Chapter 11

Nonlinear Regression Models

Exercises

1. We cannot simply take logs of both sides of the equation as the disturbance is additive rather than multiplicative. So, we must treat the model as a nonlinear regression. The linearized equation is

$$y \approx \alpha^0 x^{\beta^0} + x^{\beta^0}(\alpha - \alpha^0) + \alpha^0(\log x)x^{\beta^0}(\beta - \beta^0)$$

where α^0 and β^0 are the expansion point. For given values of α^0 and β^0, the estimating equation would be

$$y - \alpha^0 x^{\beta^0} + \alpha^0 x^{\beta^0} + \alpha^0(\log x)x^{\beta^0} = \alpha\left(x^{\beta^0}\right) + \beta\left(\alpha^0(\log x)x^{\beta^0}\right) + \varepsilon^*$$

or

$$y + \alpha^0(\log x)x^{\beta^0} = \alpha\left(x^{\beta^0}\right) + \beta\left(\alpha^0(\log x)x^{\beta^0}\right) + \varepsilon^*.$$

Estimates of α and β are obtained by applying ordinary least squares to this equation. The process is repeated with the new estimates in the role of α^0 and β^0. The iteration could be continued until convergence. Starting values are always a problem. If one has no particular values in mind, one candidate would be $\alpha^0 = \bar{y}$ and $\beta^0 = 0$ or $\beta^0 = 1$ and α^0 either $\mathbf{x}'\mathbf{y}/\mathbf{x}'\mathbf{x}$ or $\bar{y}/\bar{x}$. Alternatively, one could search directly for the α and β to minimize the sum of squares, $S(\alpha,\beta) = \Sigma_i (y_i - \alpha x^\beta)^2 = \Sigma_i \varepsilon_i^2$. The first order conditions for minimization are

$$\partial S(\alpha,\beta)/\partial\alpha = -2\Sigma_i (y_i - \alpha x^\beta)x^\beta = 0 \quad \text{and} \quad \partial S(\alpha,\beta)/\partial\beta = -2\Sigma_i (y_i - \alpha x^\beta)\alpha(\ln x)x^\beta = 0.$$

Methods for solving nonlinear equations such as these are discussed in Appendix E..

2. The proof can be done by mathematical induction. For convenience, denote the *i*th derivative by f_i. The first derivative appears in Equation (10-34). Just by plugging in i=1, it is clear that f_1 satisfies the relationship. Now, use the chain rule to differentiate f_1,

$$f_2 = (-1/\lambda^2)[x^\lambda(\ln x) - x^{(\lambda)}] + (1/\lambda)[(\ln x)x^\lambda(\ln x) - f_1]$$

Collect terms to yield $\quad f_2 = (-1/\lambda)f_1 + (1/\lambda)[x^\lambda(\ln x)^2 - f_1] = (1/\lambda)[x^\lambda(\ln x)^2 - 2f_1]$.

So, the relationship holds for i = 0, 1, and 2. We now assume that it holds for $i = K$-1, and show that if so, it also holds for $i = K$. This will complete the proof. Thus, assume

$$f_{K\text{-}1} = (1/\lambda)[x^\lambda(\ln x)^{K-1} - (K\text{-}1)f_{K\text{-}2}]$$

Differentiate this to give $\quad f_K = (-1/\lambda)f_{K\text{-}1} + (1/\lambda)[(\ln x)x^\lambda(\ln x)^{K-1} - (K\text{-}1)f_{K\text{-}1}]$.

Collect terms to give $\quad f_K = (1/\lambda)[x^\lambda(\ln x)^K - Kf_{K\text{-}1}]$, which completes the proof for the general case.

Now, we take the limiting value

$$\lim_{\lambda\to 0} f_i = \lim_{\lambda\to 0}[x^\lambda(\ln x)^i - if_{i\text{-}1}]/\lambda.$$

Use L'Hospital's rule once again.

$$\lim_{\lambda\to 0} f_i = \lim_{\lambda\to 0} d\{[x^\lambda(\ln x)^i - if_{i\text{-}1}]/d\lambda\}/\lim_{\lambda\to 0} d\lambda/d\lambda.$$

Then, $\quad \lim_{\lambda\to 0} f_i = \lim_{\lambda\to 0}\{[x^\lambda(\ln x)^{i+1} - if_i]\}$

Just collect terms, $\quad (i+1)\lim_{\lambda\to 0} f_i = \lim_{\lambda\to 0}[x^\lambda(\ln x)^{i+1}]$

or $\quad \lim_{\lambda\to 0} f_i = \lim_{\lambda\to 0}[x^\lambda(\ln x)^{i+1}]/(i+1) = (\ln x)^{i+1}/(i+1)$.

Applications

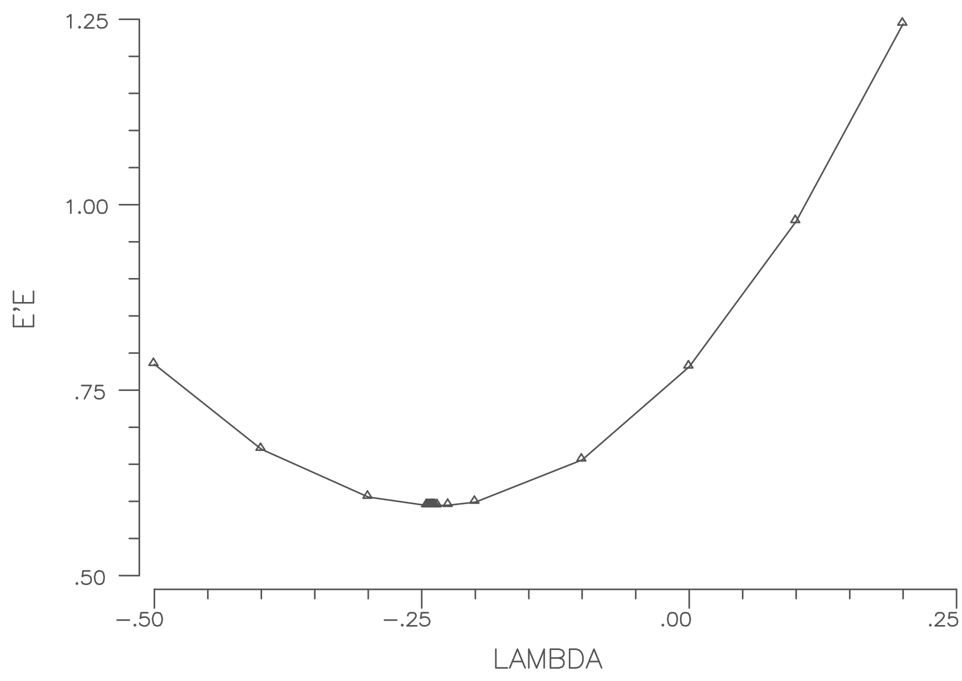
1. First, the two simple regressions produce

	Linear	**Log-linear**
Constant	114.338	1.17064
	(173.4)	(.3268)
Labor	2.33814	.602999
	(1.039)	(.1260)
Capital	.471043	.37571
	(.1124)	(.08535)
R^2	.9598	.9435
Standard Error	469.86	.1884

In the regression of *Y* on 1, *K*, *L*, and the predicted values from the loglinear equation minus the predictions from the linear equation, the coefficient on α is -587.349 with an estimated standard error of 3135. Since this is not significantly different from zero, this evidence favors the linear model. In the regression of ln*Y* on 1, ln*K*, ln*L* and the predictions from the linear model minus the exponent of the predictions from the loglinear model, the estimate of α is .000355 with a standard error of .000275. Therefore, this contradicts the preceding result and favors the loglinear model. An alternative approach is to fit the Box-Cox model in the fashion of Exercise 4. The maximum likelihood estimate of λ is about -.12, which is much closer to the log-linear model than the lonear one. The log-likelihoods are -192.5107 at the MLE, -192.6266 at λ=0 and -202.837 at λ = 1. Thus, the hypothesis that λ = 0 (the log-linear model) would not be rejected but the hypothesis that λ = 1 (the linear model) would be rejected using the Box-Cox model as a framework. □

2. The search for the minimum sum of squares produced the following results:

λ	**e′e**
-.500	.78477
-.400	.67033
-.300	.60587
-.250	.59479
-.245	.59451
-.244	.59447
-.243	.59444
-.242	.59441
-.241	.59439
-.240	.59438
-.239	.59437
-.238	.59436
-.237	.59437
-.235	.59440
-.225	.59492
-.200	.59897
-.100	.65598
0.000	.78143
.100	.97742
.200	1.24354



The sum of squared residuals is minimized at λ = -.238. At this value, the regression results are as follows:

Parameter	Estimate	OLS Std.Error	Correct Std.Error
α	2.06092	.07718	.09723
β_k	.178232	.04638	.04378
β_l	.737988	.06996	.12560
λ	-.238	----	.07710

Estimated Asymptotic Covariance Matrix

	α	β_k	β_l	λ
α	.00945			
β_k	.00262	.00192		
β_l	.00511	-.00199	.01578	
λ	.00500	.00037	.00825	.00594

The output elasticities for this function evaluated at the sample means are

$$\partial \ln Y/\partial \ln K = \beta_k K^{\lambda} = (.178232).175905^{-.238} = .2695$$

$$\partial \ln Y/\partial \ln L = \beta_l L^{\lambda} = (.443954).737988^{-.238} = .7740.$$

The estimates found for Zellner and Revankar's model were .254 and .882, respectively, so these are quite similar. For the simple log-linear model, the corresponding values are .2790 and .927. □

3. The Wald test is based on the unrestricted model. The statistic is the square of the usual t-ratio, W = $(-.232 / .0771)^2$ = 9.0546. The critical value from the chi-squared distribution is 3.84, so the hypothesis that λ = 0 can be rejected. The likelihood ratio statistic is based on both models. The sum of squared residuals for both unrestricted and restricted models is given above. The log-likelihood is $\ln L = -(n/2)[1 + \ln(2\pi) + \ln(\mathbf{e}'\mathbf{e}/n)]$, so the likelihood ratio statistic is

$$LR = n[\ln(\mathbf{e}'\mathbf{e}/n)|\lambda{=}0 - \ln(\mathbf{e}'\mathbf{e}/n)|\lambda{=}-.238] = n\ln[(\mathbf{e}'\mathbf{e}|\lambda{=}0) / (\mathbf{e}'\mathbf{e}|\lambda{=}-.238)$$
$$= 25\ln(.78143/.54369) = 6.8406.$$

