Chapter 18

Bayesian Estimation and Inference

Exercise

a. The likelihood function is

$$L(\mathbf{y}|\lambda) = \prod_{i=1}^{n} f(y_i \mid \lambda) = \prod_{i=1}^{n} \frac{\exp(-\lambda)\lambda^{y_i}}{\Gamma(y_i+1)} = \exp(-n\lambda)\lambda^{\Sigma_i y_i} \prod_{i=1}^{n} \frac{1}{\Gamma(y_i+1)}.$$

b. The posterior is

$$p(\lambda \mid y_1,\ldots,y_n) = \frac{p(y_1,\ldots,y_n \mid \lambda)\,p(\lambda)}{\int_0^\infty p(y_1,\ldots,y_n \mid \lambda)\,p(\lambda)\,d\lambda}.$$

The product of factorials will fall out. This leaves

$$\begin{aligned}
p(\lambda \mid y_1,\ldots,y_n) &= \frac{\exp(-n\lambda)\lambda^{\Sigma_i y_i}(1/\lambda)}{\int_0^\infty \exp(-n\lambda)\lambda^{\Sigma_i y_i}(1/\lambda)\,d\lambda} \\
&= \frac{\exp(-n\lambda)\lambda^{(\Sigma_i y_i)-1}}{\int_0^\infty \exp(-n\lambda)\lambda^{(\Sigma_i y_i)-1}\,d\lambda} \\
&= \frac{\exp(-n\lambda)\lambda^{n\bar{y}-1}}{\int_0^\infty \exp(-n\lambda)\lambda^{n\bar{y}-1}\,d\lambda} \\
&= \frac{n^{n\bar{y}}\exp(-n\lambda)\lambda^{n\bar{y}-1}}{\Gamma(n\bar{y})}.
\end{aligned}$$

where we have used the gamma integral at the last step. The posterior defines a two parameter gamma distribution, G(n, $n\bar{y}$).

c. The estimator of λ is the mean of the posterior. There is no need to do the integration. This falls simply out of the posterior density, $E[\lambda|\mathbf{y}] = n\bar{y}/n = \bar{y}$.

d. The posterior variance also drops out simply; it is $n\bar{y}/n^2 = \bar{y}/n$.

Application

a. $p(F_i|K_i,\theta) = \binom{K_i}{F_i}\theta^{F_i}(1-\theta)^{K_i-F_i}$ so the log likelihood function is

$$\ln L(\theta\,|\,\mathbf{y}) = \sum_{i=1}^{n} \ln\binom{K_i}{F_i} + F_i \ln\theta + (K_i - F_i)\ln(1-\theta)$$

The MLE is obtained by setting $\partial \ln L(\theta|y)/\partial\theta = \Sigma_i\,[F_i/\theta - (K_i-F_i)/(1-\theta)] = 0$. Multiply both sides by $\theta(1-\theta)$ to obtain

$\Sigma_i\,[(1-\theta)F_i - \theta\,(K_i-F_i)] = 0$

A line of algebra reveals that the solution is $\theta = (\Sigma_i F_i)/(\Sigma_i K_i) = 0.651596$.

b. The posterior density is $\dfrac{\left[\prod_{i=1}^{n}\binom{K_i}{F_i}\theta^{F_i}(1-\theta)^{K_i-F_i}\right]\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}}{\int_0^1\left[\prod_{i=1}^{n}\binom{K_i}{F_i}\theta^{F_i}(1-\theta)^{K_i-F_i}\right]\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}d\theta}$

This simplifies considerably. The combinatorials and gamma functions fall out, leaving

$$p(\theta\,|\,\mathbf{y}) = \frac{\left[\prod_{i=1}^{n}\theta^{F_i}(1-\theta)^{K_i-F_i}\right]\theta^{a-1}(1-\theta)^{b-1}}{\int_0^1\left[\prod_{i=1}^{n}\theta^{F_i}(1-\theta)^{K_i-F_i}\right]\theta^{a-1}(1-\theta)^{b-1}d\theta} = \frac{\left[\theta^{\Sigma_i F_i}(1-\theta)^{\Sigma_i(K_i-F_i)}\right]\theta^{a-1}(1-\theta)^{b-1}}{\int_0^1\left[\theta^{\Sigma_i F_i}(1-\theta)^{\Sigma_i(K_i-F_i)}\right]\theta^{a-1}(1-\theta)^{b-1}d\theta}$$

$$= \frac{\left[\theta^{(\Sigma_i F_i)+(a-1)}(1-\theta)^{[\Sigma_i(K_i-F_i)]+(b-1)}\right]}{\int_0^1\left[\theta^{(\Sigma_i F_i)+(a-1)}(1-\theta)^{\Sigma_i(K_i-F_i)]+(b-1)}\right]d\theta}$$

The denominator is a beta integral, so the posterior density is

$$p(\theta\,|\,\mathbf{y}) = \frac{\Gamma[(\Sigma_i F_i)+(a-1)]\Gamma[(\Sigma_i(K_i-F_i))+(b-1)]}{\Gamma[(\Sigma_i F_i)+(a-1)+(\Sigma_i(K_i-F_i))+(b-1)]}\left[\theta^{(\Sigma_i F_i)+(a-1)}(1-\theta)^{[\Sigma_i(K_i-F_i)]+(b-1)}\right]$$

The denominator simplifies slightly;

$$p(\theta\,|\,\mathbf{y}) = \frac{\Gamma[(\Sigma_i F_i)+(a-1)]\Gamma[(\Sigma_i(K_i-F_i))+(b-1)]}{\Gamma[(\Sigma_i K_i)+(a-1)+(b-1)]}\left[\theta^{(\Sigma_i F_i)+(a-1)}(1-\theta)^{[\Sigma_i(K_i-F_i)]+(b-1)}\right]$$

$$= \frac{\Gamma[(a+\Sigma_i F_i)-1)]\Gamma[(b+\Sigma_i(K_i-F_i))-1)]}{\Gamma[(a+b)+(\Sigma_i K_i)-1-1)]}\left[\theta^{(a+\Sigma_i F_i)-1}(1-\theta)^{[b+\Sigma_i(K_i-F_i)]-1}\right]$$

c-e. The posterior distribution is a beta distribution with parameters $a^*=(a+\Sigma_i F_i)$ and $b^*=[b+\Sigma_i(K_i-F_i)]$. The mean of this beta random variable is $a^*/(a^*+b^*) = (a+\Sigma_i F_i)/(a+b+\Sigma_i K_i)$. In the data, $\Sigma_i = 49$ and $\Sigma_i K_i = 75$. For the values given, the posterior means are

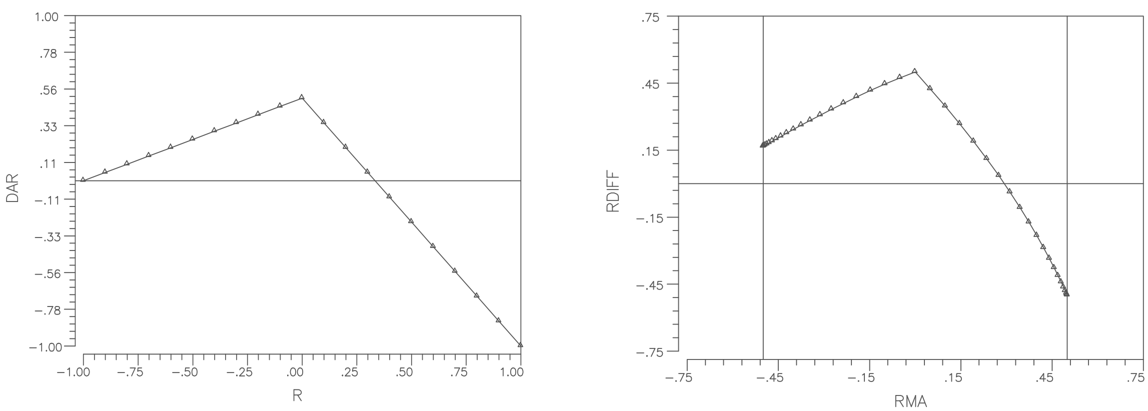
```
(a=1,b=1): Result  =       .647668
(a=2,b=2): Result  =       .643939
(a=1,b=2): Result  =       .639386
```

Chapter 19

Serial Correlation

Exercises

1. For the first order autoregressive model, the autocorrelation is ρ. Consider the first difference, $v_t = \varepsilon_t - \varepsilon_{t-1}$ which has $\text{Var}[v_t] = 2\text{Var}[\varepsilon_t] - 2\text{Cov}[(\varepsilon_t,\varepsilon_{t-1})] = 2\sigma_u^2[1/(1-\rho^2) - \rho/(1-\rho^2)] = 2\sigma_u^2/(1+\rho)$ and $\text{Cov}[v_t,v_{t-1}] = 2\text{Cov}[\varepsilon_t,\varepsilon_{t-1}] - \text{Var}[\varepsilon_t] - \text{Cov}[\varepsilon_t,\varepsilon_{t-1}] = \sigma_u^2[1/(1-\rho^2)][2\rho - 1 - \rho^2] = \sigma_u^2[(\rho-1)/(1+\rho)]$. Therefore, the autocorrelation of the differenced process is $\text{Cov}[v_t,v_{t-1}] / \text{Var}[v_t] = (\rho - 1)/2$. As the figure below on the left shows, first differencing reduces the absolute value of the autocorrelation coefficient when ρ is greater than 1/3. For economic data, this is likely to be fairly common.



For the moving average process, the first order autocorrelation is $\text{Cov}[(\varepsilon_t,\varepsilon_{t-1})]/\text{Var}[\varepsilon_t] = -\lambda/(1+\lambda^2)$. To obtain the autocorrelation of the first difference, write $\varepsilon_t - \varepsilon_{t-1} = u_t - (1+\lambda)u_{t-1} + \lambda u_{t-2}$ and $\varepsilon_{t-1} - \varepsilon_{t-2} = u_{t-1} - (1+\lambda)u_{t-2} + \lambda u_{t-3}$. The variance of the difference is $\text{Var}[\varepsilon_t - \varepsilon_{t-1}] = \sigma_u^2[(1+\lambda)^2 + (1+\lambda^2)]$. The covariance can be found by taking the expected product of terms with equal subscripts. Thus, $\text{Cov}[\varepsilon_t - \varepsilon_{t-1},\varepsilon_{t-1} - \varepsilon_{t-2}] = -\sigma_u^2(1+\lambda)^2$. The autocorrelation is $\text{Cov}[\varepsilon_t - \varepsilon_{t-1},\varepsilon_{t-1} - \varepsilon_{t-2}]/\text{Var}[\varepsilon_t - \varepsilon_{t-1}] = -(1+\lambda)^2/[(1+\lambda)^2 + (1+\lambda^2)]$. A plot of the relationship between the differenced and undifferenced series is shown in the right panel above. The horizontal axis plots the autocorrelation of the original series. The values plotted are the absolute values of the difference between the autocorrelation of the differenced series and the original series. The results are similar to those for the AR(1) model. For most of the range of the autocorrelation of the original series, differencing increases autocorrelation. But, for most of the range of values that are economically meaningful, differencing reduces autocorrelation.

