10.6

Example ~~13.2~~ Klein's Model I

A widely used example of a simultaneous equations model of the economy is Klein's (1950) *Model I*. The model may be written

$$C_t = \alpha_0 + \alpha_1 P_t + \alpha_2 P_{t-1} + \alpha_3 (W_t^p + W_t^g) + \varepsilon_{1t} \quad \text{(consumption)},$$
$$I_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \beta_3 K_{t-1} \qquad + \varepsilon_{2t} \quad \text{(investment)},$$
$$W_t^p = \gamma_0 + \gamma_1 X_t + \gamma_2 X_{t-1} + \gamma_3 A_t \qquad + \varepsilon_{3t} \quad \text{(private wages)},$$
$$X_t = C_t + I_t + G_t \quad \text{(equilibrium demand)},$$
$$P_t = X_t - T_t - W_t^p \quad \text{(private profits)},$$
$$K_t = K_{t-1} + I_t \quad \text{(capital stock)}.$$

The endogenous variables are each on the left-hand side of an equation and are labeled on the right. The exogenous variables are G_t =government nonwage spending, T_t =indirect business taxes plus net exports, W_t^g =government wage bill, A_t =time trend measured as years from 1931, and the constant term. There are also three predetermined variables: the lagged values of the capital stock, private profits, and total demand. The model contains three behavioral equations, an equilibrium condition and two accounting identities. This model provides an excellent example of a small, dynamic model of the economy. It has also been widely used as a test ground for simultaneous equations estimators. Klein estimated the parameters using yearly data for 1921 to 1941. The data are listed in Appendix Table F10.2. Table 10.4 presents limited and full information estimates for Klein's Model I based on the original data for 1920–1941.[44] 46

5/

TB 10.8

FN 46

Au: Check renumbering of tables. OK?

TABLE ~~13.3~~ 10.5 Estimates of Klein's Model I (Estimated Asymptotic Standard Errors in Parentheses)

	Limited Information Estimates				Full Information Estimates			
	2SLS				**3SLS**			
C	16.6	0.017	0.216	0.810	16.4	0.125	0.163	0.790
	(1.32)	(0.118)	(0.107)	(0.040)	(1.30)	(0.108)	(0.100)	(0.038)
I	20.3	0.150	0.616	−0.158	28.2	−0.013	0.756	−0.195
	(7.54)	(0.173)	(0.162)	(0.036)	(6.79)	(0.162)	(0.153)	(0.033)
W^p	1.50	0.439	0.147	0.130	1.80	0.400	0.181	0.150
	(1.15)	(0.036)	(0.039)	(0.029)	(1.12)	(0.032)	(0.034)	(0.028)
	LIML				**FIML**			
C	17.1	−0.222	0.396	0.823	18.3	−0.232	0.388	0.802
	(1.84)	(0.202)	(0.174)	(0.055)	(2.49)	(0.312)	(0.217)	(0.036)
I	22.6	0.075	0.680	−0.168	27.3	−0.801	1.052	−0.146
	(9.24)	(0.219)	(0.203)	(0.044)	(7.94)	(0.491)	(0.353)	(0.30)
W^p	1.53	0.434	0.151	0.132	5.79	0.234	0.285	0.235
	(2.40)	(0.137)	(0.135)	(0.065)	(1.80)	(0.049)	(0.045)	(0.035)
	OLS				**I3SLS**			
C	16.2	0.193	0.090	0.796	16.6	0.165	0.177	0.766
	(1.30)	(0.091)	(0.091)	(0.040)	(1.22)	(0.096)	(0.090)	(0.035)
I	10.1	0.480	0.333	−0.112	42.9	−0.356	1.01	−0.260
	(5.47)	(0.097)	(0.101)	(0.027)	(10.6)	(0.260)	(0.249)	(0.051)
W^p	1.50	0.439	0.146	0.130	2.62	0.375	0.194	0.168
	(1.27)	(0.032)	(0.037)	(0.032)	(1.20)	(0.031)	(0.032)	(0.029)

45 [44] The asymptotic covariance matrix for the LIML estimator will differ from that for the 2SLS estimator in a finite sample because the estimator of σ_{jj} that multiplies the inverse matrix will differ and because in computing the matrix to be inverted, the value of "k" (see the equation after (10-55)) is one for 2SLS and the smallest root in (10-54) for LIML. Asymptotically, k equals one and the estimators of σ_{jj} are equivalent.

10-57

It might seem, in light of the entire discussion, that one of the structural estimators described previously should always be preferred to ordinary least squares, which, alone among the estimators considered here, is inconsistent. Unfortunately, the issue is not so clear. First, it is often found that the OLS estimator is surprisingly close to the structural estimator. It can be shown that, at least in some cases, OLS has a smaller variance about its mean than does 2SLS about its mean, leading to the possibility that OLS might be more precise in a mean-squared-error sense.[45] But this result must be tempered by the finding that the OLS standard errors are, in all likelihood, not useful for inference purposes.[46] Nonetheless, OLS is a frequently used estimator. Obviously, this discussion is relevant only to finite samples. Asymptotically, 2SLS must dominate OLS, and in a correctly specified model, any full information estimator must dominate any limited information one. The finite sample properties are of crucial importance. Most of what we know is asymptotic properties, but most applications are based on rather small or moderately sized samples.