Finally, to compute the Lagrange Multiplier statistic, we regress the residuals from the log-linear regression on a constant, ln$K$, ln$L$, and $(1/2)(b_k \ln^2 K + b_l \ln^2 L)$ where the coefficients are those from the log-linear model (.27898 and .92731). The R^2 in this regression is .23001, so the Lagrange multiplier statistic is $LM = nR^2$ = 25(.23001) = 5.7503. All three statistics suggest the same conclusion, the hypothesis should be rejected. □

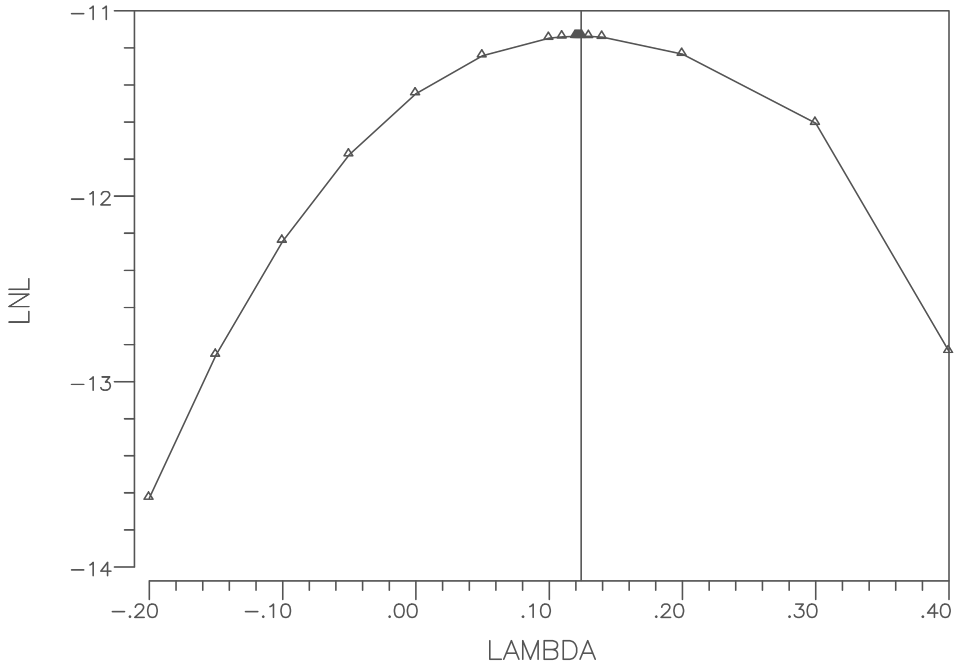
4. Instead of minimizing the sum of squared deviations, we now maximize the concentrated log-likelihood function, $\ln L = -(n/2)\ln(1+\ln(2\pi)) + (\lambda - 1)\Sigma_i \ln Y_i - (n/2)\ln(\varepsilon'\varepsilon/n)$.
The search for the maximum of ln$L$ produced the results on the next page
The log-likelihood is maximized at λ = .124. At this value, the regression results are as follows:

Parameter	Estimate	OLS Std.Error	Correct Std.Error
α	2.59465	.1283	.7151
β_k	.378094	.1070	.3228
β_l	1.13653	.1117	.4121
λ	.124	----	.2482
σ^2	.036922	----	.0179

Estimated Asymptotic Covariance Matrix

	α	β_k	β_l	λ	σ^2
α	.5114				
β_k	.2203	.1042			
β_l	.2612	.0951	.1698		
λ	.1747	.0730	.0953	.0617	
σ^2	.0104	.0044	.0059	.0038	.00032

λ	ln$L$
-.200	-13.6284
-.150	-12.8568
-.100	-12.2423
-.050	-11.7764
0.000	-11.4476
.050	-11.2427
.100	-11.1480
.110	-11.1410
.120	-11.1378
.121	-11.1377
.122	-11.1376
.123	-11.1376
.124	-11.1375
.125	-11.1376
.130	-11.1383
.140	-11.1423
.200	-11.2344
.300	-11.6064
.400	-12.8371



The output elasticities for this function evaluated at the sample means, $\overline{K}$ = .175905, $\overline{L}$ = .737988, $\overline{Y}$ = 2.870777, are $\partial \ln Y/\partial \ln K = b_k(K/Y)^{\lambda} = .2674$

$\partial \ln Y/\partial \ln L = b_l(L/Y)^{\lambda} = .9017$.

These are quite similar to the estimates given above. The sum of the two output elasticities for the states given in the example in the text are given below for the model estimated with and without transforming the dependent variable. Note that the first of these makes the model look much more similar to the Cobb Douglas model for which this sum is constant.

State	**Full Box-Cox Model**	**lnQ on left hand side**
Florida	1.2840	1.6598
Louisiana	1.2019	1.4239
California	1.1574	1.1176
Maryland	1.1657	1.0261
Ohio	1.1899	.9080
Michigan	1.1604	.8506

Once again, we are interested in testing the hypothesis that $\lambda = 0$. The Wald test statistic is $W = (.123 / .2482)^2 = .2455$. We would now not reject the hypothesis that $\lambda = 0$. This is a surprising outcome. The likelihood ratio statistic is based on both models. The sum of squared residuals for the restricted model is given above. The sum of the logs of the outputs is 19.29336, so the restricted log-likelihood is $\ln L^0 = (0-1)(19.29336) - (25/2)[1 + \ln(2\pi) + \ln(.781403/25)] = -11.44757$. The likelihood ratio statistic is $-2[-11.13758 - (-11.44757)] = .61998$. Once again, the statistic is small. Finally, to compute the Lagrange multiplier statistic, we now use the method described in Example 11.8. The result is $LM = 1.5621$. All of these suggest that the log-linear model is not a significant restriction on the Box-Cox model. This rather peculiar outcome would appear to arise because of the rather substantial reduction in the log-likelihood function which occurs when the dependent variable is transformed along with the right hand side. This is not a contradiction because the model with only the right hand side transformed is not a parametric restriction on the model with both sides transformed. Some further evidence is given in the next exercise.

5. **--> nlsq ; lhs = y ; labels = b1,b2 ; fcn=b1*(1 - 1/sqr(1+2*b2*x)) ; start = 500,.0001 ;output=2$**

```
Begin NLSQ iterations. Linearized regression.
Iteration=  1; Sum of squares=  11603.0164    ; Gradient=  11602.9326
Iteration=  2; Sum of squares=  19821.5463    ; Gradient=  19821.4534
Iteration=  3; Sum of squares=  331169.005    ; Gradient=  331144.576
Iteration=  4; Sum of squares=  356630.271    ; Gradient=  356504.582
Iteration=  5; Sum of squares=  14997.8506    ; Gradient=  14938.8590
Iteration=  6; Sum of squares=  449.855530    ; Gradient=  442.701921
Iteration=  7; Sum of squares=  102026.884    ; Gradient=  102026.775
Iteration=  8; Sum of squares=  12887.7536    ; Gradient=  12886.6539
Iteration=  9; Sum of squares=  14263101.5    ; Gradient=  14263101.0
Iteration= 10; Sum of squares=  10203.1920    ; Gradient=  10202.6789
Iteration= 11; Sum of squares=  144.393444    ; Gradient=  144.338425
Iteration= 12; Sum of squares=  258.186688    ; Gradient=  258.145522
Iteration= 13; Sum of squares=  .154284512    ; Gradient=  .113316151
Iteration= 14; Sum of squares=  .409681292E-01; Gradient=  .129216769E-05
Iteration= 15; Sum of squares=  .409668370E-01; Gradient=  .439070450E-13
Iteration= 16; Sum of squares=  .409668370E-01; Gradient=  .211594637E-18
Iteration= 17; Sum of squares=  .409668370E-01; Gradient=  .107898463E-24
Convergence achieved
+-----------------------------------------------------+
| Nonlinear   least squares regression                |
| LHS=Y        Mean                  =   43.34071     |
|              Standard deviation    =   22.80652     |
| WTS=none     Number of observs.    =         14     |
| Model size   Parameters            =          2     |
|              Degrees of freedom    =         12     |
| Residuals    Sum of squares        =   .4096684E-01 |
|              Standard error of e   =   .5409439E-01 |
| Fit          R-squared             =   .9999939     |
| Not using OLS or no constant. Rsqd & F may be < 0. |
+-----------------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 B1      |    636.427250      4.31789336   147.393   .0000
 B2      |     .00020814     .164134D-05   126.809   .0000
```

--> nlsq ; lhs = y ; labels = b1,b2 ; fcn=b1*(1 - 1/sqr(1+2*b2*x)) ; start = 600,.0002 ;output=2$

```
Begin NLSQ iterations. Linearized regression.
Iteration=  1; Sum of squares=  262.456583    ; Gradient=  262.415454
Iteration=  2; Sum of squares=  .155984704    ; Gradient=  .115016579
Iteration=  3; Sum of squares=  .409675977E-01; Gradient=  .760690867E-06
Iteration=  4; Sum of squares=  .409668370E-01; Gradient=  .379981726E-13
Iteration=  5; Sum of squares=  .409668370E-01; Gradient=  .186919870E-18
Iteration=  6; Sum of squares=  .409668370E-01; Gradient=  .150578559E-23
Convergence achieved
+-----------------------------------------------------+
| Nonlinear   least squares regression                |
| LHS=Y        Mean                  =   43.34071     |
|              Standard deviation    =   22.80652     |
| Residuals    Sum of squares        =   .4096684E-01 |
|              Standard error of e   =   .5409439E-01 |
| Fit          R-squared             =   .9999939     |
|              Adjusted R-squared    =   .9999944     |
+-----------------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
 B1      |    636.427250      4.31789336   147.393   .0000
 B2      |     .00020814     .164134D-05   126.809   .0000
```

Chapter 12

Instrumental Variables Estimation

Exercises

1. There is no need for a separate proof different from the usual for OLS. Formally, however, it follows from the results at (12-4) that

$$\mathbf{b} = \beta + \left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1}\left(\frac{\mathbf{X}'\varepsilon}{n}\right)$$

Then,

$$\mathbf{b} - \text{plim}\,\mathbf{b} = \left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1}\left(\frac{\mathbf{X}'\varepsilon}{n}\right) - \mathbf{Q}_{\mathrm{XX}}^{-1}\gamma$$

and

$$\sqrt{n}\left(\mathbf{b} - \text{plim}\,\mathbf{b}\right) = \sqrt{n}\left[\left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1}\left(\frac{\mathbf{X}'\varepsilon}{n}\right) - \mathbf{Q}_{\mathrm{XX}}^{-1}\gamma\right]$$

The large sample distribution of this statistic will be the same as the large sample of the statistic with X′X/n replaced with its probablity limit, which is $\mathbf{Q}_{\mathbf{XX}}$. Thus,

$$\sqrt{n}\left(\mathbf{b} - \text{plim}\,\mathbf{b}\right) \rightarrow \mathbf{Q}_{\mathrm{XX}}^{-1}\sqrt{n}\left[\left(\frac{\mathbf{X}'\varepsilon}{n}\right) - \gamma\right]$$

To deduce the large sample behavior of this statistic, we can invoke the results from chapter 4. The only change here is the nonzero mean (probability limit) of the vector in brackets. [See (12-3).] Thus, the same proof applies. The consistency, asymptotic normality and asymptotic covariance matrix equal to Asy.Var[b] = σ_ε^2 $(X'X)^{-1}$

2. A logical solution to this one is simple. For y and x*,

$\text{Cov}^2(y,x^*)/[\text{Var}(y)\text{Var}(x^*)] = \beta^2(\sigma_*^2)^2/[(\beta^2\sigma_*^2+\sigma_\varepsilon^2)(\sigma_*^2)]$

$\text{Cov}^2(y,x)$ /[Var(y)Var(x)] = Cov[βx*+ε,x*+u] / [Var(y)Var(x)]

= $\{\text{Cov}[y,x^*] + \text{Cov}[y,u]\}^2$ / [Var(y)Var(x)] .