2. Derive the disturbance covariance matrix for the model $y_t = \boldsymbol{\beta}'\mathbf{x}_t + \varepsilon_t$, $\varepsilon_t = \rho\varepsilon_{t-1} + u_t - \lambda u_{t-1}$. What parameter is estimated by the regression of the ordinary least squares residuals on their lagged values?

Solve the disturbance process in its moving average form. Write the process as $\varepsilon_t - \rho\varepsilon_{t-1} = u_t - \lambda u_{t-1}$ or, using the lag operator, $\varepsilon_t(1 - \rho L) = u_t - \lambda u_{t-1}$ or $\varepsilon_t = u_t/(1-\rho L) - \lambda u_{t-1}/(1-\rho L)$. After multiplying these out, we obtain

$$\varepsilon_t = u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + ... - \lambda u_{t-1} - \rho\lambda u_{t-2} - \rho^2\lambda u_{t-3} - ...$$
$$= u_t + (\rho-\lambda)u_{t-1} + \rho(\rho-\lambda)u_{t-2} + \rho^2(\rho-\lambda)u_{t-3} + ...$$

Therefore,

$$\text{Var}[\varepsilon_t] = \sigma_u^2(1 + (\rho-\lambda)^2)(1 + \rho^2 + \rho^4 + ...) = \sigma_u^2(1 + (\rho-\lambda)^2/(1-\rho^2))$$
$$= \sigma_u^2(1 + \lambda^2 - 2\rho\lambda)/(1-\rho^2)$$
$$\text{Cov}[\varepsilon_t,\varepsilon_{t-1}] = \rho\text{Var}[\varepsilon_{t-1}] + \text{Cov}[\varepsilon_{t-1},u_t] - \lambda\text{Cov}[\varepsilon_{t-1},u_{t-1}].$$

To evaluate this expression, write

$\varepsilon_{t-1} = u_{t-1} + (\rho-\lambda)u_{t-2} + \rho(\rho-\lambda)u_{t-3} + \rho^2(\rho-\lambda)u_{t-4} + ...$

Therefore, the middle term is zero and the third is simply $\lambda\sigma_u^2$. Thus,

$$\text{Cov}[\varepsilon_t,\varepsilon_{t-1}] = \sigma_u^2\{[\rho(1 + \lambda^2 - 2\rho\lambda)]/(1 - \rho^2) - \lambda]\} = \sigma_u^2[(\rho - \lambda)(1 - \lambda\rho)/(1 - \rho^2)]$$

For lags greater than 1, $\text{Cov}[\varepsilon_t,\varepsilon_{t-j}] = \rho\text{Cov}[\varepsilon_{t-1},\varepsilon_{t-j}] + \text{Cov}[\varepsilon_{t-j},u_t] - \lambda\text{Cov}[\varepsilon_{t-j},u_{t-1}]$.

Since ε_{t-j} involves only *u*s up to its current period, ε_{t-j} is uncorrelated with u_t and u_{t-1} if *j* is greater than 1. Therefore, after the first lag, the autocovariances behave in the familiar fashion, $\text{Cov}[\varepsilon_t,\varepsilon_{t-j}] = \rho\text{Cov}[\varepsilon_t,\varepsilon_{t-j+1}]$

The autocorrelation coefficient of the residuals estimates $\text{Cov}[\varepsilon_t,\varepsilon_{t-1}]/\text{Var}[\varepsilon_t] = (\rho - \lambda)(1 - \rho\lambda)/(1 + \lambda^2 - 2\rho\lambda)$.

3. Since the regression contains a lagged dependent variable, we cannot use the Durbin-Watson statistic directly. The *h* statistic in (15-34) would be $h = (1 - 1.21/2)[21 / (1 - 21(.18^2)]^{1/2} = 3.201$. The 95% critical value from the standard normal distribution for this one-tailed test would be 1.645. Therefore, we would reject the hypothesis of no autocorrelation.

4. It is commonly asserted that the Durbin-Watson statistic is only appropriate for testing for first order autoregressive disturbances. What combination of the coefficients of the model is estimated by the Durbin-Watson statistic in each of the following cases: AR(1), AR(2), MA(1)? In each case, assume that the regression model does not contain a lagged dependent variable. Comment on the impact on your results of relaxing this assumption.

In each case, plim $d = 2 - 2\rho_1$ where $\rho_1 = \text{Corr}[\varepsilon_t,\varepsilon_{t-1}]$. The first order autocorrelations are as follows: AR(1): ρ (see (15-9)) and AR(2): $\theta_1/(1 - \theta_2)$. For the AR(2), a proof is as follows: First, $\varepsilon_t = \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + u_t$. Denote $\text{Var}[\varepsilon_t]$ as c_0 and $\text{Cov}[\varepsilon_t,\varepsilon_{t-1}]$ as c_1. Then, it follows immediately that $c_1 = \theta_1 c_0 + \theta_2 c_1$ since u_t is independent of ε_{t-1}. Therefore $\rho_1 = c_1/c_0 = \theta_1/(1 - \theta_2)$. For the MA(1): $-\lambda / (1 + \lambda^2)$ (See (15-43)). To prove this, write $\varepsilon_t = u_t - \lambda u_{t-1}$. Then, since the *u*s are independent, the result follows just by multiplying out $\rho_1 = \text{Cov}[\varepsilon_t,\varepsilon_{t-1}]/\text{Var}[\varepsilon_t] = -\lambda\text{Var}[u_{t-1}]/\{\text{Var}[u_t] + \lambda^2\text{Var}[u_{t-1}]\} = -\lambda/(1 + \lambda^2)$.

Applications

1. Phillips Curve

```
--> date;1950.1$
--> peri;1950.1-2000.4$
--> crea;dp=infl-infl[-1]$
--> crea;dy=loggdp-loggdp[-1]$
--> peri;1950.3-2000.4$
--> regr;lhs=dp;rhs=one,unemp$;ar1;res=u$

+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = DP       Mean=  -.1926996283E-01, S.D.=    2.818214558     |
| Model size: Observations =     202, Parameters =   2, Deg.Fr.=    200 |
| Residuals:  Sum of squares= 1592.321197    , Std.Dev.=        2.82163 |
| Fit:        R-squared=  .002561, Adjusted R-squared =         -.00243 |
| Model test: F[  1,    200] =     .51,    Prob value =          .47449 |
| Diagnostic: Log-L =   -495.1583, Restricted(b=0) Log-L =    -495.4173 |
|             LogAmemiyaPrCrt.=    2.084, Akaike Info. Crt.=      4.922 |
| Autocorrel: Durbin-Watson Statistic =   2.82755,   Rho =      -.41378 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant     .4918922148       .74047944     .664   .5073
 UNEMP    -.9013159906E-01      .12578616    -.717   .4745     5.6712871
--> peri;1951.2-2000.4$
--> regr;lhs=u;rhs=one,u[-1],u[-2]$
```

```
+----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none    |
| Dep. var. = U        Mean=  -.3890391012E-01, S.D.=   2.799476915    |
| Model size: Observations =     199, Parameters =   3, Deg.Fr.=    196 |
| Residuals:  Sum of squares= 1079.052269    , Std.Dev.=       2.34635 |
| Fit:        R-squared=  .304618, Adjusted R-squared =         .29752 |
| Model test: F[  2,    196] =   42.93,    Prob value =         .00000 |
| Diagnostic: Log-L =   -450.5769, Restricted(b=0) Log-L =    -486.7246 |
|             LogAmemiyaPrCrt.=    1.721, Akaike Info. Crt.=      4.559 |
| Autocorrel: Durbin-Watson Statistic =   1.99273,   Rho =       .00363 |
+----------------------------------------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
 Constant -.5048615289E-01      .16633422    -.304   .7618
 U[-1]       -.5946344724   .65920584E-01   -9.020   .0000 -.10234931E-01
 U[-2]       -.3824653303   .65904378E-01   -5.803   .0000 -.14370453E-01
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)
```

--> calc;list;lm=n*rsqrd$

```
    LM      =  .60618960968412850D+02
+------------------------------------------+
| AR(1) Model:     e(t) = rho * e(t-1) + u(t) |
| Initial value of rho       =        -.41378 |
| Maximum iterations         =            100 |
| Method = Prais - Winsten                    |
| Iter=  1, SS=   1299.275, Log-L=-474.710175 |
| Final value of Rho    =            -.413779 |
| Iter=  1, SS=   1299.275, Log-L=-474.710175 |
| Durbin-Watson:   e(t) =            2.827557 |
| Std. Deviation:  e(t) =            2.799716 |
| Std. Deviation:  u(t) =            2.548799 |
| Durbin-Watson:   u(t) =            2.340706 |
| Autocorrelation: u(t) =            -.170353 |
| N[0,1] used for significance levels         |
+------------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
 Constant      .4704274598       .47671946     .987   .3237
 UNEMP    -.8709854633E-01  .80962277E-01   -1.076   .2820     5.6712871
 RHO         -.4137785986   .64213081E-01   -6.444   .0000
```

Regression results are almost unchanged. Autocorrelation of transformed residuals is -.17, less than -.41 in original model.