The large difference between the inconsistent OLS and the other estimates suggests the bias discussed earlier.On the other hand, the incorrect sign on the LIML and FIML estimate of the coefficient on P and the even larger difference of the coefficient on P_{-1} in the C equation are striking. Assuming that the equation is properly specified, these anomalies would likewise be attributed to finite sample variation, because LIML and 2SLS are asymptotically equivalent.

Intuition would suggest that systems methods, 3SLS and FIML, are to be preferred to single-equation methods, 2SLS and LIML. Indeed, if the advantage is so transparent, why would one ever choose a single-equation estimator? The proper analogy is to the use of single-equation OLS versus GLS in the SURE model of Section 10.2. An obvious practical consideration is the computational simplicity of the single-equation methods. But the current state of available software has eliminated this advantage.

Although the systems methods are asymptotically better, they have two problems. First, any specification error in the structure of the model will be propagated throughout the system by 3SLS or FIML. The limited information estimators will, by and large, confine a problem to the particular equation in which it appears. Second, in the same fashion as the SURE model, the finite-sample variation of the estimated covariance matrix is transmitted throughout the system. Thus, the finite-sample variance of 3SLS may well be as large as or larger than that of 2SLS. Although they are only single estimates, the results for Klein's Model I give a striking example. The upshot would appear to be that the advantage of the systems estimators in finite samples may be more modest than the asymptotic results would suggest. Monte Carlo studies of the issue have tended to reach the same conclusion.[47]

[45] See Goldberger (1964, pp. 359-360).

[46] Cragg (1967).

[47] See Cragg (1967) and the many related studies listed by Judge et al. (1985, pp. 646-653).

10-58

estimator.[16] The implication is that with normally distributed disturbances, 2SLS is fully efficient.

The ***k* class** of estimators is defined by the following form

$$\hat{\delta}_{j,k} = \begin{bmatrix} \mathbf{Y}_j'\mathbf{Y}_j - k\mathbf{V}_j'\mathbf{V}_j & \mathbf{Y}_j'\mathbf{X}_j \\ \mathbf{X}_j'\mathbf{Y}_j & \mathbf{X}_j'\mathbf{X}_j \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{Y}_j'\mathbf{y}_j - k\mathbf{V}_j'\mathbf{v}_j \\ \mathbf{X}_j'\mathbf{y}_j \end{bmatrix}.$$

We have already considered three members of the class, OLS with $k = 0$, 2SLS with $k = 1$, and, it can be shown, LIML with $k = \lambda_1$. [This last result follows from (13-26).] There have been many other k-class estimators derived; Davidson and MacKinnon (2004, pp. 537–538 and 548–549) and Mariano (2001) give discussion. It has been shown that all members of the k class for which k converges to 1 at a rate faster than $1/\sqrt{n}$ have the same asymptotic distribution as that of the 2SLS estimator that we examined earlier. These are largely of theoretical interest, given the pervasive use of 2SLS or OLS, save for an important consideration. The large sample properties of all k-class estimators are the same, but the finite-sample properties are possibly very different. Davidson and MacKinnon (2004, pp. 537–538 and 548–549) and Mariano (1982, 2001) suggest that some evidence favors LIML when the sample size is small or moderate and the number of overidentifying restrictions is relatively large.

10.6.6 ~~13.5.5~~ TESTING IN THE PRESENCE OF WEAK INSTRUMENTS

In Section ~~12.9~~ 8.7, we introduced the problems of estimation and inference with instrumental variables in the presence of weak instruments. The first-stage regression method of Staiger and Stock (1997) is often used to detect the condition. Other tests have also been proposed, notably that of Hahn and Hausman (2002, 2003). Consider an equation with a single endogenous variable on the right-hand side,

$$y_1 = \gamma y_2 + \mathbf{x}_1'\beta_1 + \varepsilon_1.$$

Given the way the model has been developed, the placement of y_1 on the left-hand side of this equation and y_2 on the right represents nothing more than a normalization of the coefficient matrix $\mathbf{\Gamma}$ in (~~13-2~~). ~~(Note point this Section 13.3.1.)~~ For the moment, label this the "forward" equation. If we renormalize the model in terms of y_2, we obtain the completely equivalent equation 10-42

$$\begin{aligned} y_2 &= (1/\gamma)y_1 + \mathbf{x}_1'(\beta_1/\gamma) + \varepsilon_1/\gamma \\ &= \theta y_1 + \mathbf{x}_1'\lambda_1 + v_1, \end{aligned}$$

which we [i.e., Hahn and Hausman (2002)] label the "reverse equation." In principle, for estimation of γ, it should make no difference which form we estimate; we can estimate γ directly in the first equation or indirectly through $1/\theta$ in the second. However, in practice, of all the k-class estimators listed in Section ~~13.5.4~~, which includes all the estimators we have examined, only the LIML estimator is invariant to this renormalization; certainly the 2SLS estimator is not. If we consider the forward 2SLS estimator, $\hat{\gamma}$, and the reverse estimator, $1/\hat{\theta}$, we should in principle obtain similar estimates. But there 10.6.4

[16] This is proved by showing that both estimators are members of the "k class" of estimators, all of which have the same asymptotic covariance matrix. Details are given in Theil (1971) and Schmidt (1976).