The second term is zero, since y=βx*+ε which is uncorrelated with u. Thus,

$\text{Cov}^2(y,x)$ /[Var(y)Var(x)] = Cov[y,x*] / [Var(y)Var(x)].

The numerator is the same. The denominator is larger, since [Var(y)Var(x)] = Var[y](Var[x*] + Var[u]), so the squared correlation must be smaller. If both variables are measured with errors, then we are comparing $\text{Cov}^2(y^*,x^*)$/{Var[y*]Var[x*]} to $\text{Cov}^2(y,x)$/{Var[y]Var[x]}.

The numerator is the covariance of (βx* + ε + v) with (x* + u), so the numerator of the fraction is still $\beta^2(\sigma_*^2)^2$. The denominator is still obviously larger, so the same result holds when both variables are measured with error.

3. We work off (12-16), using repeatedly the result $\Sigma_{uu} = (\sigma_u j)(\sigma_u j)'$ where j has a 1 in the first position and 0 in the remaining K-1. From (12-16),

plim b = β - $[Q^* + \Sigma_{uu}]^{-1}\Sigma_{uu}\beta$. The vector is $\Sigma_{uu}\beta$ equals $[\sigma_u^2\beta_1,0,...,0]'$. The inverse matrix is

$$[Q^* + \Sigma_{uu}]^{-1} = \left[\left(\mathbf{Q}^*\right)^{-1} - \frac{1}{1+(\sigma_u\mathbf{j})'\left(\mathbf{Q}^*\right)^{-1}(\sigma_u\mathbf{j})}\left(\mathbf{Q}^*\right)^{-1}(\sigma_u\mathbf{j})(\sigma_u\mathbf{j})'\left(\mathbf{Q}^*\right)^{-1}\right]$$

This can be simplified since the quadratic form in the denominator just picks off the 1,1 diagonal element. Thus,

$$[\mathbf{Q}^* + \boldsymbol{\Sigma}_{uu}]^{-1} = \left[(\mathbf{Q}^*)^{-1} - \frac{1}{1+\sigma_u^2 q^{*11}}(\mathbf{Q}^*)^{-1}(\sigma_u \mathbf{j})(\sigma_u \mathbf{j})'(\mathbf{Q}^*)^{-1}\right]$$

Then

$$[\mathbf{Q}^* + \boldsymbol{\Sigma}_{uu}]^{-1}\boldsymbol{\Sigma}_{uu}\boldsymbol{\beta} = \left[(\mathbf{Q}^*)^{-1} - \frac{1}{1+\sigma_u^2 q^{*11}}(\mathbf{Q}^*)^{-1}(\sigma_u \mathbf{j})(\sigma_u \mathbf{j})'(\mathbf{Q}^*)^{-1}\right](\sigma_u \mathbf{j})(\sigma_u \mathbf{j})'\beta$$

$$= (\mathbf{Q}^*)^{-1}(\sigma_u \mathbf{j})(\sigma_u \mathbf{j})'\boldsymbol{\beta} - \frac{1}{1+\sigma_u^2 q^{*11}}(\mathbf{Q}^*)^{-1}(\sigma_u \mathbf{j})(\sigma_u \mathbf{j})'(\mathbf{Q}^*)^{-1}(\sigma_u \mathbf{j})(\sigma_u \mathbf{j})'\beta$$

$$= (\mathbf{Q}^*)^{-1}\mathbf{j}\,\sigma_u^2\beta_1 - \frac{\sigma_u^2 q^{*11}}{1+\sigma_u^2 q^{*11}}(\mathbf{Q}^*)^{-1}\mathbf{j}\,\sigma_u^2\beta_1$$

$$= (\mathbf{Q}^*)^{-1}\mathbf{j}\left[1-\frac{\sigma_u^2 q^{*11}}{1+\sigma_u^2 q^{*11}}\right]\sigma_u^2\beta_1$$

$$= (\mathbf{Q}^*)^{-1}\mathbf{j}\left[\frac{1}{1+\sigma_u^2 q^{*11}}\right]\sigma_u^2\beta_1$$

$$= (\mathbf{Q}^*)^{-1}\mathbf{j}\left[\frac{\sigma_u^2\beta_1}{1+\sigma_u^2 q^{*11}}\right]$$

Finally, $(\mathbf{Q}^*)^{-1}\mathbf{j}$ equals the first column of $(\mathbf{Q}^*)^{-1}$ = $[q^{*11}, q^{*21},\ldots,q^{*k1}]$. Therefore, the first element, given by (12-17a) is

$$\text{plim } b_1 = \beta_1 - \left[\frac{\sigma_u^2\beta_1}{1+\sigma_u^2 q^{*11}}\right]q^{*11} = \beta_1\left[1-\frac{\sigma_u^2 q^{*11}}{1+\sigma_u^2 q^{*11}}\right]$$

For (12-17b),

$$\text{plim } b_2 = \beta_2 - \left[\frac{\sigma_u^2\beta_1}{1+\sigma_u^2 q^{*11}}\right]q^{*k1}$$

4. To obtain the result, note first:

$\text{plim } \mathbf{b} = \boldsymbol{\beta} + \mathbf{Q}_{XX}^{-1}\boldsymbol{\gamma}$

$\text{Asy.Var}[\mathbf{b}] = (\sigma^2/n)\mathbf{Q}_{XX}^{-1}$

$\text{Asy.Var}[\mathbf{b}_{2sls}] = (\sigma^2/n)\mathbf{Q}_{ZX}^{-1}\mathbf{Q}_{ZZ}\mathbf{Q}_{XZ}^{-1}$.

The mean squared error of the OLS estimator is the variance plus the squared bias,

$$M(b|\beta) = (\sigma^2/n)\mathbf{Q}_{XX}^{-1} + \mathbf{Q}_{XX}^{-1}\gamma\gamma'\mathbf{Q}_{XX}^{-1}$$

the mean squared error of the 2SLS estimator equals its variance. For OLS to be more precise then 2SLS, we would have to have

$$(\sigma^2/n)\mathbf{Q}_{XX}^{-1} + \mathbf{Q}_{XX}^{-1}\gamma\gamma'\mathbf{Q}_{XX}^{-1} << (\sigma^2/n)\mathbf{Q}_{ZX}^{-1}\mathbf{Q}_{ZZ}\mathbf{Q}_{XZ}^{-1}.$$

For convenience, let $\delta = \mathbf{Q}_{XX}^{-1}\gamma$ so $M(b|\beta) = (\sigma^2/n)\mathbf{Q}_{XX}^{-1} + \delta\delta'$. If the mean squared error matrix of the OLS estimator is smaller than that of the 2SLS estimator, then its inverse is larger. Use (A-66) to do the inversion. The result would be

$$[(\sigma^2/n)\mathbf{Q}_{XX}^{-1} + \delta\delta']^{-1} >> [(\sigma^2/n)\mathbf{Q}_{ZX}^{-1}\mathbf{Q}_{ZZ}\mathbf{Q}_{XZ}^{-1}]^{-1}$$

Now, use A-66

$$[(\sigma^2/n)\mathbf{Q}_{XX}^{-1} + \delta\delta']^{-1} = (n/\sigma^2)\,\mathbf{Q}_{XX} - \frac{1}{1+\delta'(n/\sigma^2)\mathbf{Q}_{XX}\delta}(n/\sigma^2)\,\mathbf{Q}_{XX}\delta\delta'(n/\sigma^2)\,\mathbf{Q}_{XX}$$

Reinsert $\delta = \mathbf{Q}_{XX}^{-1}\gamma$ and the right hand side above reduces to

$$(n/\sigma^2)\,\mathbf{Q}_{XX} - \frac{1}{1+(n/\sigma^2)\gamma'\mathbf{Q}_{XX}^{-1}\gamma}(n/\sigma^2)^2\,\gamma\gamma'$$

Therefore, if the mean squared error matrix of OLS is smaller, then

$$(n/\sigma^2)\,\mathbf{Q}_{XX} - \frac{1}{1+(n/\sigma^2)\gamma'\mathbf{Q}_{XX}^{-1}\gamma}(n/\sigma^2)^2\,\gamma\gamma' >> (n/\sigma^2)\mathbf{Q}_{XZ}\mathbf{Q}_{ZZ}^{-1}\mathbf{Q}_{ZX}$$

Collect the terms, and this implies

$$(n/\sigma^2)[\,\mathbf{Q}_{XX} - \mathbf{Q}_{XZ}\mathbf{Q}_{ZZ}^{-1}\mathbf{Q}_{ZX}] >> \frac{1}{1+(n/\sigma^2)\gamma'\mathbf{Q}_{XX}^{-1}\gamma}(n/\sigma^2)^2\,\gamma\gamma'$$

divide both sides by (n/σ^2),

$$\mathbf{Q}_{XX} - \mathbf{Q}_{XZ}\mathbf{Q}_{ZZ}^{-1}\mathbf{Q}_{ZX} >> \frac{(n/\sigma^2)}{1+(n/\sigma^2)\gamma'\mathbf{Q}_{XX}^{-1}\gamma}\gamma\gamma'$$

and divide numerator and denominator of the fraction by n/σ^2

$$\mathbf{Q}_{XX} - \mathbf{Q}_{XZ}\mathbf{Q}_{ZZ}^{-1}\mathbf{Q}_{ZX} >> \frac{1}{(\sigma^2/n)+\gamma'\mathbf{Q}_{XX}^{-1}\gamma}\gamma\gamma'$$

which is the desired result. Is it possible? It is possible, since

$$\mathbf{Q}_{XX} - \mathbf{Q}_{XZ}\mathbf{Q}_{ZZ}^{-1}\mathbf{Q}_{ZX} = \text{plim}\,(1/n)[\mathbf{X}'\mathbf{X} - \mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}] = \text{plim}\,(1/n)\,\mathbf{X}'\mathbf{M}_Z\mathbf{X}$$

which is a positive definite matrix. SInce γ varies independently of $\mathbf{Z}$ and $\mathbf{X}$, certainly there is some configuration of the data and parameters for which this is the case. The result is that it is, indeed, possible for OLS to be more precise, in the mean squared error sense, than 2SLS.