2. (Improved Phillips curve model)

```
--> crea;newecon=dmy(1974.1,2000.4)$
--> regr;lhs=dp;rhs=one,unemp,newecon;plot$
+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = DP       Mean=  -.1926996283E-01, S.D.=   2.818214558     |
| Model size: Observations =     202, Parameters =   3, Deg.Fr.=    199 |
| Residuals:  Sum of squares= 1586.260338    , Std.Dev.=       2.82332 |
| Fit:        R-squared=  .006357, Adjusted R-squared =         -.00363 |
| Model test: F[  2,    199] =     .64,    Prob value =          .53017 |
| Diagnostic: Log-L =   -494.7731, Restricted(b=0) Log-L =    -495.4173 |
|             LogAmemiyaPrCrt.=    2.091, Akaike Info. Crt.=      4.928 |
| Autocorrel: Durbin-Watson Statistic =   2.83473,   Rho =      -.41737 |
+-----------------------------------------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
 Constant     .5507626279       .74399306     .740   .4600
 UNEMP    -.9835166981E-01      .12621412    -.779   .4368     5.6712871
 NEWECON     -2.474910396       2.8382661    -.872   .3843  .49504950E-02
```

3. (GARCH Models)

.a. We used LIMDEP with the macroeconomics data in table F5.1. The rate of inflation was computed with all observations, then observations 6 to 204 were used to remove the missing data due to lags. Least squares results were obtained first. The residuals were then computed and squared. Using observations 15-204, we then computed a regression of the squared residual on a constant and 8 lagged values. The chi-squared statistic with 8 degrees of freedom is 28.24. The critical value from the table for 95% significance and 8 degrees of freedom is 15.51, so at this level of significance, the hypothesis of no GARCH effects is rejected.

```
crea;pt=100*log(cpi_u/cpi_u[-1])$
crea;pt1=pt[-1];pt2=pt[-2];pt3=pt[-3];pt4=pt[-4]$
samp;6-204$
regr;lhs=pt;rhs=one,pt1,pt2,pt3,pt4;res=et$$
crea;vt=et*et$
crea;vt1=vt[-1];vt2=vt[-2];vt3=vt[-3];vt4=vt[-4];vt5=vt[-5];vt6=vt[-6];vt7=vt[-
7];vt8=vt[-8]$
samp;15-204$
regr;lhs=vt;rhs=one,vt1,vt2,vt3,vt4,vt5,vt6,vt7,vt8$
calc;list;lm=n*rsqrd$
+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = PT       Mean=   .9589185961    , S.D.=   .8318268241     |
| Model size: Observations =     199, Parameters =   5, Deg.Fr.=    194 |
| Residuals:  Sum of squares= 61.97028507    , Std.Dev.=         .56519 |
| Fit:        R-squared=  .547673, Adjusted R-squared =          .53835 |
| Model test: F[  4,    194] =   58.72,    Prob value =          .00000 |
| Diagnostic: Log-L =   -166.2871, Restricted(b=0) Log-L =    -245.2254 |
|             LogAmemiyaPrCrt.=   -1.116, Akaike Info. Crt.=      1.721 |
| Autocorrel: Durbin-Watson Statistic =   1.80740,   Rho =       .09630 |
+-----------------------------------------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
 Constant     .1296044455   .67521735E-01    1.919   .0564
 PT1          .2856136998   .69863942E-01    4.088   .0001     .97399582
 PT2          .1237760914   .70647061E-01    1.752   .0813     .98184918
 PT3          .2516837602   .70327318E-01    3.579   .0004     .99074774
 PT4          .1824670634   .69251374E-01    2.635   .0091     .98781131
 LM      =  .28240022492847690D+02
```

For the second step, we need an estimate of α_0, which is the unconditional variance if there are no ARCH effects. We computed this based on the ARCH specification by a regression of $e_t^2 - (8/36)e_{t-1}^2 - \ldots - (1/36)e_{t-8}^2$ on just a constant term. This produces a negative estimate of α_0, but this is not the variance, so we retain the result. We note, the problem that this reflects is probably the specific, doubtless unduly restrictive, ARCH structure assumed.

```
samp;6-204$
crea;vt=et*et$
crea;ht=vt-8/36*vt[-1]-7/36*vt[-2]-6/36*vt[-3]-5/36*vt[-4]-4/36*vt[-5]-
3/36*vt[-6]-2/36*vt[-7]-1/36*vt[-8]$
samp;15-204$
calc;list;a0=xbr(ht)$
samp;6-204$
crea;qt=a0+8/36*vt[-1]+7/36*vt[-2]+6/36*vt[-3]+5/36*vt[-4]+4/36*vt[-
5]+3/36*vt[-6]+2/36*vt[-7]+1/36*vt[-8]$
samp;15-204$
plot;rhs=qt$
crea;wt=1/qt$
regr;lhs=pt;rhs=one,pt1,pt2,pt3,pt4;wts=wt$
regr;lhs=pt;rhs=one,pt1,pt2,pt3,pt4;model=garch(1,1)$
```

Once we have an estimate of α_0 in hand, we then computed the set of variances according to the ARCH(8) model, using the lagged squared residuals. Finally, we used these variance estimators to compute a weighted least squares regression accounting for the heteroscedasticity. This regression is based on observations 15-204, again because of the lagged values. Finally, using the same sample, a GARCH(1,1) model is fit by maximum likelihood.

```
+-------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = WT      |
| Dep. var. = PT       Mean=   .8006997687    , S.D.=   .6327877239    |
| Model size: Observations =     190, Parameters =   5, Deg.Fr.=    185 |
| Residuals:  Sum of squares= 38.67492770    , Std.Dev.=         .45722 |
| Fit:        R-squared=  .488964, Adjusted R-squared =          .47791 |
| Model test: F[  4,    185] =   44.25,    Prob value =          .00000 |
| Diagnostic: Log-L =   -147.7324, Restricted(b=0) Log-L =    -211.5074 |
|             LogAmemiyaPrCrt.=   -1.539, Akaike Info. Crt.=      1.608 |
| Autocorrel: Durbin-Watson Statistic =   1.90310,   Rho =       .04845 |
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant     .1468553158   .60127085E-01    2.442   .0155
 PT1      .9760051110E-01  .88469908E-01    1.103   .2714     .77755556
 PT2          .3328520370   .86772549E-01    3.836   .0002     .76745308
 PT3          .1428889148   .85420554E-01    1.673   .0961     .76271761
 PT4          .2878686524   .84090832E-01    3.423   .0008     .74173558
```

The 8 period ARCH model produces quite a substantial change in the estimates. Once again, this probably results from the restrictive assumption about the lag weights in the ARCH model. The GARCH model follows.

```
+---------------------------------------------+
| GARCH MODEL                                 |
| Maximum Likelihood Estimates                |
| Model estimated: Jul 31, 2002 at 01:19:14PM.|
| Dependent variable                  PT      |
| Weighting variable                None      |
| Number of observations             190      |
| Iterations completed                22      |
| Log likelihood function      -135.5043      |
| Restricted log likelihood    -147.6465      |
| Chi squared                   24.28447      |
| Degrees of freedom                   2      |
| Prob[ChiSqd > value] =        .5328953E-05 |
| GARCH Model, P = 1, Q = 1                   |
| Wald statistic for GARCH =      521.483     |
+---------------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
          Regression parameters
 Constant     .1308478127   .61887183E-01    2.114   .0345
 PT1          .1749239917   .70912277E-01    2.467   .0136     .98810078
 PT2          .2532191617   .73228319E-01    3.458   .0005     .98160455
 PT3          .1552879436   .68274176E-01    2.274   .0229     .97782066
 PT4          .2751467919   .63910272E-01    4.305   .0000     .97277700
          Unconditional Variance
 Alpha(0)  .1005125676E-01  .11653271E-01     .863   .3884
          Lagged Variance Terms
 Delta(1)     .8556879884   .89322732E-01    9.580   .0000
          Lagged Squared Disturbance Terms
 Alpha(1)     .1077364862   .60761132E-01    1.773   .0762
          Equilibrium variance, a0/[1-D(1)-A(1)]
 EquilVar     .2748082674       2.0559946     .134   .8937
```

Chapter 20

Models with Lagged Variables

Exercises

1. For the first, the mean lag is $.55(.02)(0) + .55(.15)(1) + ... + .55(.17)(4) = 1.31$ periods. The impact multiplier is $.55(.02) = .011$ while the long run multiplier is the sum of the coefficients, .55.

For the second, the coefficient on x_t is .6, so this is the impact multiplier. The mean lag is found by applying (18-9) to $B(L) = [.6 + 2L]/[1 - .6L + .5L^2] = A(L)/D(L)$. Then, $B'(1)/B(1) = \{[D(1)A'(1) - A(1)D'(1)]/[D(1)]^2\} / [A(1)/D(1)] = A'(1)/A(1) - D'(1)/D(1) = (2/2.6) / (.4/.9) = 1.731$ periods. The long run multiplier is $B(1) = 2.6/.9 = 2.888$ periods.

For the third, since we are interested only in the coefficients on x_t, write the model as $y_t = \alpha + \beta x_t[1 + \gamma L + \gamma^2 L^2 + ...] + \delta z_t^* + u_t$. The lag coefficients on x_t are simply β times powers of γ.

2. We would regress y_t on a constant, $x_t, x_{t-1}, ..., x_{t-6}$. Constrained least squares using

$$\mathbf{R} = \begin{bmatrix} 1 & -5 & 10 & -10 & 5 & -1 & 0 & 0 \\ 0 & 1 & -5 & 10 & -10 & 5 & -1 & 0 \\ 0 & 0 & 1 & -5 & 10 & -10 & 5 & -1 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

would produce the PDL estimates.

3. The ratio of polynomials will equal $B(L) = [.6 + 2L]/[1 - .6L + .5L^2]$. This will expand to $B(L) = \beta_0 + \beta_1 L + \beta_2 L^2 + ...$ Multiply both sides of the equation by $(1 - .6L + .5L^2)$ to obtain $(\beta_0 + \beta_1 L + \beta_2 L^2 + ...)(1 - .6L + .5L^2) = .6 + 2L$. Since the two sides must be equal, it follows that $\beta_0 = .6$ (the only term not involving L) $-.6\beta_0 + \beta_1 = 2$ (the only term involving only L. Therefore, $\beta_1 = 2.36$. All remaining terms, involving $L^2, L^3, ...$ must equal zero. Therefore, $\beta_j - .6\beta_{j-1} + .5\beta_{j-2} = 0$ for all $j > 1$, or $\beta_j = .6\beta_{j-1} - .5\beta_{j-2}$. This provides a recursion for all remaining coefficients. For the specified coefficients, $\beta_2 = .6(2.36) - .5(.3) = 1.266$. $\beta_3 = .6(1.266) - .5(2.36) = -.4204$, $\beta_4 = .6(-.4204) - .5(1.266) = -.88524$ and so on.

