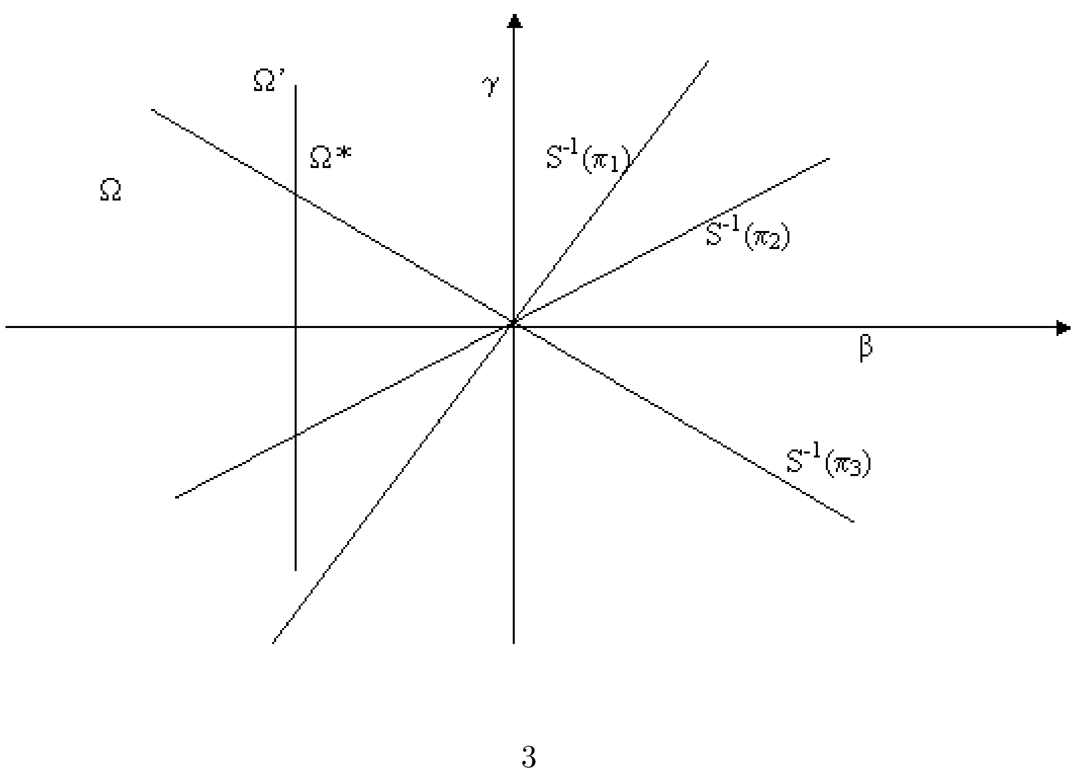
OC is not necessary for weak identification.

Theorem 3 *RC is sufficient for identification.*

Proof: $g = S^1 \circ S'$ is the required injection.

These concepts are illustrated in Figure 1 for the case of $G = K = 1$. Here the equivalance classes $S^{-1}(\pi)$ are lines through the origin (but not including the origin). The Grassman manifold $\Omega/\tilde{}$ is simply the real projective space $P^1(R)$. Restricting attention to the set Ω' gives a just identified model. The point Ω^* is an overidentified model.

Figure 1: Identification



We close this section with 2 examples then turn our attention in II to the problem of estimation.

(no examples yet)

Example 1: Klein I

Example 2:

2 Estimation and Efficiency

For any given sample, say of size N, the data can be sufficiently represented by the multivariate regression coefficient $\hat{\pi}_N$. To study the estimation problem we identify $\hat{\pi}$ with points π in Ψ and consider the map from π to $S(\Omega')$, the image of the model space in the space of data. An estimator is a map E': $\Psi \to \Omega'$ from data to models. Under the assumption that the model Ω' is identified, we can restrict attention to the map e: $\Psi \to S(\Omega')$ and carry out the complete analysis in the manifold Ψ.

Associated with a point π in Ψ is the tangent space T_π, a vector space with dimension GK. We write the basis vectors as $\{\partial^a\}$ and work with the Fisher information metric given by

$$I_{ab} = E\partial_a \ell \partial_b \ell$$

where ℓ is the loglikelihood function $\ell(\pi|Y, X)$ for the reduced form parameters. Note that $S(\Omega')$ is a submanifold of Ψ which can be parametrized in local coordinates by $\omega = (\beta, \gamma)$. The tangent space T_ω is a subspace of the space $T_{\pi(\omega)}$ where $\pi(\omega) = S(\omega)$ is a point on $S(\Omega')$ in Ψ. This tangent space has basis $\{\partial^r\}$. Clearly we have $\partial^r = b_r^a \partial_a$ and the metric on $S(\Omega')$, $I_{rs} = b_r^a b_s^b I_{ab}$.