10-59

is a bias in the 2SLS estimator that becomes more pronounced as the instruments become weaker. The Hahn and Hausman test statistic is based on the difference between these two estimators (corrected for the known bias of the 2SLS estimator in this case). [Research on this and other tests is ongoing. Hausman, Stock, and Yogo (2005) do report rather disappointing results for the power of this test in the presence of irrelevant instruments.]

The problem of inference remains. The upshot of the development so far is that the usual test statistics are likely to be unreliable. Some useful results have been obtained for devising inference procedures that are more robust than the standard first order asymptotics that we have employed (for example, in Theorem ~~12.1~~ 8.1 and Section ~~13.5.3~~ 10.6). Kleibergen (2002) has constructed a class of test statistics based on Anderson and Rubin's (1949, 1950) results that appears to offer some progress. An intriguing aspect of this strand of research is that the Anderson and Rubin test was developed in their 1949 and 1950 studies and predates by several years the development of two-stage least squares by Theil (1953) and Basmann (1957). [See Stock and Trebbi (2003) for discussion of the early development of the method of instrumental variables.] A lengthy description of Kleibergen's method and several extensions appears in the survey by Dufour (2003), which we draw on here for a cursory look at the Anderson and Rubin statistic.

The simultaneous equations model in terms of equation 1 is written

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{Y}_1\boldsymbol{\gamma}_1 + \boldsymbol{\varepsilon}_1, \\ \mathbf{Y}_1 &= \mathbf{X}_1\boldsymbol{\Pi}_1 + \mathbf{X}_1^*\boldsymbol{\Pi}_1^* + \mathbf{V}_1, \end{aligned} \qquad \text{10-64 } \cancel{(13\text{-}27)}$$

where $\mathbf{y}_1$ is the n observations on the left-hand variable in the equation of interest, $\mathbf{Y}_1$ is the n observations on M_1 endogenous variables in this equation, γ_1 is the structural parameter vector in this equation, and $\mathbf{X}_1$ is the K_1 included exogenous variables in equation 1. ~~(See Table 13.1.)~~ The second equation is the set of M_1 reduced form equations for the included endogenous variables that appear in equation 1. (Note that M_1^* endogenous variables, $\mathbf{Y}_1^*$, are excluded from equation 1.) The full set of exogenous variables in the model is

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_1^*],$$

where X_1^* is the K_1^* exogenous variables that are excluded from equation 1. (We are changing Dufour's notation slightly to conform to the conventions used in our development of the model.) Note that the second equation represents the first stage of the two-stage least squares procedure.

We are interested in inference about γ_1. We must first assume that the model is identified. We will invoke the rank and order conditions as usual. The order condition is that there must be at least as many excluded exogenous variables as there are included endogenous variables, which is that $K_1^* \geq M_1$. ~~(With the model in the preceding form, it is easy to see the logic of this condition. If we are going to apply 2SLS by regressing $\mathbf{y}_1$ on a prediction for $\mathbf{Y}_1$ that is a linear combination of the variables in $\mathbf{X}$, then in order for the resulting regressor matrix to have full column rank, the predicted $\mathbf{Y}_1$ in equation 1 above must involve at least enough variables that it is linearly independent of $\mathbf{X}_1$.)~~ For the rank condition to be met, we must have

$$\boldsymbol{\pi}_1^* - \boldsymbol{\Pi}_1^*\boldsymbol{\gamma}_1 = \mathbf{0}.$$

10-60

where π_1^* is the second part of the coefficient vector in the reduced form equation for $\mathbf{y}_1$, that is,

$$\mathbf{y}_1 = \mathbf{X}_1\pi_1 + \mathbf{X}_1^*\pi_1^* + \mathbf{v}_1.$$

~~This is equation (13-48)~~. For this result to hold, Π_1^* must have full column rank, K_1^*. The weak instruments problem is embodied in Π_1^*. If this matrix has short rank, the parameter vector γ_1 is not identified. The weak instruments problem arises when Π_1^* is nearly short ranked. The important aspect of that observation is that the weak instruments can be characterized as an identification problem.