4. By multiplying through by the denominator of the lag function, we obtain an autoregressive form

$$\begin{aligned} y_t &= \alpha(1+\delta_1+\delta_2) + \beta x_t + \gamma x_{t-1} - \delta_1 y_{t-1} - \delta_2 y_{t-2} + \varepsilon_t + \delta_1\varepsilon_{t-1} + \delta_2\varepsilon_{t-2} \\ &= \alpha(1+\delta_1+\delta_2) + \beta x_t + \gamma x_{t-1} - \delta_1 y_{t-1} - \delta_2 y_{t-2} + v_t \end{aligned}$$

The model cannot be estimated consistently by ordinary least squares because there is autocorrelation in the presence of a lagged dependent variable. There are two approaches possible. Nonlinear least squares could be applied to the moving average (distributed lag) form. This would be fairly complicated, though a method of doing so is described by Maddala and Rao (1973). A much simpler approach would be to estimate the model in the autoregressive form using an instrumental variables estimator. The lagged variables x_{t-2} and x_{t-3} can be used for the lagged dependent variables. ~

5. The model can be estimated as an autoregressive or distributed lag equation. Consider, first, the autoregressive form. Multiply through by $(1 - \gamma L)(1 - \phi L)$ to obtain

$$y_t = \alpha(1-\gamma)(1-\phi) + \beta x_t - (\beta\phi)x_{t-1} + \delta z_t - (\delta\gamma)z_{t-1} + (\gamma + \phi)y_{t-1} - (\gamma\phi)y_{t-2} + \varepsilon_t - (\gamma+\phi)\varepsilon_{t-1} + (\gamma\phi)\varepsilon_{t-2}.$$

Clearly, the model cannot be estimated by ordinary least squares, since there is an autocorrelated disturbance and a lagged dependent variable. The parameters can be estimated consistently, but inefficiently by linear instrumental variables. The inefficiency arises from the fact that the parameters are overidentified. The linear estimator estimates seven functions of the five underlying parameters. One possibility is a GMM estimator. Let $v_t = \varepsilon_t - (\gamma+\phi)\varepsilon_{t-1} + (\gamma\phi)\varepsilon_{t-2}$. Then, a GMM estimator can be defined in terms of, say, a set of moment equations of the form $E[v_t w_t] = 0$, where w_t is current and lagged values of x and z. A minimum distance estimator could then be used for estimation.

The distributed lag approach might be taken, instead. Each of the two regressors produces a recursions $x_t^* = x_t + \gamma x_{t-1}^*$ and $z_t^* = z_t + \gamma z_{t-1}^*$. Thus, values of the moving average regressors can be built up recursively. Note that the model is linear in 1, x_t^*, and z_t^*. Therefore, an approach is to search a grid of values of (γ,ϕ) to minimize the sum of squares. ~

Applications

1. The long run multiplier is $\beta_0 + \beta_1 + ... + \beta_6$. The model is a classical regression, so it can be estimated by ordinary least squares. The estimator of the long run multiplier would be the sum of the least squares coefficients. If the sixth lag is omitted, then the standard omitted variable result applies, and all the coefficients are biased. The orthogonality result needed to remove the bias explicitly fails here, since x_t is an AR(1) process. All the lags are correlated. Since the form of the relationship is, in fact, known, we can derive the omitted variable formula. In particular, by construction, x_t will have mean zero. By implication, y_t will also, so we lose nothing by assuming that the constant term is zero. To save some cumbersome algebra, we'll also assume with no loss of generality that the unconditional variance of x_t is 1. Let $X_1 = [x_t,x_{t-1},...,x_{t-5}]$ and $X_2 = x_{t-6}$. Then, for the regression of y on X_1, we have by the omitted variable formula,

$$E\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} | X_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} + \begin{bmatrix} 1 & r & r^2 & r^3 & r^4 & r^5 \\ r & 1 & r & r^2 & r^3 & r^4 \\ r^2 & r & 1 & r & r^2 & r^3 \\ r^3 & r^2 & r & 1 & r & r^2 \\ r^4 & r^3 & r^2 & r & 1 & r \\ r^5 & r^4 & r^3 & r^2 & r & 1 \end{bmatrix}^{-1} \begin{bmatrix} r^6 \\ r^5 \\ r^4 \\ r^3 \\ r^2 \\ r \end{bmatrix} \beta_6$$

We can derive a formal solution to the bias in this estimator. Note that the column that is to the right of the inverse matrix is r times the last column matrix. Therefore, the matrix product is r times the last column of an identity matrix. This gives us the complete result,

$$E\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} | X_1 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ r \end{bmatrix} \beta_6.$$

Therefore, the first 5 coefficients are unbiased, and the last one is an estimator of $\beta_5 + r\beta_6$. Adding these up, we see that when the last lag is omitted from the model, the estimator of the long run multiplier is biased downware by $(1-r)\beta_6$. For part d, we will use a similar construction. But, now there are five variables in X_1 and x_{t-5} and x_{t-6} in X_2. The same kind of computation will show that the first four coefficients are unbiased while the fifth now estimates $\beta_4 + r\beta_5 + r^2\beta_6$. The long run multiplier is estimated with downward bias equal to $(1-r)\beta_5 + (1-r^2)\beta_6$.

```
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 XT           .9726595701      1.9258818      .505   .6141     8.3384522
 XT1          .7709686332      3.1555811      .244   .8072     8.3301663
 XT2          .5450409860      3.1761465      .172   .8639     8.3218191
 XT3        -.6061007409      3.1903388     -.190   .8495     8.3134324
 XT4        -.2272352746      3.1729930     -.072   .9430     8.3050260
 XT5        -1.916555094      3.1414210     -.610   .5425     8.2964570
 XT6         1.218771893      1.8814874      .648   .5179     8.2878393
Matrix LRM      has  1 rows and  1 columns.
              1
        +--------------
```

```
      1|      .7575
 XT           1.101551478       1.9126777     .576   .5653     8.3384522
 XT1           .6941982792      3.1485851     .220   .8257     8.3301663
 XT2           .5287939572      3.1712435     .167   .8677     8.3218191
 XT3         -.7300170198       3.1797815    -.230   .8187     8.3134324
 XT4         -.5552651191       3.1275848    -.178   .8593     8.3050260
 XT5         -.2826674399       1.8697065    -.151   .8800     8.2964570
Matrix LRM      has  1 rows and  1 columns.
               1
        +--------------
      1|      .7566
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 XT           1.077633667       1.9012923     .567   .5715     8.3384522
 XT1           .7070443138      3.1394606     .225   .8221     8.3301663
 XT2           .5633400685      3.1549830     .179   .8585     8.3218191
 XT3         -.6608149939       3.1386871    -.211   .8335     8.3134324
 XT4         -.9304013056       1.8990464    -.490   .6247     8.3050260
Matrix LRM      has  1 rows and  1 columns.
               1
        +--------------
      1|      .7568
--> calc;list;cor(xt,xt1)$
    Result  =  .99978740920470700D+00
```

The results of the three suggested regressions are shown above. We used observations 7 - 204 of the logged real investment and real GDP data in deviations from the means for all regressions. Note that although there are some large changes in the estimated individual parameters, the long run multiplier is almost identical in all cases. Looking at the analytical results we can see why this would be the case. The correlation between current and lagged log gdp is r = 0.9998. Therefore, the biases that we found, $(1-r)\beta_6$ and $(1-r)\beta_5 + (1-r^2)\beta_6$ are trivial.

2. Because the model has both lagged dependent variables and autocorrelated disturbances, ordinary least squares will be inconsistent. Consistent estimates could be obtained by the method of instrumental variables. We can use x_{t-1} and x_{t-2} as the instruments for y_{t-1} and y_{t-2}. Efficient estimates can be obtained by a two step procedure. We write the model as $y_t - \rho y_{t-1} = \alpha(1-\rho) + \beta(x_t - \rho x_{t-1}) + \gamma(y_{t-1} - \rho y_{t-2}) + \delta(y_{t-2} - \rho y_{t-3}) + u_t$. With a consistent estimator of ρ, we could use FGLS. The residuals from the *IV* estimator can be used to estimate ρ. Then OLS using the transformed data is asymptotically equivalent to GLS. The method of Hatanaka discussed in the text is another possibility.

Using the real consumption and real disposable income data in Table F5.1, we obtain the following results: Estimated standard errors are shown in parentheses. (The estimated autocorrelation based on the IV estimates is .9172.) All three sets of estimates are based on the last 201 observations, 1950.4 to 2000.4

	OLS	IV	2 Step FGLS
$\hat{\alpha}$	-1.4946	-64.5073	-4.6614
	(3.8291)	(46.1075)	(3.2041)
$\hat{\beta}$	.007569	.7003	.3477
	(.001662)	(.4910)	(.0432)
$\hat{\gamma}$	1.1977	.5726	.2332
	(.006921)	(.9043)	(.05933)
$\hat{\delta}$	-0.1988	-.3324	.4072
	(.07109)	(.4962)	(.05500)

Chapter 21

Time Series Models

There are no exercises or applications in Chapter 21.

Chapter 22

Nonstationary Data

Exercise

1. The autocorrelations are simple to obtain just by multiplying out v_t^2, v_tv_{t-1} and so on. The autocovariances are $1+\theta_1^2 + \theta_2^2$, $-\theta_2(1 - \theta_2)$, $-\theta_2$, 0, 0, 0... which provides the autocorrelations by division by the first of these. The partial autocorrelations are messy, and can be obtained by the Yule Walker equations. Alternatively (and much more simply), we can make use of the observation in Section 21.2.3 that the partial autocorrelations for the MA(2) process mirror tha autocorrelations for an AR(2). Thus, the results in Section 21.2.3 for the AR(2) can be used directly.