5. The matrices are X = [i,x] and Z = [i,z]. For the OLS estimators, we know from chapter 2 that

$$a = \bar{y} - b\bar{x} \text{ and } b = \text{Cov}[x,y]/\text{var}[x].$$

For the IV estimator, $(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{y}$, we obtain the result in detail. Given the forms,

$$(\mathbf{Z}'\mathbf{X}) = \begin{bmatrix} n & \Sigma x_i \\ n_1 & \Sigma_{z=1} x_i \end{bmatrix} = \begin{bmatrix} n & n\bar{x} \\ n_1 & n_1\bar{x}_1 \end{bmatrix},\ (\mathbf{Z}'\mathbf{X})^{-1} = \frac{1}{nn_1(\bar{x}-\bar{x})}\begin{bmatrix} n_1\bar{x}_1 & -n\bar{x} \\ -n_1 & n \end{bmatrix},\ \mathbf{Z}'\mathbf{y} = \begin{bmatrix} n\bar{y} \\ n_1\bar{y}_1 \end{bmatrix}$$

where subscript 1 indicates the mean of the observations for which z equals 1, and n_1 is the number of observations. Multiplying the matrix times the vector and cancelling terms produces the solutions

$$a_{IV} = a_{IV} = \frac{\bar{x}_1\bar{y} - \bar{x}\,\bar{y}_1}{\bar{x}_1 - \bar{x}} \text{ and } b_{IV} = \frac{\bar{y}_1 - \bar{y}}{\bar{x}_1 - \bar{x}}$$

Application

a. The statement of the problem is actually a bit optimistic. GIven the way it is stated, it would imply that the exogenous variables in the "demand" equation would be, in principle, (Ed, Union, Fem) which are also in the supply equation, plus the remainder, (Exp, Exp^2, Occ, Ind, South, SMSA, Blk). The problem is that the model as stated would not be identified – the supply equation would, but the demand equation would not be. The way out would be to assume that at least one of (Ed, Union, Fem) does not appear in the demand equation. Since surely education would, that leaves one or both of Union and Fem. We will assume both of them are omitted. So, our equation is

$$\ln Wage_{it} = \alpha_1 + \alpha_2 Ed_{it} + \alpha_3 Exp_{it} + \alpha_4 Exp_{it}^2 + \alpha_5 Occ_{it} + \alpha_6 Ind_{it} + \alpha_7 South_{it} + \alpha_8 SMSA_{it} + \alpha_9 Blk_{it} + \gamma\, Wks_{it} + u_{it}.$$

```
NAMELIST ; X = one,Ed,Exp,Expsq,Occ,Ind,South,SMSA,Blk,Wks $
NAMELIST ; Z = one,Ed,Exp,expsq,Occ,Ind,south,SMSA,Blk,Union,Fem $
Regress  ; Lhs = lwage ; Rhs = X $
2SLS     ; Lhs = lwage ; Rhs = X ; Inst = Z $
REGRESS  ; Lhs = Wks ; Rhs = Z ; cls:b(10)=0,b(11)=0$
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=LWAGE    Mean                 =   6.676346     |
|              Standard deviation   =   .4615122     |
| WTS=none     Number of observs.   =       4165     |
| Model size   Parameters           =         10     |
|              Degrees of freedom   =       4155     |
| Residuals    Sum of squares       =   581.2717     |
|              Standard error of e  =   .3740280     |
| Fit          R-squared            =   .3446066     |
|              Adjusted R-squared   =   .3431870     |
| Model test   F[  9,  4155] (prob) = 242.74 (.0000) |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    5.13171052       .07238152    70.898   .0000
 ED      |     .06112766       .00277226    22.050   .0000   12.8453782
 EXP     |     .04291665       .00229783    18.677   .0000   19.8537815
 EXPSQ   |    -.00070803     .506204D-04   -13.987   .0000   514.405042
 OCC     |    -.07814434       .01502100    -5.202   .0000    .51116447
 IND     |     .09066812       .01247863     7.266   .0000    .39543818
 SOUTH   |    -.07629062       .01318346    -5.787   .0000    .29027611
 SMSA    |     .13789225       .01278553    10.785   .0000    .65378151
 BLK     |    -.26269494       .02304380   -11.400   .0000    .07226891
 WKS     |     .00484184       .00113470     4.267   .0000   46.8115246
+----------------------------------------------------+
| Two stage   least squares regression               |
| LHS=LWAGE    Mean                 =   6.676346     |
|              Standard deviation   =   .4615122     |
| WTS=none     Number of observs.   =       4165     |
| Model size   Parameters           =         10     |
|              Degrees of freedom   =       4155     |
| Residuals    Sum of squares       =   602.3138     |
|              Standard error of e  =   .3807377     |
| Fit          R-squared            =   .3192467     |
|              Adjusted R-squared   =   .3177722     |
| Model test   F[  9,  4155] (prob) = 216.50 (.0000) |
+----------------------------------------------------+
| Instrumental Variables:
|ONE      ED       EXP      EXPSQ    OCC      IND      SOUTH    SMSA
|BLK      UNION    FEM
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    4.46105888       .27680953    16.116   .0000
 ED      |     .06167266       .00283031    21.790   .0000   12.8453782
```

EXP	.04207640	.00236282	17.808	.0000	19.8537815
EXPSQ	-.00068241	.525268D-04	-12.992	.0000	514.405042
OCC	-.07605669	.01531301	-4.967	.0000	.51116447
IND	.08348143	.01302032	6.412	.0000	.39543818
SOUTH	-.08242895	.01364036	-6.043	.0000	.29027611
SMSA	.13244624	.01319402	10.038	.0000	.65378151
BLK	-.25212290	.02383132	-10.579	.0000	.07226891
WKS	.01922950	.00583960	3.293	.0010	46.8115246