Anderson and Rubin (1949, 1950) (AR) proposed a method of testing $H_0: \gamma_1 = \gamma_1^0$. The AR statistic is constructed as follows: Combining the two equations in (13-27), we have

$$\mathbf{y}_1 = \mathbf{X}_1\beta_1 + \mathbf{X}_1\Pi_1\gamma_1 + \mathbf{X}_1^*\Pi_1^*\gamma_1 + \varepsilon_1 + \mathbf{V}_1\gamma_1.$$

10-64

Using (~~13-27~~) again, subtract $\mathbf{Y}_1\gamma_1^0$ from both sides of this equation to obtain

$$\begin{aligned}\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0 &= \mathbf{X}_1\beta_1 + \mathbf{X}_1\Pi_1\gamma_1 + \mathbf{X}_1^*\Pi_1^*\gamma_1 + \varepsilon_1 + \mathbf{V}_1\gamma_1 \\ &\quad - \mathbf{X}_1\Pi_1\gamma_1^0 - \mathbf{X}_1^*\Pi_1^*\gamma_1^0 - \mathbf{V}_1\gamma_1^0 \\ &= \mathbf{X}_1[\beta_1 + \Pi_1(\gamma_1 - \gamma_1^0)] + \mathbf{X}_1^*[\Pi_1^*(\gamma_1 - \gamma_1^0)] + \varepsilon_1 + \mathbf{V}_1(\gamma_1 - \gamma_1^0) \\ &= \mathbf{X}_1\theta_1 + \mathbf{X}_1^*\theta_1^* + \mathbf{w}_1.\end{aligned}$$

Under the null hypothesis, this equation reduces to

$$\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0 = \mathbf{X}_1\theta_1 + \mathbf{w}_1,$$

so a test of the null hypothesis can be carried out by testing the hypothesis that θ_1^* equals zero in the preceding partial reduced-form equation. Anderson and Rubin proposed a simple F test,

$$\begin{aligned}\mathrm{AR}(\gamma_1^0) &= \frac{[(\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0)'\mathbf{M}_1(\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0) - (\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0)'\mathbf{M}(\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0)]/K_1^*}{(\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0)'\mathbf{M}(\mathbf{y}_1 - \mathbf{Y}_1\gamma_1^0)/(n-K)} \\ &\sim F[K_1^*, n-K],\end{aligned}$$

where $\mathbf{M}_1 = [\mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1']$ and $\mathbf{M} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']$. This is the standard F statistic for testing the hypothesis that the set of coefficients is zero in the classical
29 linear regression. [See (5-29).] [Dufour (2003) shows how the statistic can be extended to allow more general restrictions that also include β_1.]

There are several striking features of this approach, beyond the fact that it has been available since 1949: (1) its distribution is free of the model parameters in finite samples (assuming normality of the disturbances); (2) *it is robust to the weak instruments problem;* (3) it is robust to the exclusion of other instruments; and (4) it is robust to specification errors in the structural equations for $\mathbf{Y}_1$, the other variables in the equation. There are some shortcomings as well, namely: (1) the tests developed by this method are only applied to the full parameter vector; (2) the power of the test may diminish as more (and too many more) instrumental variables are added; (3) it relies on a normality assumption for the disturbances; and (4) there does not appear to be a counterpart for nonlinear systems of equations.

10.7 SUMMARY AND CONCLUSIONS

This chapter has surveyed the specification and estimation of multiple equations models. The SUR model is an application of the generalized regression model introduced in Chapter 9. The advantage of the SUR formulation is the rich variety of behavioral models that fit into this framework. We began with estimation and inference with the SUR model, treating it essentially as a generalized regression. The major difference between this set of results and the single equation model in Chapter 9 is practical. While the SUR model is, in principle a single equation GR model with an elaborate covariance structure, special problems arise when we explicitly recognize its intrinsic nature as a set of equations linked by their disturbances. The major result for estimation at this step is the feasible GLS estimator. In spite of its apparent complexity, we can estimate the SUR model by a straightforward two-step GLS approach that is similar to the one we used for models with heteroscedasticity in Chapter 9. We also extended the SUR model to autocorrelation and heteroscedasticity. Once again, the multiple equation nature of the model complicates these applications. Section 10.4 presented a common applications of the seemingly unrelated regressions model, the estimation of demand systems. One of the signature features of this literature is the seamless transition from the theoretical models of optimization of consumers and producers to the sets of empirical demand equations derived from Roy's identity for consumers and Shephard's lemma for producers.

The multiple equations models surveyed in this chapter involve most of the issues that arise in analysis of linear equations in econometrics. Before one embarks on the process of estimation, it is necessary to establish that the sample data actually contain sufficient information to provide estimates of the parameters in question. This is the question of identification. Identification involves both the statistical properties of estimators and the role of theory in the specification of the model. Once identification is established, there are numerous methods of estimation. We considered a number of single-equation techniques, including least squares, instrumental variables, and maximum likelihood. Fully efficient use of the sample data will require joint estimation of all the equations in the system. Once again, there are several techniques—these are extensions of the single equation methods including three stage least squares, and full information maximum likelihood. In both frameworks, this is one of those benign situations in which the computationally simplest estimator is generally the most efficient one.