Applications

1. ADF Test

```
+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = R        Mean=   8.212678571    , S.D.=   .7762719558     |
| Model size: Observations =      56, Parameters =   6, Deg.Fr.=     50 |
| Residuals:  Sum of squares= .9651001703    , Std.Dev.=         .13893 |
| Fit:        R-squared=  .970881, Adjusted R-squared =          .96797 |
| Model test: F[  5,     50] =  333.41,    Prob value =          .00000 |
| Diagnostic: Log-L =     34.2439, Restricted(b=0) Log-L =     -64.7739 |
|             LogAmemiyaPrCrt.=   -3.846, Akaike Info. Crt.=     -1.009 |
| Autocorrel: Durbin-Watson Statistic =   1.91589,   Rho =       .04205 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant     .2565690959       .47172815      .544   .5889
 T         .4401352136E-03  .25092142E-02      .175   .8615     32.500000
 R1           .9653227410   .48183346E-01   20.034   .0000     8.2305357
 DR1          .5600009441       .14342088    3.905   .0003 -.12321429E-01
 DR2         -.1739775168       .14781417   -1.177   .2448 -.20535714E-01
 DR3      -.7792177815E-03      .11072916    -.007   .9944 -.11607143E-01
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)
```

--> wald;fn1=b_r1-1$

```
+-----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic              =       .51796    |
| Prob. from Chi-squared[ 1] =       .47171     |
+-----------------------------------------------+
+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
 Fncn(1)  -.3467725900E-01  .48183346E-01    -.720   .4717
```

The unit root hypothesis is definitely not rejected.

2. Macroeconomic Model

```
--> samp;1-204$
--> crea;c=log(realcons);y=log(realdpi)$
--> crea;c1=c[-1];c2=c[-2]$
--> samp;3-204$
--> regr;lhs=c;rhs=one,y,c1,c2$
+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = C        Mean=   7.889033683    , S.D.=   .5102401315     |
| Model size: Observations =     202, Parameters =   4, Deg.Fr.=    198 |
| Residuals:  Sum of squares= .1519097328E-01, Std.Dev.=         .00876 |
| Fit:        R-squared=  .999710, Adjusted R-squared =          .99971 |
| Model test: F[  3,    198] =********,    Prob value =          .00000 |
| Diagnostic: Log-L =    672.4019, Restricted(b=0) Log-L =    -150.2038 |
|             LogAmemiyaPrCrt.=   -9.456, Akaike Info. Crt.=     -6.618 |
| Autocorrel: Durbin-Watson Statistic =   1.89384,   Rho =       .05308 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  .8165780259E-03  .10779352E-01     .076   .9397
 Y         .7869591065E-01  .29020268E-01    2.712   .0073     7.9998985
 C1           .9680839747   .72732869E-01   13.310   .0000     7.8802520
 C2      -.4701660339E-01  .70076193E-01    -.671   .5030     7.8714299
--> crea;e1=e[-1];e2=e[-3];e3=e[-3]$
--> crea;e1=e[-1];e2=e[-2];e3=e[-3]$
--> regr;lhs=e;rhs=one,e1,e2,e3$
+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = E        Mean=  -.6947138134E-15, S.D.=   .8693502258E-02 |
| Model size: Observations =     202, Parameters =   4, Deg.Fr.=    198 |
| Residuals:  Sum of squares= .1339943625E-01, Std.Dev.=         .00823 |
| Fit:        R-squared=  .117934, Adjusted R-squared =          .10457 |
| Model test: F[  3,    198] =    8.82,    Prob value =          .00002 |
| Diagnostic: Log-L =    685.0763, Restricted(b=0) Log-L =     672.4019 |
|             LogAmemiyaPrCrt.=   -9.581, Akaike Info. Crt.=     -6.743 |
| Autocorrel: Durbin-Watson Statistic =   1.85371,   Rho =       .07314 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  .2437121418E-04  .57884755E-03     .042   .9665
 E1      -.2553462753E-01  .70917392E-01    -.360   .7192 -.21497022E-04
 E2           .3385045374   .66904365E-01    5.060   .0000 -.56959898E-04
 E3        .6894158132E-01  .71101163E-01     .970   .3334 -.81793147E-04

--> calc;list;chisq=n*rsqrd$
    CHISQ   =  .23822731697405480D+02

Matrix Result   has  2 rows and  2 columns.
               1             2
        +---------------------------
       1|     1.0688     .0000000D+00
       2|    19.8378     .0000000D+00
```

Short run multiplier is $\beta = .07869$. Long run is $\beta/(1-\gamma_1 - \gamma_2) = 12.669$. (Not very plausible.)

3. ADF Test. To carry out the test, the rate of inflation is regressed on a constant, a time trend, the previous year's value of the rate of inflation, and three lags of the first difference. The test statistic for the ADF is (0.7290534455-1)/.011230759 = -2.373. The critical value in the lower part of Table 20.4 with about 100 observations is -3.45. Since our value is large than this, it follows that the hypothesis of a unit root cannot be rejected.