This is the test of relevance of the instrumental variables. In the regression of WKS on the full set of exogenous variables, we test the hypothesis that the coefficients on the instruments, UNION and FEM are jointly zero. The results show that the hypothesis is rejected. We conclude that the instruments are relevant.

```
+-----------------------------------------------------+
| Linearly restricted regression                      |
| Ordinary    least squares regression                |
| LHS=WKS      Mean                 =   46.81152      |
|              Standard deviation   =   5.129098      |
| WTS=none     Number of observs.   =       4165      |
| Model size   Parameters           =          9      |
|              Degrees of freedom   =       4156      |
| Residuals    Sum of squares       =   108653.5      |
|              Standard error of e  =   5.113097      |
| Fit          R-squared            =   .8138966E-02   |
|              Adjusted R-squared   =   .6229705E-02   |
| Model test   F[  8,  4156] (prob) =   4.26 (.0000)  |
| Restrictns.  F[  2,  4154] (prob) =  84.57 (.0000)  |
| Not using OLS or no constant. Rsqd & F may be < 0.  |
| Note, with restrictions imposed,  Rsqd may be < 0.  |
+-----------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Constant	46.6129896	.67547781	69.007	.0000	
ED	-.03787988	.03789322	-1.000	.3175	12.8453782
EXP	.05840099	.03139904	1.860	.0629	19.8537815
EXPSQ	-.00178055	.00069145	-2.575	.0100	514.405042
OCC	-.14509978	.20533021	-.707	.4798	.51116447
IND	.49950389	.17041135	2.931	.0034	.39543818
SOUTH	.42663864	.18010107	2.369	.0178	.29027611
SMSA	.37851979	.17468415	2.167	.0302	.65378151
BLK	-.73479892	.31481083	-2.334	.0196	.07226891
UNION	.444089D-15	.182255D-08	.000	1.0000	.36398559
FEM	.000000	......(Fixed Parameter).......			

Chapter 13

Simultaneous Equations Models

1. (a) Since nothing is excluded from either equation and there are no other restrictions, neither equation passes the order condition for identification.

(1) We use (13-12) and the equations which follow it. For the first equation, $[\mathbf{A}_3',\mathbf{A}_5'] = \beta_{22}$, a scalar which has rank M-1 = 1 unless $\beta_{22} = 0$. For the second, $[\mathbf{A}_3',\mathbf{A}_5'] = \beta_{31}$. Thus, both equations are identified.

(2) This restriction does not restrict the first equation, so it remains unidentified. The second equation is now identified, as $[\mathbf{A}_3',\mathbf{A}_5'] = [\beta_{11},\beta_{21}]$ has rank 1 if either of the two ceofficients are nonzero.

(3) If γ_1 equals 0, the model becomes partially recursive. The first equation becomes a regression which can be estimated by ordinary least squares. However, the second equation continues to fail the order condition. To see the problem, consider that even with the restriction, any linear combination of the two equations has the same variables as the original second eqation.

(4) We know from above that if $\beta_{32} = 0$, the second equation is identifiable. If it is, then γ_2 is identified. We may treat it as known. As such, γ_1 is known. By regressing $\mathbf{y}_1 - \gamma_1\mathbf{y}_2$ on the $\mathbf{x}$s, we would obtain estimates of the remaining parameters, so these restrictions identify the model. It is instructive to analyze this from the standpoint of false structures as done in the text. A false structure which incorporates the known restrictions would be $\begin{bmatrix} 1 & -\gamma \\ -\lambda & 1 \\ \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & 0 \end{bmatrix} \times \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$. If the false structure is to obey the restrictions, then $f_{11} - \gamma f_{21} = 1$, $f_{22} - \gamma f_{12} = 1$, $f_{21} - \gamma f_{11} = f_{12} - \gamma f_{22}$, $\beta_{31} f_{12} = 0$. It follows then that $f_{12} = 0$ so $f_{11} = 1$. Then, $f_{21} - \gamma f_{11} = -\gamma$ or $f_{21} = (f_{11} - 1)\gamma$ so that $f_{11} - \gamma^2(f_{11} - 1) = 1$. This can only hold for all values of γ if $f_{11} = 1$ and, then, $f_{21} = 0$. Therefore, $\mathbf{F} = \mathbf{I}$ which establishes identification.

(5) If $\beta_{31} = 0$, the first equation is identified by the usual rank and order conditions. Consider, then, the off-diagonal element of $\mathbf{\Sigma} = \mathbf{\Gamma}'\mathbf{\Omega}\mathbf{\Gamma}$. $\mathbf{\Omega}$ is identified since it is the reduced form covariance matrix. The off-diagonal element is $\sigma_{12} = \omega_{11} + \omega_{22} - (\gamma_1 + \gamma_2)\omega_{12} = 0$. Since γ_1 is zero, $\gamma_2 = \omega_{12}/(\omega_{11} + \omega_{22})$. With γ_2 known, the remaining parameters are estimable by least squares regression of $(\mathbf{y}_2 - \gamma_2\mathbf{y}_1)$ on the $\mathbf{x}$s. Therefore, the restrictions identify the model.

(6) Since this is only a single restriction, it will not likely identify the entire model. Consider again the false structure. The restrictions implied by the theory are $f_{11} - \gamma_2 f_{21} = 1$, $f_{22} - \gamma_1 f_{12} = 1$, $\beta_{21}f_{11} + \beta_{22}f_{21} = \beta_{21}f_{12} + \beta_{22}f_{22}$. The three restrictions on four unknown elements of $\mathbf{F}$ do not serve to pin down any of them. This restriction does not even partially identify the model.

(7) The last four restrictions remove x_2 and x_3 from the model. The remaining model is not identified by the usual rank and order conditions. From part (5), we see that the first restriction implies $\sigma_{12} = \omega_{11} + \omega_{22} - (\gamma_1 + \gamma_2)\omega_{12} = 0$. But, with neither γ_1 nor γ_2 specified, this does not identify either parameter.

(8) The first equation is identified by the conventional rank and order conditions. The second equation fails the order condition. But, the restriction $\sigma_{12} = 0$ provides the necessary additional information needed to identify the model. For simplicity, write the model with the restrictions imposed as

$$y_1 = \gamma_1 y_2 + \varepsilon_1 \text{ and } y_2 = \gamma_2 y_1 + \beta x + \varepsilon_2.$$

The reduced form is $y_1 = \pi_1 x + v_1$ and $y_2 = \pi_2 x + v_2$

where $\pi_1 = \gamma_1\beta/\Delta$ and $\pi_2 = \beta/\Delta$ with $\Delta = (1 - \gamma_1\gamma_2)$, and $v_1 = (\varepsilon_1 + \gamma_1\varepsilon_2)/\Delta$ and $v_2 = (\varepsilon_2 + \gamma_2\varepsilon_1)/\Delta$. The reduced form variances and covariances are $\omega_{11} = (\gamma_1^2\sigma_{22} + \sigma_{11})/\Delta^2$, $\omega_{22} = (\gamma_2^2\sigma_{11} + \sigma_{22})/\Delta^2$, $\omega_{12} = (\gamma_1\sigma_{22} + \gamma_2\sigma_{11})/\Delta^2$. All reduced form parameters are estimable directly by using least squares, so the reduced form is identified in all cases. Now, $\gamma_1 = \pi_1/\pi_2$. σ_{11} is the residual variance in the euqation $(y_1 - \gamma_1 y_2) = \varepsilon_1$, so σ_{11} must be estimable (identified) if γ_1 is. Now, with a bit of manipulation, we find that $\gamma_1\omega_{12} - \omega_{11} = -\sigma_{11}/\Delta$. Therefore, with σ_{11} and

γ_1 "known" (identified), the only remaining unknown is γ_2, which is therefore identified. With γ_1 and γ_2 in hand, β may be deduced from π_2. With γ_2 and β in hand, σ_{22} is the residual variance in the equation (y_2 - βx - $\gamma_2 y_1$) = ε_2, which is directly estimable, therefore, identified. □

2. Following the method in Example 13.6, for identification of the investment equation, we require that the matrix $\begin{bmatrix} (1) & (2) & (3) & (4) & (5) & (6) & (7) & (8) & (9) \\ -1 & \alpha_3 & 0 & 0 & \alpha_3 & 0 & 0 & 0 & 0 \\ 0 & -1 & \gamma_1 & 0 & 0 & 0 & 0 & \gamma_3 & \gamma_2 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ have rank 5. Columns (1), (4), (6), (7), and (8) each have one element in a different row, so they are linearly independent. Therefore, the matrix has rank five. For the third equation, the required matrix is $\begin{bmatrix} (1) & (2) & (3) & (4) & (5) & (6) & (7) & (8) & (9) & (10) \\ -1 & 0 & \alpha_1 & 0 & \alpha_3 & 0 & 0 & 0 & \alpha_2 & 0 \\ 0 & -1 & \beta_1 & 0 & 0 & 0 & 0 & 0 & \beta_2 & \beta_3 \\ 1 & 1 & 0 & 0 & 0 & 01 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. Columns (4), (6), (7), (9), and (10) are linearly independent. □

3. We find $[\mathbf{A}_3', \mathbf{A}_5']'$ for each equation.

$$\overset{(1)}{\begin{bmatrix} \gamma_{32} & 1 & \gamma_{34} \\ \beta_{12} & \beta_{13} & \beta_{14} \\ 0 & \beta_{43} & \beta_4 \\ \beta_{32} & 0 & 0 \end{bmatrix}}, \quad \overset{(2)}{\begin{bmatrix} 0 & \beta_{43} & \beta_{44} \end{bmatrix}}, \quad \overset{(3)}{\begin{bmatrix} 1 & \gamma_{12} & 0 \\ \gamma_{41} & \gamma_{42} & 1 \\ \beta_{21} & 1 & 0 \\ 0 & \beta_{52} & 00 \end{bmatrix}}, \quad \overset{(4)}{\begin{bmatrix} 1 & \gamma_{12} & 0 \\ \beta_{31} & \beta_{32} & \beta_{33} \\ 0 & \beta_{52} & 0 \end{bmatrix}}$$

Identification requires that the rank of each matrix be M-1 = 3. The second is obviously not identified. In (1), none of the three columns can be written as a linear combination of the other two, so it has rank 3. (Although the second and last columns have nonzero elements in the same positions, for the matrix to have short rank, we would require that the third column be a multiple of the second, since the first cannot appear in the linear combination which is to replicate the second column.) By the same logic, (3) and (4) are identified. □

4. Obtain the reduced form for the model in Exercise 1 under each of the assumptions made in parts (a) and (b1), (b6), and (b9).

(1). The model is $y_1 = \gamma_1 y_2 + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3 + \varepsilon_1$
$y_2 = \gamma_2 y_1 + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3 + \varepsilon_2$.

Therefore, $\mathbf{\Gamma} = \begin{bmatrix} 1 & -\gamma_2 \\ -\gamma_1 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} -\beta_{11} & -\beta_{12} \\ 0 & -\beta_{22} \\ -\beta_{31} & 0 \end{bmatrix}$ and $\mathbf{\Sigma}$ is unrestricted. The reduced form is

$\mathbf{\Pi} = \dfrac{1}{1-\gamma_1\gamma_2}\begin{bmatrix} \beta_{11}+\gamma_1\beta_{21} & \gamma_2\beta_{11}+\beta_{12} \\ \gamma_1\beta_{22} & \beta_{22} \\ \beta_{31} & \gamma_2\beta_{31} \end{bmatrix}$ and

$$\boldsymbol{\Omega} = (\boldsymbol{\Gamma}^{-1})'\boldsymbol{\Sigma}(\boldsymbol{\Gamma}^{-1}) = \frac{1}{(1-\gamma_1\gamma_2)^2}\begin{bmatrix} \sigma_{11}+\gamma_1^2\sigma_{22} + 2\gamma_1\sigma_{12} & \gamma_2\sigma_{11}+\gamma_1\sigma_{22} + (\gamma_1+\gamma_2)\sigma_{12} \\ \gamma_2\sigma_{11}+\gamma_1\sigma_{22} + (\gamma_1+\gamma_2)\sigma_{12} & \gamma_2^2\sigma_{11}+\sigma_{22} + 2\gamma_1\sigma_{12} \end{bmatrix}$$

(6) The model is $y_1 = \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3 + \varepsilon_1$
$y_2 = \gamma_2 y_1 + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3 + \varepsilon_2$

The first equation is already a reduced form. Substituting it into the second provides the second reduced form.

The coefficient matrix is $\mathbf{P} = \begin{bmatrix} \beta_{11} & \beta_{12}+\gamma_2\beta_{11} \\ \beta_{21} & \beta_{22}+\gamma_2\beta_{21} \\ \beta_{31} & \beta_{32}+\gamma_2\beta_{31} \end{bmatrix}$, $\boldsymbol{\Gamma}^{-1} = \begin{bmatrix} 1 & \gamma_2 \\ 0 & 1 \end{bmatrix}$ so $\boldsymbol{\Omega} = (\boldsymbol{\Gamma}^{-1})'\boldsymbol{\Sigma}(\boldsymbol{\Gamma}^{-1}) = \begin{bmatrix} \sigma_{11} & \gamma_2\sigma_{11} \\ \gamma_2\sigma_{11} & \gamma_2^2\sigma_{11}+\sigma_{22} \end{bmatrix}$

(9) The model is

$y_1 = \gamma_1 y_2 + \varepsilon_1$
$y_2 = \gamma_2 y_1 + \beta_{12}x_1 + \varepsilon_2$

Then, $\boldsymbol{\Pi} = -\mathbf{B}\boldsymbol{\Gamma}^{-1} = [\beta_{12}\gamma_1/(1-\gamma_1\gamma_2) \quad \beta_{12}/(1-\gamma_1\gamma_2)]$ and $\boldsymbol{\Omega} = \begin{bmatrix} \sigma_{11}+\gamma_1^2\sigma_{22} & \gamma_2\sigma_{11}+\gamma_1\sigma_{22} \\ \gamma_2\sigma_{11}+\gamma_1\sigma_{22} & \gamma_2^2\sigma_{11}+\sigma_{22} \end{bmatrix}$. □

5. The relevant submatrices are $\mathbf{X'X} = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 10 & 8 \\ 3 & 8 & 15 \end{bmatrix}$, $\mathbf{X'y}_1 = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{X'y}_2 = \begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix}$, $\mathbf{y_1'y_1} = 20$, $\mathbf{y}_2'\mathbf{y}_2 = 10$,

$\mathbf{y}_1'\mathbf{y}_2 = 6$, $\mathbf{X'Z}_1 = \begin{bmatrix} 3 & 5 \\ 6 & 2 \\ 7 & 3 \end{bmatrix}$, $\mathbf{X'Z}_2 = \begin{bmatrix} 4 & 2 & 3 \\ 3 & 10 & 8 \\ 5 & 8 & 15 \end{bmatrix}$ $\mathbf{Z}_1'\mathbf{Z}_1 = \begin{bmatrix} 10 & 3 \\ 3 & 5 \end{bmatrix}$, $\mathbf{Z}_2'\mathbf{Z}_2 = \begin{bmatrix} 10 & 3 & 5 \\ 3 & 10 & 8 \\ 5 & 8 & 15 \end{bmatrix}$,

$\mathbf{Z}_1'\mathbf{Z}_2 = \begin{bmatrix} 6 & 6 & 7 \\ 4 & 2 & 3 \end{bmatrix}$, $\mathbf{Z}_1'\mathbf{y}_1 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$, $\mathbf{Z}_1'\mathbf{y}_2 = \begin{bmatrix} 10 \\ 3 \end{bmatrix}$, $\mathbf{Z}_2'\mathbf{y}_1 = \begin{bmatrix} 20 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{Z}_2'\mathbf{y}_2 = \begin{bmatrix} 6 \\ 6 \\ 7 \end{bmatrix}$.

The two OLS coefficient vectors are

$\mathbf{d}_1 = (\mathbf{X'X})^{-1}\mathbf{X'y}_1 = [.439024, .536585]'$
$\mathbf{d}_2 = (\mathbf{X'X})^{-1}\mathbf{X'y}_2 = [.193016, .384127, .19746]'$.

The two stage least squares estimators are

$\hat{\delta}_1 = [\mathbf{Z}_1'\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'Z}_1]^{-1}[\mathbf{Z}_1'\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'y}_1] = [.368816, .578711]'$.

$\hat{\delta}_2 = [\mathbf{Z}_2'\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'Z}_2]^{-1}[\mathbf{Z}_2'\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'y}_2] = [.484375, .367188, .109375]'$.

$\hat{\sigma}_{11} = (\mathbf{y}_1'\mathbf{y}_1 - 2\mathbf{y}_1'\mathbf{Z}\hat{\delta}_1 + \hat{\delta}_1'\mathbf{Z}_1'\mathbf{Z}_1\hat{\delta}_1) / 25 = .610397$, $\hat{\sigma}_{22} = .268384$.

The estimated asymptotic covariance matrices are

$$\text{Est.Var}[\hat{\delta}_1] = \hat{\sigma}_{11}[\mathbf{Z}_1'\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'Z}_1]^{-1} = \begin{bmatrix} .215858 & .129035 \\ .129036 & .1995 \end{bmatrix}$$

$$\text{Est.Var}[\text{Est.Var}[\hat{\delta}_2]] = \begin{bmatrix} .132423 & -.007699 & -.040035 \\ -.007688 & .047259 & -.022538 \\ -.040035 & -.022638 & .043311 \end{bmatrix}.$$

The three stage least squares estimate is

$$\begin{bmatrix} \hat{\sigma}^{11}[\mathbf{Z}_1'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_1] & \hat{\sigma}^{12}[\mathbf{Z}_1'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_2] \\ \hat{\sigma}^{12}[\mathbf{Z}_2'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_1] & \hat{\sigma}^{22}[\mathbf{Z}_2'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_2] \end{bmatrix}^{-1} \begin{bmatrix} \hat{\sigma}^{11}[\mathbf{Z}_1'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_1]+ \\ \hat{\sigma}^{12}[\mathbf{Z}_1'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_2] \\ \\ \hat{\sigma}^{12}[\mathbf{Z}_2'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_1]+ \\ \hat{\sigma}^{22}[\mathbf{Z}_2'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_2] \end{bmatrix}$$

$$= [.368817,.578708,.4706,.306363,.168294]'.$$

The estimated standard errors are the square roots of the diagonal elements of the inverse matrix, [.4637,.4466,.3626,.1716,.1628], compared to the 2SLS values, [.4637,.4466,.3639,.2174,.2081].

To compute the limited information maximum likelihood estimator, we require the matrix of sums of squares and cross products of residuals of the regressions of $\mathbf{y}_1$ and $\mathbf{y}_2$ on $\mathbf{x}_1$ and on $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$. These are

$$\mathbf{W}^0 = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{x}_1(\mathbf{x}_1'\mathbf{x}_1)^{-1}\mathbf{x}_1'\mathbf{Y} = \begin{bmatrix} 16.5 & 3.60 \\ 3.60 & 8.20 \end{bmatrix}, \mathbf{W}^1 = \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 16.2872 & 2.55312 \\ 2.55312 & 5.3617 \end{bmatrix}.$$

The two characteristic roots of $(\mathbf{W}^1)^{-1}\mathbf{W}^0$ are 1.53157 and 1.00837. We carry the smaller one into the *k*-class computation [see, for example, Theil (1971) or Judge, et al (1985)];

$$\hat{\delta}_{1k} = \begin{bmatrix} 10-1.00837(5.3617) & 3 \\ 3 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 6-1.00837(2.55312) \\ 4 \end{bmatrix} = \begin{bmatrix} .367116 \\ .57973 \end{bmatrix}$$

Finally, the two estimates of the reduced form are

$$\text{(OLS)} \quad \mathbf{P} = \begin{bmatrix} .680851 & .329787 \\ .010638 & .37243 \\ .191489 & .202128 \end{bmatrix}$$

$$\text{and} \quad \text{(2SLS)} \quad \hat{\Pi} = \begin{bmatrix} -.578711 & 0 \\ 0 & -.367188 \\ 0 & -.109375 \end{bmatrix} \begin{bmatrix} 1 & -.484375 \\ -.368816 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} .704581 & .341281 \\ .104880 & .447051 \\ .049113 & .133164 \end{bmatrix}.$$

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.

The structure is $$[y_1\ y_2] \begin{bmatrix} 1 & -\gamma_2 \\ -\gamma_1 & 1 \end{bmatrix} + [x_1\ x_2\ x_3\ x_4] \begin{bmatrix} \beta_{11} & 0 \\ \beta_{21} & 0 \\ 0 & \beta_{32} \\ 0 & \beta_{42} \end{bmatrix} = [\varepsilon_1\ \varepsilon_1].$$

or $\mathbf{y}'\boldsymbol{\Gamma} + \mathbf{x}'\mathbf{B} = \boldsymbol{\varepsilon}'$. The reduced form coefficient matrix is

$$\Pi = -\mathbf{B}\boldsymbol{\Gamma}^{-1} = \frac{1}{1-\gamma_1\gamma_2} \begin{bmatrix} \beta_{11} & \gamma_2\beta_{11} \\ \beta_{21} & \gamma_2\beta_{21} \\ \gamma_1\beta_{32} & \beta_{32} \\ \gamma_1\beta_{42} & \beta_{42} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{21} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \\ \pi_{41} & \pi_{42} \end{bmatrix}$$

The two restrictions are $\pi_{12}/\pi_{11} = \pi_{22}/\pi_{21}$ and $\pi_{31}/\pi_{32} = \pi_{41}/\pi_{42}$. If we write the reduced form as

$$y_1 = \pi_{11}x_1 + \pi_{21}x_2 + \pi_{31}x_3 + \pi_{41}x_4 + v_1$$
$$y_2 = \pi_{12}x_1 + \pi_{22}x_2 + \pi_{32}x_3 + \pi_{42}x_4 + v_2.$$

We could treat the system as a nonlinear seemingly unrelated regressions model. One possible way to handle the restrictions is to eliminate two parameters directly by making the substitutions

$$\pi_{12} = \pi_{11}\pi_{22}/\pi_{21} \quad \text{and} \quad \pi_{31} = \pi_{32}\pi_{41}/\pi_{42}.$$

The pair of equations would be

$$y_1 = \pi_{11}x_1 + \pi_{21}x_2 + (\pi_{32}\pi_{41}/\pi_{42})x_3 + \pi_{41}x_4 + v_1$$
$$y_2 = (\pi_{11}\pi_{22}/\pi_{21})x_1 + \pi_{22}x_2 + \pi_{32}x_3 + \pi_{42}x_4 + v_2.$$

This nonlinear system could now be estimated by nonlinear GLS. The function to be minimized would be

$$\Sigma_{i=1}^{n}\ v_{i1}^2\sigma^{11} + v_{i2}^2\sigma^{22} + 2v_{i1}v_{i2}\sigma^{12} = n\text{tr}(\mathbf{\Sigma}^{-1}\mathbf{W}).$$

Needless to say, this would be quite involved. □

7. We would require that all three characteristic roots have modulus less than one. An intuitive guess that the diagonal element greater than one would preclude this would be correct. The roots are the solutions to

$$det\begin{bmatrix} -.1899-\lambda & -.9471 & -.8991 \\ 0 & 1.0287-\lambda & 0 \\ -.0656 & -.0791 & .0952-\lambda \end{bmatrix} = 0.$$ Expanding this produces $-(.1899 + \lambda)(1.0287 - \lambda)(.0952 - \lambda)$ $- .0565(1.0287 - \lambda).8991 = 0$. There is no need to go any further. It is obvious that $\lambda = 1.0287$ is a solution, so there is at least one characteristic root larger than 1. The system is unstable.

8. Prove plim $\mathbf{Y}_j'\boldsymbol{\varepsilon}/T = \boldsymbol{\omega}_\mathbf{j} - \mathbf{\Omega}_{jj}\boldsymbol{\gamma}_\mathbf{j}$.

Consistent with the partitioning $\mathbf{y}' = [y_j\ \mathbf{Y}_j'\ \mathbf{Y}_i^{*\prime}]$, partition $\mathbf{\Omega}$ into

$$\mathbf{\Omega} = \begin{matrix} \omega_{jj} & \boldsymbol{\omega}_j' & \boldsymbol{\omega}^{*\prime}_j \\ \boldsymbol{\omega}_j & \mathbf{\Omega}_{jj} & \mathbf{\Omega}_j' \\ \boldsymbol{\omega}^*_j & \mathbf{\Omega}^*_j & \mathbf{\Omega}_j^* \end{matrix}$$

and, as in the equation preceding (13-8), partition the jth column of $\mathbf{\Gamma}$ as $\mathbf{\Gamma}_j = \begin{bmatrix} 1 \\ -\boldsymbol{\gamma} \\ \mathbf{0} \end{bmatrix}$. Since the full set of reduced form disturbances is $\mathbf{V} = \mathbf{E\Gamma}^{-1}$, it follows that $\mathbf{E} = \mathbf{V\Gamma}$. In particular, the jth column of $\mathbf{E}$ is $\boldsymbol{\varepsilon}_j = \mathbf{V\Gamma}_j$. In the reduced form, now referring to (15-8), $\mathbf{Y}_j = \mathbf{X\Pi}_j + \mathbf{V}_j$, where $\mathbf{\Pi}_j$ is the M_j columns of $\mathbf{\Pi}$ corresponding to the included endogenous variables and $\mathbf{V}_j$ is the $T\times M_j$ matrix of their reduced form disturbances. Since $\mathbf{X}$ is uncorrelated with all columns of $\mathbf{E}$, we have

$$\text{plim } \mathbf{Y}_j'\varepsilon_j/T = \text{plim } \mathbf{V}_j'\,\mathbf{\Gamma}_j/T = [\boldsymbol{\omega}_j\ \ \mathbf{\Omega}_{jj}\ \ \mathbf{\Omega}_j^*]\begin{bmatrix} 1 \\ -\boldsymbol{\gamma} \\ \mathbf{0} \end{bmatrix} = \boldsymbol{\omega}_j - \mathbf{\Omega}_{jj}\boldsymbol{\gamma}_j$$ as required.

9. Prove that an underidentified equation cannot be estimated by two stage least squares.

If the equation fails the order condition, then the number of excluded exogenous variables is less than the number of included endogenous. The matrix of instrumental variables to be used for two stage least squares is of the form $\hat{\mathbf{Z}} = [\mathbf{XA},\mathbf{X}_j]$, where $\mathbf{XA}$ is M_j linear combination of all K columns in $\mathbf{X}$ and $\mathbf{X}_j$ is K_j columns of $\mathbf{X}$. In total, $K = K_j^* + K_j$. If the equation fails the order condition, then $K_j^* < M_j$, so $\hat{\mathbf{Z}}$ is $M_j + K_j$ columns which are linear combinations of $K = K_j^* + K_j < M_j + K_j$. Therefore, $\hat{\mathbf{Z}}$ cannot have full column rank. In order to compute the two stage least squares estimator, we require $(\hat{\mathbf{Z}}'\hat{\mathbf{Z}})^{-1}$, which cannot be computed.

Application