10-62

AU: Edits to KTs OK?

of equations

Key Terms and Concepts

- Admissible
- Behavioral equation
- Complete system
- Constant returns to scale
- Dynamic model
- Equilibrium condition
- Exclusion restrictions
- FIML
- Full information estimator
- Generalized regression model
- Homogeneity restriction
- Identification
- Invariance
- Kronecker product
- Likelihood ratio test
- Maximum likelihood
- Nonsample information
- Observationally equivalent
- Pooled model
- Projection
- Recursive model
- Restrictions
- Shephard's lemma
- Specification test
- Structural equation
- Three-stage least squares (3SLS) estimator
- Underidentified
- Autocorrelation
- Causality
- Completeness condition
- Covariance structures model
- Econometric model
- Equilibrium multipliers
- Exogenous
- Fixed effects
- Full information maximum likelihood
- Granger causality
- Identical explanatory variables
- Instrumental variable estimator
- Jointly dependent
- Lagrange multiplier test
- Limited information estimator
- Multivariate regression model
- Nonstructural
- Order condition
- Predetermined variable
- Random effects model
- Reduced form
- Seemingly unrelated regressions
- Singular disturbance covariance matrix
- Singularity of the
- Strongly exogenous
- Structural form
- Triangular system
- Weak instruments
- Balanced panel
- Cobb-Douglas model
- Consistent estimators
- Demand system
- Endogenous
- Exactly identified model
- Feasible GLS
- Flexible functional form
- Fully recursive model
- Heteroscedasticity
- Identical regressors
- Interdependent
- k class
- Least variance ratio
- LIML estimator
- Limited information maximum likelihood
- Nonlinear systems
- Normalization
- Overidentification
- Problem of identification
- Rank condition
- Reduced form disturbance
- Share equations
- Simultaneous equations bias
- Structural disturbance
- System methods of estimation
- Two-stage least squares (2SLS) estimator
- Weakly exogenous

AU: The following terms were not bold KTs in text:

~~Balanced panel~~
Behavioral equation
~~Causality~~
~~Consistent estimators~~
Equilibrium condition
~~Equilibrium multipliers~~
~~Exactly identified model~~
Exclusion restrictions
FIML
~~Fixed effects~~
Flexible functional form
Full information maximum likelihood
~~Fully recursive model~~
Identification
Nonsample information

~~Nonstructural~~
Order condition
~~Overidentification~~
~~Projection~~
~~Random effects model~~
Rank condition
Restrictions
Seemingly unrelated regressions
~~Simultaneous equations bias~~
~~Specification test~~
System methods of estimation
~~Underidentified~~
Weak instruments

10-63

- Lagrange multiplier test
- Likelihood ratio test
- Maximum likelihood
- Multivariate regression model
- Pooled model
- Projection
- Random effects model
- Seemingly unrelated regressions
- Share equations
- Shephard's lemma
- Singular disturbance covariance matrix
- System of demand equations
- Taylor series
- Translog function
- Wald statistic
- Zellner's efficient estimator

Exercises

1. A sample of 100 observations produces the following sample data:

$$\bar{y}_1 = 1, \quad \bar{y}_2 = 2,$$
$$\mathbf{y}_1'\mathbf{y}_1 = 150,$$
$$\mathbf{y}_2'\mathbf{y}_2 = 550,$$
$$\mathbf{y}_1'\mathbf{y}_2 = 260.$$

The underlying bivariate regression model is

$$y_1 = \mu + \varepsilon_1,$$
$$y_2 = \mu + \varepsilon_2.$$

a. Compute the OLS estimate of μ, and estimate the sampling variance of this estimator.
b. Compute the FGLS estimate of μ and the sampling variance of the estimator.

2. Consider estimation of the following two-equation model:

$$y_1 = \beta_1 + \varepsilon_1,$$
$$y_2 = \beta_2 x + \varepsilon_2.$$

A sample of 50 observations produces the following moment matrix:

$$\begin{array}{c} \\ 1 \\ y_1 \\ y_2 \\ x \end{array} \begin{array}{c} \begin{array}{cccc} 1 & y_1 & y_2 & x \end{array} \\ \begin{bmatrix} 50 & & & \\ 150 & 500 & & \\ 50 & 40 & 90 & \\ 100 & 60 & 50 & 100 \end{bmatrix} \end{array}.$$

a. Write the explicit formula for the GLS estimator of $[\beta_1, \beta_2]$. What is the asymptotic covariance matrix of the estimator?
b. Derive the OLS estimator and its sampling variance in this model.
c. Obtain the OLS estimates of β_1 and β_2, and estimate the sampling covariance matrix of the two estimates. Use n instead of $(n-1)$ as the divisor to compute the estimates of the disturbance variances.
d. Compute the FGLS estimates of β_1 and β_2 and the estimated sampling covariance matrix.
e. Test the hypothesis that $\beta_2 = 1$.