4. Reestimated model in example 13.1.

```
--> samp;1-204$
--> crea;ddp1=infl[-1]-infl[-2]$
--> crea;ddp2=ddp1[-1]$
--> crea;ddp3=ddp1[-2]$
--> crea;dp=infl[-1]$
--> samp;97-204$
--> crea;t=trn(1,1)$
--> regr;lhs=infl;rhs=one,t,dp,ddp1,ddp2,ddp3$
+-----------------------------------------------------------------------+
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = INFL     Mean=   4.907672727    , S.D.=   3.617392978     |
| Model size: Observations =     108, Parameters =   6, Deg.Fr.=    102 |
| Residuals:  Sum of squares= 608.5020156    , Std.Dev.=        2.44248 |
| Fit:        R-squared=  .565403, Adjusted R-squared =          .54410 |
| Model test: F[  5,    102] =   26.54,    Prob value =          .00000 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |t-ratio |P[|T|>t] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant     2.226039717       1.1342702    1.963   .0524
 T        -.1836785769E-01  .11230759E-01   -1.635   .1050     54.500000
 DP           .7290534455       .11419140    6.384   .0000     4.9830886
 DDP1        -.4744389916       .12707255   -3.734   .0003 -.58569323E-01
 DDP2        -.4273030624       .11563482   -3.695   .0004 -.46827528E-01
 DDP3        -.2248432703   .98954483E-01   -2.272   .0252 -.86558444E-02
--> wald;fn1=b_dp-1$
+---------+--------------+----------------+--------+---------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] |
+---------+--------------+----------------+--------+---------+
 Fncn(1)     -.2709465545       .11419140   -2.373   .0177
--> samp;1-204$
--> crea;ct=realcons;yt=realgdp;gt=realgovt;rt=tbilrate$
--> crea;ct1=ct[-1];yt1=yt[-1]$
--> samp;2-204$
--> samp;1-204$
--> crea;ct=realcons;yt=realgdp;gt=realgovt;rt=tbilrate;it=realinvs$
--> crea;ct1=ct[-1];yt1=yt[-1]$
--> crea;dy=yt-yt1$
--> samp;2-204$
--> name;x=one,rt,ct1,yt1,gt$
--> 2sls;lhs=ct;rhs=one,yt,ct1;inst=x;res=ec$
--> 2sls;lhs=it;rhs=one,rt,dy;inst=x;res=ei$
--> iden;rhs=ec;pds=10$
--> iden;rhs=ei;pds=10$
+-----------------------------------------------------------------------+
| Two stage   least squares regression    Weighting variable = none     |
| Dep. var. = CT       Mean=   3008.995074    , S.D.=   1456.900152     |
| Model size: Observations =     203, Parameters =   3, Deg.Fr.=    200 |
| Residuals:  Sum of squares= 96595.67529    , Std.Dev.=       21.97677 |
| Fit:        R-squared=  .999771, Adjusted R-squared =          .99977 |
|             (Note:  Not using OLS.  R-squared is not bounded in [0,1] |
| Model test: F[  2,    200] =********,    Prob value =          .00000 |
| Diagnostic: Log-L =   -913.8005, Restricted(b=0) Log-L =   -1766.2087 |
|             LogAmemiyaPrCrt.=    6.195, Akaike Info. Crt.=      9.033 |
| Autocorrel: Durbin-Watson Statistic =   1.61078,   Rho =       .19461 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
```

```
+---------+-------------+---------------+--------+---------+----------+
 Constant     6.666079115       8.6211817     .773   .4394
 YT       -.2932041745E-01  .35260653E-01    -.832   .4057     4577.1882
 CT1          1.051478712   .51482187E-01   20.424   .0000     2982.9744

+----------------------------------------------------------------------+
| Two stage   least squares regression    Weighting variable = none     |
| Dep. var. = IT       Mean=   654.5295567    , S.D.=   391.3705005     |
| Model size: Observations =     203, Parameters =   3, Deg.Fr.=    200 |
| Residuals:  Sum of squares= 54658669.31    , Std.Dev.=      522.77466 |
| Fit:        R-squared= -.793071, Adjusted R-squared =         -.81100 |
|             (Note:  Not using OLS.  R-squared is not bounded in [0,1] |
| Diagnostic: Log-L =  -1557.1409, Restricted(b=0) Log-L =   -1499.3832 |
|             LogAmemiyaPrCrt.=   12.533, Akaike Info. Crt.=     15.371 |
| Autocorrel: Durbin-Watson Statistic =   1.49055,   Rho =       .25473 |
+----------------------------------------------------------------------+
+---------+-------------+---------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+---------------+--------+---------+----------+
 Constant    -141.8297176       103.57113   -1.369   .1709
 RT           52.04340559       12.971223    4.012   .0001     5.2499007
 DY           13.80361384       1.7499250    7.888   .0000     37.898522
Time series identification for EC
Box-Pierce Statistic =    40.8498      Box-Ljung Statistic  =    41.7842
Degrees of freedom   =         10      Degrees of freedom   =         10
Significance level   =      .0000      Significance level   =      .0000
* => |coefficient| > 2/sqrt(N) or > 95% significant.
PACF is computed using Yule-Walker equations.
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
Lag |  Autocorrelation Function     |Box/Prc|    Partial Autocorrelations   X
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
  1 | .194*|           |**          |  7.65*| .194*|            |**         X
  2 | .264*|           |***         | 21.82*| .236*|            |***        X
  3 | .273*|           |***         | 36.93*| .207*|            |**         X
  4 | .067 |           |*           | 37.85*|-.063 |          * |           X
  5 | .054 |           |*           | 38.44*|-.068 |          * |           X
  6 | .073 |           |*           | 39.52*| .018 |            |*          X
  7 | .009 |           |*           | 39.53*| .003 |            |*          X
  8 |-.078 |          *|            | 40.78*|-.109 |          * |           X
  9 | .019 |           |*           | 40.85*| .023 |            |*          X
 10 | .002 |           |*           | 40.85*| .050 |            |*          X
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
Time series identification for EI
Box-Pierce Statistic =    27.4753      Box-Ljung Statistic  =    28.3566
Degrees of freedom   =         10      Degrees of freedom   =         10
Significance level   =      .0022      Significance level   =      .0016
* => |coefficient| > 2/sqrt(N) or > 95% significant.
PACF is computed using Yule-Walker equations.
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
Lag |  Autocorrelation Function     |Box/Prc|    Partial Autocorrelations   X
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
  1 | .244*|           |***         | 12.13*| .244*|            |***        X
  2 | .143*|           |**          | 16.27*| .096 |            |*          X
  3 | .037 |           |*           | 16.55*|-.019 |          * |           X
  4 |-.001 |          *|            | 16.55*|-.017 |          * |           X
  5 |-.066 |          *|            | 17.42*|-.078 |          * |           X
  6 | .003 |           |*           | 17.43*| .043 |            |*          X
  7 |-.042 |          *|            | 17.79*|-.033 |          * |           X
  8 |-.107 |          *|            | 20.10*|-.107 |          * |           X
  9 | .108 |           |*           | 22.46*| .194*|            |**         X
 10 | .157*|           |**          | 27.48*| .142*|            |**         X
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

Chapter 23
Models for Discrete Choice

Exercises

1. The log-likelihood is

$\ln L = \Sigma_{0,0}\ln\text{Prob}[y=0,d=0] + \Sigma_{0,1}\ln\text{Prob}[y=0,d=1] + \Sigma_{1,0}\ln\text{Prob}[y=1,d=0] + \Sigma_{1,1}\ln\text{Prob}[y=1,d=1]$

where $\Sigma_{i,j}$ indicates the sum over observations for which $y = i$ and $d = j$. Since there are no other regressors, this reduces to $\ln L = 24\ln(1 - F(\alpha)) + 32\ln(1 - F(\delta)) + 28\ln F(\alpha) + 16\ln F(\delta)$. Although it is straightforward to maximize the log-likelihood directly in terms of α and δ, an alternative, convenient approach is to estimate $F(\alpha)$ and $F(\delta)$. These functions can then be inverted to estimate the original parameters. The invariance of maximum likelihood estimators to transformation will justify this approach. One virtue of this approach is that the same procedure is used for both probit and logit models. Let $A = F(\alpha)$ and $D = F(\delta)$. Then, the log likelihood is simply $\ln L = 24\ln(1 - A) + 32\ln(1 - D) + 28\ln A + 16\ln D$. The necessary conditions are

$$\partial\ln L/\partial A = -24/(1 - A) + 28/A = 0$$
$$\partial\ln L/\partial D = -32/(1 - D) + 16/D = 0.$$

Simple manipulations produce the two solutions $A = 28/(24+28) = .539$ and $D = 16/(32+16) = .333$. Then, these functions can be inverted to produce the MLEs of α and β. Thus, $\hat{\alpha} = F^{-1}(A)$ and $\hat{\beta} = F^{-1}(D) - \hat{\alpha}$. The two inverse functions are $\Phi^{-1}(A)$ for the probit model, which must be approximated, and ln[F/(1-F)] for the logit model. The estimates are,

	Probit	Logit
α	.098	.156
δ	-.431	-.694
β	-.529	-.850

(Notice the proportionality relationship, .156/.098 = 1.592 and -.848/-.529 = 1.607.)

We will compute the asymptotic covariance matrix for $\hat{\alpha}$ and $\hat{\beta}$ directly using (19-24) for the probit model and (19-22) for the logit model. We will require $h_i = \partial^2\ln L/\partial(\alpha+\beta d)^2$ for the four cells. For the computation, we will require $\phi(c)/\Phi(c)$ and $-\phi(c)/[1-\Phi(c)]$. These are listed in the table below.

y	d	$\alpha+\beta d$	Φ	ϕ	λ_1 ϕ/Φ	λ_0 $-\phi/(1-\Phi)$	$\lambda_0\lambda_1$
0	0	.098	.539	.397	.737	-.861	-.636
1	0	.098	.539	.397	.737	-.861	-.636
0	1	-.431	.333	.364	1.093	-.546	-.597
1	1	-.431	.333	.364	1.093	-.546	-.597

The estimated asymptotic covariance matrix is the inverse of the estimate of $-E[\mathbf{H}]$.

$$-\hat{\mathbf{H}} = 24(.636)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + 28(.636)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + 32(.597)\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} + 16(.597)\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}.$$ Then,

$$\left[-\hat{\mathbf{H}}\right]^{-1} = \begin{bmatrix}61.728 & 28.656\\28.656 & 28.656\end{bmatrix}^{-1} = \begin{bmatrix}.03024 & -.03024\\-.03024 & .06513\end{bmatrix}.$$ The asymptotic standard errors are the square roots of the diagonal elements, which are .1739 and .2552, respectively. To test the hypothesis that β = 0, we would refer $z = -.529 / .2552 = -2.073$ to the standard normal table. This is larger than the 1.96 critical value, so we would reject the hypothesis. To compute the likelihood ratio statistic, we will require the two log-likelihoods. The restricted log-likelihood (for both the probit and logit models) is given in (19-28). This would be

$\ln L_0 = 100[.44\ln.44 + .56\ln.56] = -68.593$. Let the predicted values above be denoted

P_{00} = Prob[y=0,d=0] = .461 (i.e., 1 - .539)
P_{10} = Prob[y=1,d=0] = .539
P_{01} = Prob[y=0,d=1] = .667
P_{11} = Prob[y=0,d=1] = .333

and let n_{ij} equal the number of observations in each cell Then, the unrestricted log-likelihood is $\ln L$ = 24ln.461 + 28ln.539 + 32ln.667 + 16ln.333 = -66.442. The likelihood ratio statistic would be λ = -2(-66.6442 - (-68.593)) = 4.302. The critical value from the chi-squared distribution with one degree of freedom is 3.84, so once again, the test statistic is slightly larger than the table value.

We now compute the Hessian for the logit model. The predicted probabilities are

$$\begin{aligned}
\text{Prob}[y = 0, d = 0] &= P_{00} = 1/(1 + e^{.156}) &&= .462\\
\text{Prob}[y = 1, d = 0] &= P_{10} = 1 - P_{00} &&= .538\\
\text{Prob}[y = 0, d = 1] &= P_{01} = 1/(1 + e^{-.431}) &&= .667\\
\text{Prob}[y = 1, d = 1] &= P_{11} = 1 - P_{01} &&= .333.
\end{aligned}$$

Notice that in spite of the quite different coefficients, these are identical to the results for the probit model. Remember that we originally estimated the probabilities, not the parameters, and these were independent of the distribution. Then, the Hessian is computed in the same manner as for the probit model using $h_{ij} = F_{ij}(1-F_{ij})$ instead of $\lambda_0\lambda_1$ in each cell. The asymptotic covariance matrix is the inverse of

$$(28+24)(.462)(.538)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}+(32+16)(.667)(.333)\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}.$$

The standard errors are .2782 and .4137. For testing the hypothesis that β equals zero, the t-statistic is z = -.850/.4137 = -2.055, which is almost the same as that for the probit model. The unrestricted log-likelihood is $\ln L$ = 24ln.4285 + ... + 16ln.3635 = -66.442 (again). The chi-squared statistic is 4.302, as before. □

2. Using the usual regression statistics, we would have $a = \bar{y} - b\bar{x}$, $b = \Sigma_i(x_i - \bar{x})(y_i - \bar{y})/\Sigma_i(x_i - \bar{x})^2$.
For data in which y is a binary variable, we can decompose the numerator somewhat further. First, divide both numerator and denominator by the sample size. Second, since only one variable need be in deviation form, drop the deviation in x. That leaves $b = \left[\Sigma_i x_i(y_i - \bar{y})/n\right]/\left[\Sigma_i(x_i - \bar{x})^2/n\right]$. The denominator is the sample variance of x. Since y_i is only 0s and 1s, $\bar{y}$ is the proportion of 1s in the sample, P. Thus, the numerator is

$(1/n)\Sigma_i x_i y_i - (1/n)\Sigma_i x_i \bar{y} = (1/n)\Sigma_1 x_i - P\bar{x} = (n_1/n)\bar{x}_1 - P[P\bar{x} + (1-P)\bar{x}_0] = P(1-P)(\bar{x}_1 - \bar{x}_0)$.

Therefore, the regression is essentially measuring how much the mean of x varies across the two groups of observations. The constant term does not simplify into any intuitively useful form.

3. The model was estimated using Newton's method as described in the text. The estimated coefficients and their standard are shown below:

$$\hat{y}^* = -.51274 + .15964X$$

(1.042) (.202)

Log-likelihood = -6.403 Restricted log-likelihood = -6.9315.

The t-ratio for testing the hypothesis is .15964/.202 = .79. The chi-squared for the likelihood ratio test is 1.057. Neither is large enough to lead to rejection of the hypothesis.

4. The derivatives of the log-likelihood are given in (23-18)-(23-21). If all coefficients except the constant term are zero, then the first order condition for maximizing the log-likelihood would be $\partial\ln L/\partial\beta = \Sigma_i(y_i - \lambda)(1) = 0$ since with no regressors, λ_i will not vary with i. This leads to the constrained maximum $\hat{\lambda} = \Sigma_i y_i/n = P$, which might be expected. Thus, we estimate the constant term such that $P = \dfrac{\exp(\hat{\alpha})}{1+\exp(\hat{\alpha})}$, or $\hat{\alpha}$ = logit(P). The LM statistic based on the BHHH estimator of the covariance matrix of the first derivatives would be $LM = [\Sigma_i\mathbf{g}_i]'[\Sigma_i\mathbf{g}_i\mathbf{g}_i']^{-1}[\Sigma_i\mathbf{g}_i]$ where $\mathbf{g}_i = \Sigma_i(y_i - P)\mathbf{x}_i$. In full, the statistic is

$$LM = [\Sigma_i(y_i - P)\mathbf{x}_i]'[\Sigma_i(y_i - P)^2\mathbf{x}_i\mathbf{x}_i']^{-1}[\Sigma_i(y_i - P)\mathbf{x}_i].$$

The actual (and expected) Hessian can be used instead by replacing $(y_i - P)^2$ with $P(1 - P)$ in the inverse matrix. The statistic could then be written

$$LM = [\mathbf{X}'(\mathbf{y} - P\mathbf{i})]'[(\mathbf{X}'\mathbf{X})^{-1}][\mathbf{X}'(\mathbf{y} - P\mathbf{i})]/P(1-P) = \mathbf{e}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}/P(1-P)$$

In the preceding, $\mathbf{e}'\mathbf{e} = \Sigma_i(y_i - P)^2 = nP(1 - P)$. Therefore, $LM = n[\mathbf{e}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}/\mathbf{e}'\mathbf{e}]$, which establishes the result. To compute the statistic, we regress $(y_i - P)$ on the $\mathbf{x}$s, then carry nR^2 into the chi-squared table.

5. Since there is no regressor, we may write the log-likelihood as

$$\ln L = 50\ln\Phi(-\alpha) + 40\ln[\Phi(\mu_1-\alpha) - \Phi(-\alpha)] + 45\ln[\Phi(\mu_2-\alpha) - \Phi(\mu_1-\alpha)] +$$

$$80\ln[\Phi(\mu_3-\alpha) - \Phi(\mu_2-\alpha)] + 35\ln[1 - \Phi(\mu_3-\alpha)].$$

There are four unknown parameters to estimate and four free probabilities. Suppose, then, we treat $\Phi(-\alpha)$, $\Phi(\mu_1-\alpha)$, $\Phi(\mu_2-\alpha)$, and $\Phi(\mu_3-\alpha)$ as the unknown parameters, π_0, π_1, π_2, and π_3, respectively. If we can find estimators of these, we can solve for the underlying parameters. We may write the log-likelihood as

$$\ln L = 50\ln\pi 0 + 40\ln(\pi 1 - \pi 0) + 45\ln(\pi_2 - \pi_1) + 80\ln(\pi_3 - \pi_2) + 35\ln(1 - \pi_3).$$

The necessary conditions are

$$\begin{aligned}
\partial\ln L/\partial\pi_0 &= 50/\pi_0 - 40/(\pi_1-\pi_0) &&= 0\\
\partial\ln L/\partial\pi_1 &= 40/(\pi_1 - \pi_0) - 45/(\pi_2 - \pi_1) &&= 0\\
\partial\ln L/\partial\pi_2 &= 45/(\pi_2 - \pi_1) - 80/(\pi_3 - \pi_2) &&= 0\\
\partial\ln L/\partial\pi_3 &= 80/(\pi_3 - \pi_2) - 35/(1 - \pi_3) &&= 0.
\end{aligned}$$

By a simple rearrangement, these can be recast as a set of linear equations. Thus,

$$\begin{aligned}
90\pi_0 - 50\pi_1 &= 0\\
45\pi_0 - 85\pi_1 + 40\pi_2 &= 0\\
80\pi_1 - 125\pi_2 + 45\pi_3 &= 0\\
- 35\pi_2 + 115\pi_3 &= 80
\end{aligned}$$

or

$$\begin{bmatrix} 90 & -50 & 0 & 00 \\ 45 & -85 & 40 & 0 \\ 0 & 80 & -125 & 45 \\ 0 & 0 & -35 & 115 \end{bmatrix} \begin{bmatrix} \pi_0 \\ \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 80 \end{bmatrix}$$

The solution (as might be expected) is

$$\begin{aligned}
\pi_0 &= .2 \quad (50/250)\\
\pi_1 &= .36 \quad ((50+40)/250)\\
\pi_2 &= .54 \quad ((50+40+45)/250)\\
\pi_3 &= .86 \quad ((50+40+45+80)/250).
\end{aligned}$$

Now, we can solve for the underlying parameters.

$$\begin{aligned}
-\alpha &= \Phi^{-1}(.2) = -.841, \text{ so } \alpha = .841.\\
\mu_1-\alpha &= \Phi^{-1}(.36) = -.358, \text{ so } \mu_1 = .483\\
\mu_2-\alpha &= \Phi^{-1}(.54) = .101, \text{ so } \mu_2 = .942\\
\mu_3-\alpha &= \Phi^{-1}(.86) = 1.081, \text{ so } \mu_3 = 1.922.
\end{aligned}$$

6. To estimate the coefficients, we will use a two step FGLS procedure. Ordinary least squares estimates based on Section 19.4.3 are consistent, but inefficient. The OLS regression produces

$$\Phi^{-1}(P_i) = \hat{z}_i = -2.18098 + .0098898T$$
$$(.7404) \quad (.002883).$$

The predicted values from this regression can then be used to compute the weights in (21-39). The weighted least squares regression produces $\hat{z}_i = -2.3116 + .010646T$

$$(.8103) \quad (.003322)$$

In order to achieve a predicted proportion of 95%, we will require $z_i = 1.645$. The T required to achieve this is

$$T = (1.645 + 2.3116) / .010646 = 372.$$

The z_i which corresponds to 90% is 1.282. Doing the same calculation as above, this requires T = 338 trucks. At \$20,000 per truck, this requires \$6.751 million, so the budget is inadequate.

The marginal effect is $\partial\Phi_i/\partial T = .010646\phi(z_i)$. At $T = 300$, $z_i = .8822$, so $\phi(z_i) = .2703$ and the marginal effect is .00288.

7. This is similar to Exercise 1. It is simplest to prove it in that framework. Since the model has only a dummy variable, we can use the same log likelihood as in Exercise 1. But, in this exercise, there are no observations in the cell (y=1,x=0). The resulting log likelihood is, therefore,

$$\ln L = \Sigma_{0,0}\ln\text{Prob}[y=0,x=0] + \Sigma_{0,1}\ln\text{Prob}[y=0,x=1] + \Sigma_{1,1}\ln\text{Prob}[y=1,x=1]$$

or

$$\ln L = n_3\ln\text{Prob}[y=0,x=0] + n_2\ln\text{Prob}[y=0,x=1] + n_1\ln\text{Prob}[y=1,x=1].$$

Now, let $\delta = \alpha + \beta$. The log likelihood function is $\ln L = n_3\ln(1 - F(\alpha)) + n_2\ln(1 - F(\delta)) + n_1\ln F(\delta)$. For estimation, let $A = F(\alpha)$ and $D = F(\delta)$. We can estimate A and D, then $\alpha = F^{-1}(A)$ and $\beta = F^{-1}(D) - \alpha$. The first order condition for estimation of A is $\partial\ln L/\partial A = -n_3/(1 - A) = 0$, which obviously has no solution. If A cannot be estimated then α cannot either, nor, in turn, can β. This applies to both probit and logit models.

8. We'll do this more generally for any model $F(\alpha)$. Since the 'model' contains only a constant, the log likelihood is $\log L = \Sigma_0\log[1-F(\alpha)] + \Sigma_1\log F(\alpha) = n_0\log[1-F(\alpha)]+n_1\log F(\alpha)$. The likelihood equation is $\partial\log L/\partial\alpha = \Sigma_0[-f(\alpha)/[1-F(\alpha)] + \Sigma_1 f(\alpha)/F(\alpha) = 0$ where $f(\alpha)$ is the density (derivative of $F(\alpha)$ so that at the solution, $n_0f(\alpha)/[1-F(\alpha)] = n_1f(\alpha)/F(\alpha)$. Divide both sides of this equation by $f(\alpha)$ and solve it for $F(\alpha) = n_1/(n_0+n_1)$, as might be expected. You can then insert this solution for $F(\alpha)$ back into the log likelihood, and (23-28) follows immediately.

9. Look at the two cases. Neither case has an estimator which is consistent in both cases. In both cases, the unconditional fixed effects effects estimator is inconsistent, so the rest of the analysis falls apart. This is the incidental parameters problem at work. Note that the fixed effects estimator is inconsistent because in both models, the estimator of the constant terms is a function of 1/T. Certainly in both cases, if the fixed effects model is appropriate, then the random effects estimator is inconsistent, whereas if the random effects model is appropriate, the maximum likelihood random effects estimator is both consistent and efficient. Thus, in this instance, the random effects satisfies the requirements of the test. In fact, there does exist a consistent estimator for the logit model with fixed effects - see the text. However, this estimator must be based on a restricted sample observations with the sum of the ys equal to zero or T muust be discarded, so the mechanics of the Hausman test are problematic. This does not fall into the template of computations for the Hausman test.

Applications

1. Binary Choice for Extramarital Affairs using Redbook data