```
?=========================================================
? Application 13.1 - Simultaneous Equations
?=========================================================
? Read the data
? For convenience, rename the variables so they correspond
? to the example in the text.
sample ; 1 - 204 $
create ; ct=realcons$
create ; it=realinvs$
create ; gt=realgovt$
create ; rt=tbilrate $
? Impose (artifically) the adding up condition on total demand.
create ; yt=ct+it+gt $
create ; ct1=ct[-1] $
create ; yt1 = yt[-1] $
create ; dyt = yt - yt1 $
sample ; 2-204 $
names  ; xt = one,gt,rt,ct1,yt1$
? Estimate equations by 2sls and save coefficients with
? the names used in the example.
2sls   ; lhs = ct ; rhs=one,yt,ct1 ; inst = xt $
+----------------------------------------------------+
| Two stage   least squares regression               |
| LHS=CT       Mean                 =   3008.995     |
|              Standard deviation   =   1456.900     |
| WTS=none     Number of observs.   =        203     |
| Model size   Parameters           =          3     |
|              Degrees of freedom   =        200     |
| Residuals    Sum of squares       =   75713.32     |
|              Standard error of e  =   19.45679     |
| Fit          R-squared            =   .9998208     |
|              Adjusted R-squared   =   .9998190     |
| Model test   F[  2,   200] (prob) =******* (.0000) |
+----------------------------------------------------+
| Instrumental Variables:
|ONE      GT       RT       CT1      YT1
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|   -13.8657181      5.31536302    -2.609   .0091
 YT      |     .05843862       .01790473     3.264   .0011   4663.67389
 CT1     |     .92200662       .02657199    34.698   .0000   2982.97438
calc   ; a0=b(1) ; a1=b(2) ; a2=b(3) $
2sls   ; lhs = it ; rhs=one,rt,dyt ; inst = xt $
+----------------------------------------------------+
| Two stage   least squares regression               |
| LHS=IT       Mean                 =   654.5296     |
|              Standard deviation   =   391.3705     |
| WTS=none     Number of observs.   =        203     |
| Model size   Parameters           =          3     |
|              Degrees of freedom   =        200     |
| Residuals    Sum of squares       =   .7744227E+08 |
|              Standard error of e  =   622.2631     |
| Fit          R-squared            =  -1.540485     |
|              Adjusted R-squared   =  -1.565889     |
+----------------------------------------------------+
| Instrumental Variables:
|ONE      GT       RT       CT1      YT1
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|   -300.699429      125.980850    -2.387   .0170
 RT      |    56.5192542      15.4643912     3.655   .0003   5.24965517
 DYT     |    16.5359646      2.02509785     8.166   .0000   39.8236453
calc   ; b0=b(1) ; b1=b(2) ; b2=b(3) $
```