10 - 64

3. The model

$$y_1 = \beta_1 x_1 + \varepsilon_1,$$
$$y_2 = \beta_2 x_2 + \varepsilon_2$$

satisfies all the assumptions of the classical multivariate regression model. All variables have zero means. The following sample second-moment matrix is obtained from a sample of 20 observations:

$$\begin{array}{c} \\ y_1 \\ y_2 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{matrix} y_1 & y_2 & x_1 & x_2 \end{matrix} \\ \begin{bmatrix} 20 & 6 & 4 & 3 \\ 6 & 10 & 3 & 6 \\ 4 & 3 & 5 & 2 \\ 3 & 6 & 2 & 10 \end{bmatrix} \end{array}.$$

a. Compute the FGLS estimates of β_1 and β_2.
b. Test the hypothesis that $\beta_1 = \beta_2$.
c. Compute the maximum likelihood estimates of the model parameters.
d. Use the likelihood ratio test to test the hypothesis in part b.

4. Prove that in the model

$$\mathbf{y}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_1,$$
$$\mathbf{y}_2 = \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}_2,$$

generalized least squares is equivalent to equation-by-equation ordinary least squares if $\mathbf{X}_1 = \mathbf{X}_2$. Does your result hold if it is also known that $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$?

5. Consider the two-equation system

$$y_1 = \beta_1 x_1 \qquad\qquad + \varepsilon_1,$$
$$y_2 = \qquad \beta_2 x_2 + \beta_3 x_3 + \varepsilon_2.$$

Assume that the disturbance variances and covariance are known. Now suppose that the analyst of this model applies GLS but erroneously omits x_3 from the second equation. What effect does this specification error have on the consistency of the estimator of β_1?

6. Consider the system

$$y_1 = \alpha_1 + \beta x + \varepsilon_1,$$
$$y_2 = \alpha_2 \qquad + \varepsilon_2.$$

The disturbances are freely correlated. Prove that GLS applied to the system leads to the OLS estimates of α_1 and α_2 but to a mixture of the least squares slopes in the regressions of y_1 and y_2 on x as the estimator of β. What is the mixture? To simplify the algebra, assume (with no loss of generality) that $\bar{x} = 0$.

7. For the model

$$y_1 = \alpha_1 + \beta x + \varepsilon_1,$$
$$y_2 = \alpha_2 \qquad + \varepsilon_2,$$
$$y_3 = \alpha_3 \qquad + \varepsilon_3,$$

assume that $y_{i2} + y_{i3} = 1$ at every observation. Prove that the sample covariance matrix of the least squares residuals from the three equations will be singular, thereby precluding computation of the FGLS estimator. How could you proceed in this case?

10-65

theory in the specification of the model. Once identification is established, there are numerous methods of estimation. We considered a number of single-equation techniques, including least squares, instrumental variables, and maximum likelihood. Fully efficient use of the sample data will require joint estimation of all the equations in the system. Once again, there are several techniques—these are extensions of the single-equation methods including three stage least squares, and full information maximum likelihood. In both frameworks, this is one of those benign situations in which the computationally simplest estimator is generally the most efficient one. In the final section of this chapter, we examined the special properties of dynamic models. An important consideration in this analysis was the stability of the equations. Modern macroeconometrics involves many models in which one or more roots of the dynamic system equal one, so that these models, in the simple autoregressive form are unstable. In terms of the analysis in Section 13.9.3, in such a model, a shock to the system is permanent—the effects do not die out. We will examine a model of monetary policy with these characteristics in Section 20.6.8.

Key Terms and Concepts

- Admissible
- Behavioral equation
- Causality
- Complete system
- Completeness condition
- Consistent estimators
- Cumulative multiplier
- Dominant root
- Dynamic model
- Dynamic multiplier
- Econometric model
- Endogenous
- Equilibrium condition
- Equilibrium multipliers
- Exactly identified model
- Exclusion restrictions
- Exogenous
- FIML
- Final form
- Full information
- Full information maximum likelihood
- Fully recursive model
- Granger causality
- Identification
- Impact multiplier
- Impulse response function
- Indirect least squares
- Initial conditions
- Instrumental variable estimator
- Interdependent
- Jointly dependent
- *k* class
- Least variance ratio
- Limited information
- LIML
- Nonlinear systems
- Nonsample information
- Nonstructural
- Normalization
- Observationally equivalent
- Order condition
- Overidentification
- Predetermined variable
- Problem of identification
- Rank condition
- Recursive model
- Reduced form
- Reduced form disturbance
- Restrictions
- Simultaneous equations bias
- Specification test
- Stability
- Strongly exogenous
- Structural disturbance
- Structural equation
- Structural form
- System methods of estimation
- Three-stage least squares
- Triangular system
- Two-stage least squares
- Underidentified
- Weak instruments
- Weakly exogenous

Exercises

8. ~~1~~. Consider the following two-equation model:

$$y_1 = \gamma_1 y_2 + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3 + \varepsilon_1,$$
$$y_2 = \gamma_2 y_1 + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3 + \varepsilon_2.$$

10 - 66

a. Verify that, as stated, neither equation is identified.
b. Establish whether or not the following restrictions are sufficient to identify (or partially identify) the model:
(1) $\beta_{21} = \beta_{32} = 0$,
(2) $\beta_{12} = \beta_{22} = 0$,
(3) $\gamma_1 = 0$,
(4) $\gamma_1 = \gamma_2$ and $\beta_{32} = 0$,
(5) $\sigma_{12} = 0$ and $\beta_{31} = 0$,
(6) $\gamma_1 = 0$ and $\sigma_{12} = 0$,
(7) $\beta_{21} + \beta_{22} = 1$,
(8) $\sigma_{12} = 0, \beta_{21} = \beta_{22} = \beta_{31} = \beta_{32} = 0$.
(9) $\sigma_{12} = 0, \beta_{11} = \beta_{21} = \beta_{22} = \beta_{31} = \beta_{32} = 0$.