```
?=======================================================
? Application 23.1
?=======================================================
?
Create ; A = (Yrb > 0) $
Namelist ; X = one,v1,v2,v5,v6 $
Probit ; Lhs = A ; Rhs = X ; marginal Effects $
Logit  ; Lhs = A ; Rhs = X ; marginal Effects $
+---------------------------------------------+
| Binomial Probit Model                       |
| Maximum Likelihood Estimates                |
| Dependent variable                    A     |
| Number of observations             6366     |
| Log likelihood function       -3547.865     |
| Number of parameters                  5     |
| Info. Criterion: AIC =          1.11620     |
| Info. Criterion: BIC =          1.12151     |
| Restricted log likelihood     -4002.530     |
+---------------------------------------------+
+--------+-------------+---------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+---------------+--------+--------+----------+
---------+Index function for probability
```

```
 Constant|    1.43453507       .15493583     9.259   .0000
 V1      |    -.42595261       .01807583   -23.565   .0000   4.10964499
 V2      |     .02797013       .00254409    10.994   .0000   29.0828621
 V5      |    -.20942202       .02015534   -10.390   .0000   2.42617028
 V6      |    -.03522668       .00801808    -4.393   .0000   14.2098649
+--------------------------------------------+
| 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]|Elasticity|
+--------+--------------+----------------+--------+--------+----------+
---------+
 Constant|     .27876593       .01081795    25.769   .0000
 V1      |    -.14911732       .00634679   -23.495   .0000  -2.01181601
 V2      |     .00979177       .00088860    11.019   .0000    .93487672
 V5      |    -.07331438       .00703451   -10.422   .0000   -.58393740
 V6      |    -.01233214       .00280535    -4.396   .0000   -.57528664
+----------------------------------------------+
| Binary Logit Model for Binary Choice         |
| Maximum Likelihood Estimates                 |
| Dependent variable                    A      |
| Number of observations             6366      |
| Log likelihood function       -3549.741      |
| Number of parameters                  5      |
| Info. Criterion: AIC =          1.11679      |
| Info. Criterion: BIC =          1.12210      |
| Restricted log likelihood     -4002.530      |
+----------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
---------+Characteristics in numerator of Prob[Y = 1]
 Constant|    2.41622262       .26160831     9.236   .0000
 V1      |    -.70802698       .03091557   -22.902   .0000   4.10964499
 V2      |     .04624150       .00426119    10.852   .0000   29.0828621
 V5      |    -.35139771       .03413337   -10.295   .0000   2.42617028
 V6      |    -.05899324       .01354756    -4.355   .0000   14.2098649
+--------------------------------------------+
| Partial derivatives of probabilities with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used are All Obs.            |
+--------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|Elasticity|
+--------+--------------+----------------+--------+--------+----------+
---------+Marginal effect for variable in probability
 Constant|     .50898166       .05554126     9.164   .0000
 V1      |    -.14914716       .00650799   -22.918   .0000  -2.03205673
 V2      |     .00974086       .00089378    10.898   .0000    .93918419
 V5      |    -.07402256       .00714156   -10.365   .0000   -.59539053
 V6      |    -.01242703       .00285019    -4.360   .0000   -.58542862
```

2. Ordered Choice For Self Reported Marriage Rating