```
?
? Create the coefficients of the reduced form.  We only need the parts
? for the dynamics.  These are in the second half of the example.
calc   ; a=1-a1-b2 $
?
? Construct the matrix that governs the dynamics of the system.  Note that
? the I equation is static.  It is a function of y(t-1) and c(t-1) but not
? of I(t-1).  This is the DELTA(1) submatrix in (13-42).  The dominant
? root is the largest rood of DELTA(1).
calc   ; list ; C11=(1-b2)/a ; C12=-a1*b2/a ; C21=a2/a ; C22=-b2/a $
matrix ; C = [c11,c12 / c21,c22] $
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 C11     =       .996253
 C12     =       .061967
 C21     =      -.059124
 C22     =      1.060378
Matrix ; list ; roots = cxrt(c)$
Calc   ; list ; domroot = sqr(roots(1,1)^2 + roots(1,2)^2)$
--> Matrix ; list ; roots = cxrt(c)$

Matrix ROOTS    has  2 rows and  2 columns.
               1             2
        +--------------------------
       1|    1.02832      -.05134
       2|    1.02832       .05134
--> Calc   ; list ; domroot = sqr(roots(1,1)^2 + roots(1,2)^2)$
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 DOMROOT =      1.029596

? The largest root is larger than on in absolute value.  The system is unstable.

3sls   ; lhs = ct,it ; eq1=one,yt,ct1 ; eq2=one,rt,dyt ; inst=xt ; maxit=0 $
+---------------------------------------------------+
| Estimates for equation: CT                        |
| InstVar/GLS least squares regression              |
| LHS=CT       Mean                 =   3008.995    |
| Residuals    Sum of squares       =   73370.06    |
|              Standard error of e  =   19.15334    |
+---------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|   -17.4780776      4.55837624    -3.834   .0001
 YT      |     .07312129       .01415744     5.165   .0000   4663.67389
 CT1     |     .90026227       .02103720    42.794   .0000   2982.97438
+---------------------------------------------------+
| Estimates for equation: IT                        |
| InstVar/GLS least squares regression              |
| LHS=IT       Mean                 =   654.5296    |
| Residuals    Sum of squares       =   .9735005E+08 |
|              Standard error of e  =   697.6749    |
+---------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|   -236.744328      122.661644    -1.930   .0536
 RT      |    30.5417941      12.9861014     2.352   .0187   5.24965517
 DYT     |    18.3544221      1.93633720     9.479   .0000   39.8236453
```

Chapter 14

Estimation Frameworks in Econometrics

Exercise

1. A fully parametric model/estimator provides consistent, efficient, and comparatively precise results. The semiparametric model/estimator, by comparison, is relatively less precise in general terms. But, the payoff to this imprecision is that the semiparametric formulation is more likely to be robust to failures of the assumptions of the parametric model. Consider, for example, the binary probit model of Chapter 21, which makes a strong assumption of normality and homoscedasticity. If the assumptions are correct, the probit estimator is the most efficient use of the data. However, if the normality assumption or the homoscedasticity assumption are incorrect, then the probit estimator becomes inconsistent in an unknown fashion. Lewbel's semiparametric estimator for the binary choice model, in contrast, is not very precise in comparison to the probit model. But, it will remain consistent if the normality assumption is violated, and it is even robust to certain kinds of heteroscedasticity.

Applications

1. Using the gasoline market data in Appendix Table F2.2, use the partially linear regression method in Section 16.3.3 to fit an equation of the form

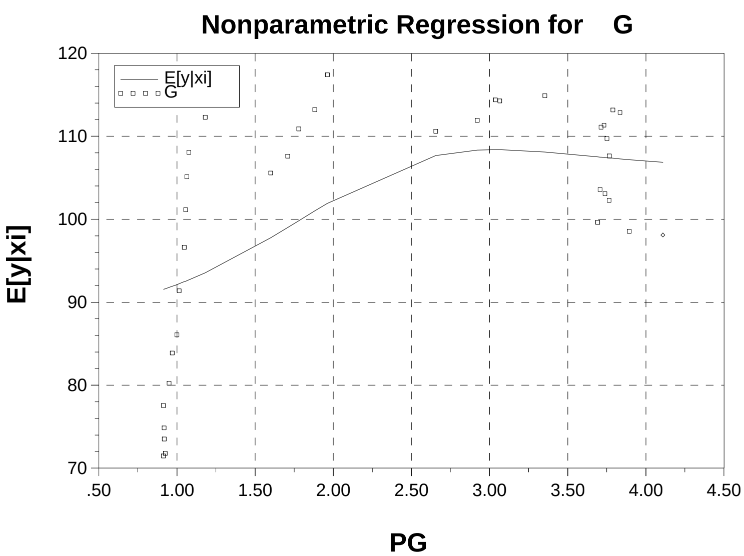
$$\ln(G/Pop) = \beta_1 \ln(Income) + \beta_2 \ln P_{new\ cars} + \beta_3 \ln P_{used\ cars} + g(\ln P_{gasoline}) + \varepsilon$$