~~2. Verify the rank and order conditions for identification of the second and third behavioral equations in Klein's Model I.~~

3. Check the identifiability of the parameters of the following model:

$$[y_1 \quad y_2 \quad y_3 \quad y_4]\begin{bmatrix} 1 & \gamma_{12} & 0 & 0 \\ \gamma_{21} & 1 & \gamma_{23} & \gamma_{24} \\ 0 & \gamma_{32} & 1 & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & 0 & 1 \end{bmatrix}$$

$$+[x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5]\begin{bmatrix} 0 & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & 1 & 0 & \beta_{24} \\ \beta_{31} & \beta_{32} & \beta_{33} & 0 \\ 0 & 0 & \beta_{43} & \beta_{44} \\ 0 & \beta_{52} & 0 & 0 \end{bmatrix} = [\varepsilon_1 \quad \varepsilon_2 \quad \varepsilon_3 \quad \varepsilon_4].$$

9 ~~4.~~ Obtain the reduced form for the model in Exercise ~~1~~ 8 under each of the assumptions made in parts a and in parts b1 and b9.

10 ~~5.~~ The following model is specified:

$$y_1 = \gamma_1 y_2 + \beta_{11} x_1 + \varepsilon_1,$$
$$y_2 = \gamma_2 y_1 + \beta_{22} x_2 + \beta_{32} x_3 + \varepsilon_2.$$

All variables are measured as deviations from their means. The sample of 25 observations produces the following matrix of sums of squares and cross products:

$$\begin{array}{c} \\ y_1 \\ y_2 \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccccc} y_1 & y_2 & x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} 20 & 6 & 4 & 3 & 5 \\ 6 & 10 & 3 & 6 & 7 \\ 4 & 3 & 5 & 2 & 3 \\ 3 & 6 & 2 & 10 & 8 \\ 5 & 7 & 3 & 8 & 15 \end{bmatrix} \end{array}.$$

10-67

a. Estimate the two equations by OLS.
b. Estimate the parameters of the two equations by 2SLS. Also estimate the asymptotic covariance matrix of the 2SLS estimates.
c. Obtain the LIML estimates of the parameters of the first equation.
d. Estimate the two equations by 3SLS.
e. Estimate the reduced form coefficient matrix by OLS and indirectly by using your structural estimates from part b.

11 ~~6.~~ For the model

$$y_1 = \gamma_1 y_2 + \beta_{11} x_1 + \beta_{21} x_2 + \varepsilon_1,$$

$$y_2 = \gamma_2 y_1 + \beta_{32} x_3 + \beta_{42} x_4 + \varepsilon_2,$$

show that there are two restrictions on the reduced form coefficients. Describe a procedure for estimating the model while incorporating the restrictions.

7. An updated version of Klein's Model I was estimated. The relevant submatrix of $\mathbf{\Delta}$ is

$$\mathbf{\Delta}_1 = \begin{bmatrix} -0.1899 & -0.9471 & -0.8991 \\ 0 & 0.9287 & 0 \\ -0.0656 & -0.0791 & 0.0952 \end{bmatrix}.$$

Is the model stable?

12 ~~8.~~ Prove that

$$\operatorname{plim} \frac{\mathbf{Y}_j' \boldsymbol{\varepsilon}_j}{T} = \boldsymbol{\omega}_j - \mathbf{\Omega}_{jj} \boldsymbol{\gamma}_j.$$

13 ~~9.~~ Prove that an underidentified equation cannot be estimated by 2SLS.

Application

2

3. The data in Appendix Table F5.1 may be used to estimate a small macroeconomic model. Use these data to estimate the model in Example 13.1. Estimate the parameters of the two equations by two-stage and three-stage least squares. Then, using the two-stage least squares results, examine the dynamic properties of the model. Using the results in Section 13.9, determine if the dominant root of the system is less than one.