With any estimator e and point ω we can associate the submanifold

$$A_\omega = \{e^{-1}(\omega) = \pi \in \Psi | e(\pi) = \omega\}$$

the datasets π mapping into the estimate ω via the estimator e. If A_ω is a partition of Ψ, the set $\{A_\omega$: $\omega \in \Omega'\}$ is a foliation; for smooth estimators the A_ω are at least a local foliation. This is a reasonable requirement so we assume e is $C^1(\Psi, S(\Omega'))$. We call the family A_ω an ancillary family. Denote points in A_ω by a, so (ω, a) is an alternative local coordinate system for Ψ. The tangent space to the submanifold A_ω is $T_\omega(A)$ spanned by the vectors $\{\partial_v\}$. Clearly, $\partial_v = b_v^a \partial_a$. The metric (ω, a) is I_{rs} as already given; $I_{vw} = b_v^a b_w^b I_{ab}$ and $I_{rv} = b_r^a b_v^b I_{ab}$. This last term given the angle between the tangent spaces $T_\omega(A)$ and T_ω. We are now in a position to consider evaluation of the efficiency of estimators. Thedata π in general are normally distributed (exactly under our assumptions, approximately more widely). It can be shown by expansion of the distribution of $\hat{\omega} = e(\pi)$ that the variance of $\sqrt{N}(\hat{\omega} - \omega)$ for consistent $\hat{\omega}$ is $I^{rs} = (I_{rs} - I_{rv} - I_s I^{vw})^{-1}$ plus terms in higher powers of N^{-1}, where the I^{vw} are the terms in the matrix inverse $\{I_{vw}\}^{-1}$. Thus a consistent estimator is first-order efficient iff the ancillary family is orthogonal to T_ω at $\hat{\omega}$. As usual the MLE is efficient. Thus π can be mapped into $(\hat{\omega}, a)$ with a an approximate (first-order) ancillary. These concepts are illustrated in Figure 2.

To see this more closely let $\eta = (\omega, a)$ and index η by $\{\eta_\alpha\}$, etc. It involves no loss of generality to measure a so that points on $S(\Omega')$ are $(\omega, 0)$. The Edgeworth expansion for $\sqrt{N}(\hat{\eta} - \eta^0)$ (derived from the distribution of π) is

$$p(\hat{\eta}) = N(\eta^0, I_{\alpha\beta}) \left[1 + \frac{1}{\sqrt[6]{N}} \kappa_{\alpha\beta\gamma} h^{\alpha\beta\gamma} + 0(N^{-1}) \right] \tag{1}$$

where $h^{\alpha\beta\gamma}$ are third-order tensorial Hermite polynomials

$$h^{\alpha\beta\gamma} = \eta^\alpha \eta^\beta \eta^\gamma - I^{\alpha\beta}\eta^\gamma - I^{\alpha\gamma}\eta^\beta - I^{\beta\gamma}\eta^\alpha$$

and $\kappa_{\alpha\beta\gamma} = 3 - \Gamma^{(-1/3)}_{\alpha\beta\gamma}$; $\Gamma_{\alpha\beta\gamma}$ relates the basis vectors in different tangent spaces thus controlling for nonnormality in the asymptotic distributions. Of course, we are primarily interested in the marginal distribution of the estimator $\sqrt{N}(\hat{\omega} - \omega^0)$. This has variance I^{rs} given above to first-order. For a first-order efficient estimator, such as the MLE, we may exploit the orthogonality of the ancillary family to show that the second-order expansion can be obtained in the same form as (1.1) with subscripts rst rather than $\alpha\beta\gamma$.

3 Testing

Typically we are interested in inspecting whether the economy theory Ω' at hand may be rejected. In classical terms this amounts to testing the overidentifying restrictions. If $\dim S(\Omega') < \dim \Psi$ we may select an economy $\Omega'' \subseteq \Omega$ that satisfies exact strong identification. We look for rejection of the reduction from Ω'' to Ω'. A test T is given by a rejection space $R \subset \Psi$ such that the hypothesis H_0: $\omega \in \Omega'$ is rejected iff $\pi \in R$. The power of T as ω is perturbed to order $N^{-1/2}$ away from Ω' may be expanded in terms of increasing order in $N^{-1/2}$, and T is first-order uniformly efficient if the leading term (weakly) dominates that of any other test at the same level. The test T is first-order uniformly efficient iff the ancillary family is asymptotically orthogonal, in particular the likelihood ratio test $LR = \sup_{\omega \in \Omega'} L(S(\omega); \pi) L(\pi; \pi)$ uniformly efficient. Here $L(S(\omega); \cdot)$ indicates the dependence of the likelihood function L on data π. In applications, when N is finite, the overidentifying restrictions are rejected if $T = -2 \log LR/b$ is critical in a $\chi^2(\dim \Omega'' - \dim \Omega')$ table, since Bartlett adjustments b leaves the χ^2-distribution for T valid to order $N^{-3/2}$ (as opposed to the usual rate N^{-1}).