```
crea;gp=lg;ip=ly;ncp=lpnc;upp=lpuc;pgp=lpg$
sort;lhs=pgp;rhs=gp,ip,ncp,upp$
crea;dgp=.809*gp - .5*gp[-1] - .309*gp[-2]$
crea;dip=.809*ip - .5*ip[-1] - .309*ip[-2]$
crea;dnc=.809*ncp -.5*ncp[-1]-.309*ncp[-2]$
crea;duc=.809*upp -.5*upp[-1]-.309*upp[-2]$
samp;3-36$
regr;lhs=dgp;rhs=dip,dnc,duc;res=e$
+-----------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = DGP      Mean=   .9708646870E-02, S.D.=   .4738748109E-01 |
| Model size: Observations =       34, Parameters =   3, Deg.Fr.=     31 |
| Residuals:  Sum of squares= .1485994289E-01, Std.Dev.=         .02189 |
| Fit:        R-squared=  .799472, Adjusted R-squared =          .78653 |
| Model test: F[  2,     31] =   61.80,    Prob value =          .00000 |
| Diagnostic: Log-L =     83.2587, Restricted(b=0) Log-L =      55.9431 |
|             LogAmemiyaPrCrt.=   -7.559, Akaike Info. Crt.=     -4.721 |
| Model does not contain ONE. R-squared and F can be negative!          |
| Autocorrel: Durbin-Watson Statistic =   1.34659,   Rho =       .32671 |
+-----------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 DIP          .9629902959       .11631885    8.279   .0000  .14504254E-01
 DNC        -.1010972781   .87755182E-01   -1.152   .2581  .20153536E-01
 DUC      -.3197058148E-01  .51875022E-01    -.616   .5422  .35656776E-01
--> matr;varpl={1+1/(2*2)}*varb$
--> matr;stat(b,varpl)$
+-----------------------------------------------+
```

```
|Number of observations in current sample =      34 |
|Number of parameters computed here       =       3 |
|Number of degrees of freedom             =      31 |
+----------------------------------------------------+
+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
 B_1          .9629902959       .13004843    7.405   .0000
 B_2         -.1010972781   .98113277E-01   -1.030   .3028
 B_3      -.3197058148E-01  .57998037E-01    -.551   .5815
```

2.

```
+----------------------------------------+
| Nonparametric Regression for G         |
| Observations       =           36      |
| Points plotted     =           36      |
| Bandwidth          =       .468092     |
| Statistics for abscissa values----     |
| Mean               =      2.316611     |
| Standard Deviation =      1.251735     |
| Minimum            =       .914000     |
| Maximum            =      4.109000     |
| ------------------------------------   |
| Kernel Function    =      Logistic     |
| Cross val. M.S.E.  =    121.084982     |
| Results matrix     =        KERNEL     |
+----------------------------------------+
```



3. A. Using the probit model and the Klein and Spady semiparametric models, the two sets of coefficient estimates are somewhat similar.

```
+----------------------------------------------+
| Binomial Probit Model                        |
| Maximum Likelihood Estimates                 |
| Model estimated: Jul 31, 2002 at 05:16:40PM.|
| Dependent variable                   P       |
| Weighting variable                None       |
| Number of observations             601       |
| Iterations completed                 5       |
```

```
| Log likelihood function       -307.2955     |
| Restricted log likelihood     -337.6885     |
| Chi squared                    60.78608     |
| Degrees of freedom                    5     |
| Prob[ChiSqd > value] =         .0000000     |
| Hosmer-Lemeshow chi-squared =   5.74742     |
| P-value=  .67550 with deg.fr. =       8     |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Index function for probability
 Z2       -.2202376072E-01  .10177371E-01   -2.164   .0305     32.487521
 Z3        .5990084920E-01  .17086004E-01    3.506   .0005     8.1776955
 Z5         -.1836462412   .51493239E-01   -3.566   .0004     3.1164725
 Z7        .3751312008E-01  .32844576E-01    1.142   .2534     4.1946755
 Z8         -.2729824396   .52473295E-01   -5.202   .0000     3.9317804
 Constant     .9766647244       .36104809    2.705   .0068
+---------------------------------------------+
| Seimparametric Binary Choice Model          |
| Maximum Likelihood Estimates                |
| Model estimated: Jul 31, 2002 at 11:01:24PM.|
| Dependent variable                    P     |
| Weighting variable                 None     |
| Number of observations              601     |
| Iterations completed                 13     |
| Log likelihood function       -334.7367     |
| Restricted log likelihood     -337.6885     |
| Chi squared                    5.903551     |
| Degrees of freedom                    4     |
| Prob[ChiSqd > value] =         .2064679     |
| Hosmer-Lemeshow chi-squared = 118.69649     |
| P-value=  .00000 with deg.fr. =       8     |
| Logistic kernel fn. Bandwidth =  .34423     |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Characteristics in numerator of Prob[Y = 1]
 Z2       -.3284308221E-01  .52254249E-01    -.629   .5297     32.487521
 Z3           .1089817386   .86483083E-01    1.260   .2076     8.1776955
 Z5         -.2384951835       .23320058   -1.023   .3064     3.1164725
 Z7         -.1026067037       .17130225    -.599   .5492     4.1946755
 Z8         -.1892263132       .21598982    -.876   .3810     3.9317804
 Constant     .0000000000 ........(Fixed Parameter)........
```

The probit model produces a set of marginal effects, as discussed in the text. These cannot be computed for the Klein and Spady estimator.

```
+---------------------------------------------+
| Partial derivatives of E[y] = F[*]   with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs.  |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
          Index function for probability
 Z2       -.6695300413E-02  .30909282E-02   -2.166   .0303     32.487521
 Z3        .1821006800E-01  .51704684E-02    3.522   .0004     8.1776955
 Z5       -.5582910069E-01  .15568275E-01   -3.586   .0003     3.1164725
 Z7        .1140411992E-01  .99845393E-02    1.142   .2534     4.1946755
 Z8       -.8298761795E-01  .15933104E-01   -5.209   .0000     3.9317804
 Constant     .2969094977       .11108860    2.673   .0075
```

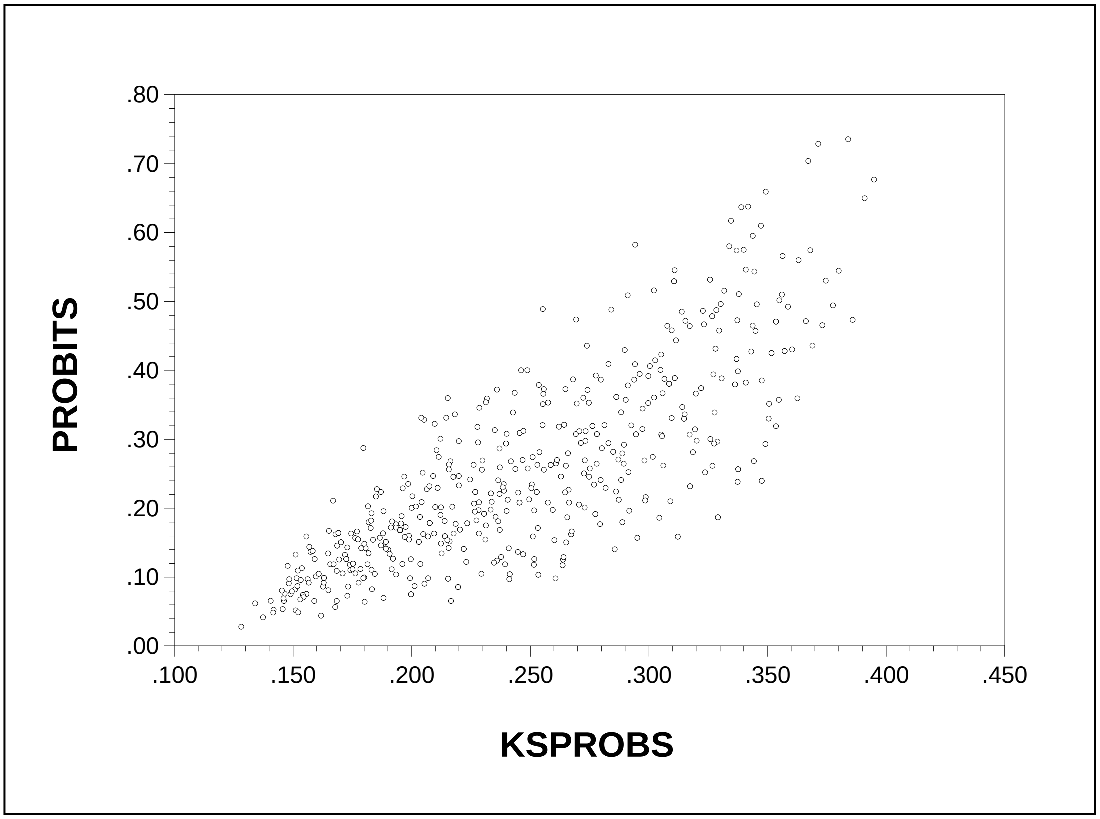
These are the various fit measures for the probit model

```
+----------------------------------------+
| Fit Measures for Binomial Choice Model |
| Probit   model for variable P          |
+----------------------------------------+
| Proportions P0= .750416   P1= .249584  |
| N =     601 N0=     451   N1=     150  |
| LogL =  -307.29545 LogL0 =  -337.6885  |
| Estrella = 1-(L/L0)^(-2L0/n) = .10056  |
+----------------------------------------+
|     Efron |  McFadden  |  Ben./Lerman  |
|    .10905 |    .09000  |       .66451  |
|    Cramer | Veall/Zim. |     Rsqrd_ML  |
|    .10486 |    .17359  |       .09619  |
+----------------------------------------+
| Information  Akaike I.C. Schwarz I.C.  |
| Criteria        1.04258    652.98248  |
+----------------------------------------+
Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.
Threshold value for predicting Y=1 = .5000
            Predicted
------  ----------  +  -----
Actual      0    1  |  Total
------  ----------  +  -----
  0       437   14  |    451
  1       130   20  |    150
------  ----------  +  -----
Total     567   34  |    601
```

These are the fit measures for the probabilities computed for the Klein and Spady model. The probit model fits better by all measures computed.

```
+----------------------------------------+
| Fit Measures for Binomial Choice Model |
| Observed = P        Fitted = KSPROBS   |
+----------------------------------------+
| Proportions P0= .750416   P1= .249584  |
| N =     601 N0=     451   N1=     150  |
| LogL =  -320.37513 LogL0 =  -337.6885  |
| Estrella = 1-(L/L0)^(-2L0/n) = .05743  |
+----------------------------------------+
|     Efron |  McFadden  |  Ben./Lerman  |
|    .05686 |    .05127  |       .64117  |
|    Cramer | Veall/Zim. |     Rsqrd_ML  |
|    .03897 |    .10295  |       .05599  |
+----------------------------------------+
```

The first figure below plots the probit probabilities against the Klein and Spady probabilities. The models are obviously similar, though there is substantial difference in the fitted values.



Finally, these two figures plot the predicted probabilities from the two models against the respective index functions, **b'x**. Note that the two plots are based on different coefficient vectors, so it is not possible to merge the two figures.