10-68

Applications

1. Continuing the analysis of Section 10.~~4~~5.2, we find that a translog cost function for one output and three factor inputs that does not impose constant returns to scale is

$$\begin{aligned}\ln C = \alpha &+ \beta_1 \ln p_1 + \beta_2 \ln p_2 + \beta_3 \ln p_3 + \delta_{11}\tfrac{1}{2}\ln^2 p_1 + \delta_{12}\ln p_1 \ln p_2 \\ &+ \delta_{13}\ln p_1 \ln p_3 + \delta_{22}\tfrac{1}{2}\ln^2 p_2 + \delta_{23}\ln p_2 \ln p_3 + \delta_{33}\tfrac{1}{2}\ln^2 p_3 \\ &+ \gamma_{q1}\ln Q \ln p_1 + \gamma_{q2}\ln Q \ln p_2 + \gamma_{q3}\ln Q \ln p_3 \\ &+ \beta_q \ln Q + \beta_{qq}\tfrac{1}{2}\ln^2 Q + \varepsilon_c.\end{aligned}$$

The factor share equations are

$$\begin{aligned} S_1 &= \beta_1 + \delta_{11}\ln p_1 + \delta_{12}\ln p_2 + \delta_{13}\ln p_3 + \gamma_{q1}\ln Q + \varepsilon_1, \\ S_2 &= \beta_2 + \delta_{12}\ln p_1 + \delta_{22}\ln p_2 + \delta_{23}\ln p_3 + \gamma_{q2}\ln Q + \varepsilon_2, \\ S_3 &= \beta_3 + \delta_{13}\ln p_1 + \delta_{23}\ln p_2 + \delta_{33}\ln p_3 + \gamma_{q3}\ln Q + \varepsilon_3.\end{aligned}$$

[See Christensen and Greene (1976) for analysis of this model.]

a. The three factor shares must add identically to 1. What restrictions does this requirement place on the model parameters?

b. Show that the adding-up condition in (10-38) can be imposed directly on the model by specifying the translog model in (C/p_3), (p_1/p_3), and (p_2/p_3) and dropping the third share equation. (See Example ~~10.8~~.) Notice that this reduces the number of free parameters in the model to 10. 10.3

c. Continuing part b, the model as specified with the symmetry and equality restrictions has 15 parameters. By imposing the constraints, you reduce this number to 10 in the estimating equations. How would you obtain estimates of the parameters not estimated directly?

The remaining parts of this exercise will require specialized software. The **E-Views, TSP, Stata** or **LIMDEP**, programs noted in the preface are four that could be used. All estimation is to be done using the data used in Section ~~10.4.1~~ 10.5.1

d. Estimate each of the three equations you obtained in part b by ordinary least squares. Do the estimates appear to satisfy the cross-equation equality and symmetry restrictions implied by the theory?

10.5.1 e. Using the data in Section ~~10.4.1~~, estimate the full system of three equations (cost and the two independent shares), imposing the symmetry and cross-equation equality constraints.

f. Using your parameter estimates, compute the estimates of the elasticities in (10-39) at the means of the variables.

g. Use a likelihood ratio statistic to test the joint hypothesis that $\gamma_{qi} = 0$, $i = 1, 2, 3$. [Hint: Just drop the relevant variables from the model.] 10.3

2. The Grunfeld investment data in Appendix Table ~~F9.3 are a~~ classic data set that have been used for decades to develop and demonstrate estimators for seemingly unrelated regressions.[~~34~~ 48 50] Although somewhat dated at this juncture, they remain an ideal application of the techniques presented in this chapter.[~~35~~ 49 51] The data consist of

FN 50 FN 51

48 50 [~~34~~] See Grunfeld (1958), Grunfeld and Griliches (1960), and Boot and de Witt (1960).

49 51 [~~35~~] See, in particular, Zellner (1962, 1963) and Zellner and Huang (1962).

10-69
End 10

time series of 20 yearly observations on ten firms. The three variables are

I_{it} = gross investment,
F_{it} = market value of the firm at the end of the previous year,
C_{it} = value of the stock of plant and equipment at the end of the previous year.

The main equation in the studies noted is

FN 52

$$I_{it} = \beta_1 + \beta_2 F_{it} + \beta_3 C_{it} + \varepsilon_{it}.$$ [50]

a. Fit the ten equations separately by ordinary least squares and report your results.
b. Use a Wald (Chow) test to test the "aggregation" restriction that the ten coefficient vectors are the same.
c. Use the seemingly unrelated regressions (FGLS) estimator to reestimate the parameters of the model, once again, allowing the coefficients to differ across the ten equations. Now, use the pooled model and, again, FGLS to estimate the constrained equation with equal parameter vectors, and test the aggregation hypothesis.
d. Using the OLS residuals from the separate regression, use the LM statistic in (10-17) to test for the presence of cross equation correlation.
e. An alternative specification to the model in part c. that focuses on the variances rather than the means is a groupwise heteroscedasticity model. For the current application, you can fit this model using (10-19), (10-20), and (10-21) while imposing the much simpler model with $\sigma_{ij} = 0$ when $i \neq j$. Do the results of the pooled model differ in the three cases considered, simple OLS, groupwise heteroscedasticity and full unrestricted covariances [which would be (10-20)] with $\mathbf{\Omega}_{ij} = \mathbf{I}$?

3. The data in AppendixTable F5.2 may be used to estimate a small macroeconomic model. Use these data to estimate the model in Example 10.4. Estimate the parameters of the two equations by two stage and three-stage least squares.

[50] Note that the model specifies investment, a flow, as a function of two stocks. This could be a theoretical misspecification. It might be preferable to specify the model in terms of planned investment. But, 50 years after the fact, we'll take the specified model as it is.