```
+----------------------------------------------+
| Ordered Probability Model                    |
| Maximum Likelihood Estimates                 |
| Dependent variable             MARRIAGE      |
| Weighting variable                 None      |
```

```
| Number of observations             6366      |
| Iterations completed                 15      |
| Log likelihood function       -7720.145      |
| Number of parameters                 12      |
| Info. Criterion: AIC =          2.42920      |
| Info. Criterion: BIC =          2.44194      |
| Restricted log likelihood     -7926.487      |
| Underlying probabilities based on Normal     |
+----------------------------------------------+
+----------------------------------------------+
| Ordered Probability Model                    |
| Cell frequencies for outcomes                |
|  Y Count Freq  Y Count Freq  Y Count Freq    |
|  0    99 .015  1   348 .054  2   993 .155    |
|  3  2242 .352  4  2684 .421                  |
+----------------------------------------------+
+--------+-------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+----------------+--------+--------+----------+
---------+Index function for probability
 Constant|    1.87997564       .12760529    14.733   .0000
 YRB     |    -.09669427       .00649907   -14.878   .0000    .70537389
 V2      |    -.00624520       .00471646    -1.324   .1855   29.0828621
 V3      |    -.00952932       .00506534    -1.881   .0599   9.00942507
 V4      |    -.05879586       .01520251    -3.868   .0001   1.39687402
 V5      |     .10524384       .01624338     6.479   .0000   2.42617028
 V6      |     .02526318       .00727002     3.475   .0005   14.2098649
 V7      |     .02069865       .01614318     1.282   .1998   3.42412818
 V8      |     .02725715       .01072244     2.542   .0110   3.85014138
---------+Threshold parameters for index
 Mu(1)   |     .71088354       .02219910    32.023   .0000
 Mu(2)   |    1.47186849       .01737814    84.697   .0000
 Mu(3)   |    2.46392113       .01923976   128.064   .0000

+----------------------------------------------------------------------+
| Summary of Marginal Effects for Ordered Probability Model (probit)     |
+----------------------------------------------------------------------+
Variable|    Y=00    Y=01    Y=02    Y=03    Y=04    Y=05    Y=06    Y=07 |
-----------------------------------------------------------------------+
YRB          .0031   .0087   .0167   .0093  -.0377
V2           .0002   .0006   .0011   .0006  -.0024
V3           .0003   .0009   .0016   .0009  -.0037
V4           .0019   .0053   .0101   .0056  -.0229
V5          -.0033  -.0095  -.0182  -.0101   .0411
V6          -.0008  -.0023  -.0044  -.0024   .0099
V7          -.0007  -.0019  -.0036  -.0020   .0081
V8          -.0009  -.0025  -.0047  -.0026   .0106
+------------------------------------------------------------------------+
|   Cross tabulation of predictions. Row is actual, column is predicted.  |
|   Model = Probit    .  Prediction is number of the most probable cell.  |
+-------+-------+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Actual|Row Sum|  0  |  1  |  2  |  3  |  4  |  5  |  6  |  7  |  8  |  9  |
+-------+-------+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|      0|     99|    0|    0|    0|   68|   31|
|      1|    348|    2|    0|    5|  170|  171|
|      2|    993|    7|    0|    7|  453|  526|
|      3|   2242|    3|    0|   10|  674| 1555|
|      4|   2684|    2|    0|    5|  593| 2084|
+-------+-------+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|Col Sum|   6366|   14|    0|   27| 1958| 4367|    0|    0|    0|    0|    0|
+-------+-------+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
```

Chapter 24
Truncation, Censoring and Sample Selection

Exercises

1. The sample mean of all 20 observations is 4.18222. For the 14 nonzero observations, the mean is (20/14)4.18222 = 5.9746. Both of these should overestimate μ. In the first case, all negative values have been transformed to zeroes. Therefore, if we had had the original data, our estimator would include the negative values as well as the positive ones. Since we have only the zeroes, instead, our estimator includes, for every negative y^* a number which is larger than the true y^*. This will inflate the estimate. Likewise, for the truncated mean, whereas a complete sample might include some negative values, the observed one will not. Once again, this will serve to inflate the estimator of the mean.

2. The log-likelihood for the Tobit model is given in (24-13). With only a constant term, this is

$$\ln L = (-n_1/2)[\ln(2\pi) + \ln\sigma^2] - (1/(2\sigma^2))\Sigma_1(y_i - \mu)^2 + \Sigma_0\ln\Phi(-\mu/\sigma)$$

In terms of γ and θ, this is

$$\ln L = (-n_1/2)[\ln(2\pi) - \ln\theta^2] - (1/2)\Sigma_1(\theta y_i - \gamma)^2 + \Sigma_0\ln\Phi(-\gamma)$$

$$= (-n_1/2)\ln(2\pi) + n_1\ln\theta - (1/2)\Sigma_1(\theta y_i - \gamma)^2 + \Sigma_0\ln\Phi(-\gamma).$$

The necessary conditions for maximizing this with respect to γ and θ are

$$\partial\ln L/\partial\gamma = \Sigma_1(\theta y_i - \gamma) - \Sigma_0\phi(-\gamma)/\Phi(-\gamma) = \theta\Sigma_1 y_i - n_1\gamma - n_0[\phi(-\gamma)/\Phi(\gamma)] = 0$$

$$\partial\ln L/\partial\theta = n_1/\theta - \Sigma_1 y_i(\theta y_i - \gamma) = n_1/\theta - \theta\Sigma_1 y_i^2 + \gamma\Sigma_1 y_i = 0.$$

There are a few different ways one might solve these two equations. A grid search over the values of γ and θ is a possibility. A direct maximum likelihood estimator for the tobit model is the simpler choice if one is available. The model with only a constant term is otherwise the same as the usual model. Using the data above, the tobit maximum likelihood estimates are $\hat{\mu} = 3.2731$, $\hat{\sigma} = 5.0303$.

3. The log-likelihood for the truncated regression with only a constant term is

$$\ln L = (-n/2)[\ln(2\pi) + \ln\sigma^2] - (1/(2\sigma^2))\Sigma_1(y_i - \mu)^2 - \Sigma_i\ln\Phi(\mu/\sigma)$$

Once again transforming to γ and σ, this is

$$\ln L = -(n/2)\ln(2\pi) + n\ln\theta - (1/2)\Sigma_i(\theta y_i - \gamma)^2 - n\ln\Phi(\gamma).$$

The necessary conditions for maximizing this are

$$\partial\ln L/\partial\gamma = \Sigma_i(\theta y_i - \gamma) - n\phi(\gamma)/\Phi(\gamma) = 0$$

$$\partial\ln L/\partial\theta = n/\theta - \Sigma_i y_i(\theta y_i - \gamma)$$

The first of the two equations can be $\overline{y} = \gamma/\theta + \lambda/\theta$, where $\lambda = \phi(\gamma)/\Phi(\gamma)$. Now, reverting back to μ and σ, this is $\overline{y} = \mu + \sigma\lambda$ which is (24-6). The second equation can be manipulated to produce $\Sigma y_i^2/n - \mu\,\overline{y} = \sigma^2$. Once again, trial and error could be used to find a solution. As before, estimating the model as a truncated regression with only a constant term will also produce a solution. The solution by this method is $\hat{\mu} = 3.3439$, $\hat{\sigma} = 5.6368$. With the data of the first problem, we would have the following: Estimated Prob[$y^* > 0$] = 14/20 = .7. This is an estimate of $\Phi(\mu/\sigma)$, so we would have $\mu/\sigma = \Phi^{-1}(.7) = .525$ or $\mu = .525\sigma$. Now, we can use the relationship $E[y|y > 0] = \mu + \sigma\phi(\mu/\sigma)/\Phi(\mu/\sigma) = \mu + \sigma\lambda$. Since μ/σ is now known, we have $\lambda = \phi(.525) / \Phi(.525) = .496$ so a second equation is $5.9746 = \mu + .496\sigma$. The joint solution is $\hat{\mu} = 3.0697$, $\hat{\sigma} = 5.8470$. The three solutions are surprisingly close.

4. Using Theorem 24.5, we have $1 - \Phi(\alpha_z) = 14/35 = .4$, $\alpha_z = \Phi^{-1}(.6) = .253$, $\lambda(\alpha_z) = .9659$, $\delta(\alpha_z) = .6886$. The two moment equations are based on the mean and variance of y in the observed data, 5.9746 and 9.869, respectively. The equations would be $5.9746 = \mu + \sigma(.7)(.9659)$ and $9.869 = \sigma^2(1 - .7^2(.6886))$. The joint solution is $\hat{\mu} = 3.3651$, $\hat{\sigma} = 3.8594$.

5. The conditional mean function is $E[y|\mathbf{x}] = \Phi(\boldsymbol{\beta}'\mathbf{x}_i/\sigma_i)\boldsymbol{\beta}'\mathbf{x}_i + \sigma_i\Phi(\boldsymbol{\beta}'\mathbf{x}_i/\sigma_i)$ using the equation before (24-12). Suppose that $\sigma_i = \sigma\exp(\boldsymbol{\alpha}'\mathbf{x}_i)$ for the same vector $\mathbf{x}_i$. (We'll relax that assumption shortly.) Now, differentiate this expression with respect to $\mathbf{x}$. We differentiate the two parts, first with respect to $\boldsymbol{\beta}'\mathbf{x}$ then with respect to σ_i.

$$\frac{\partial E[y_i|\mathbf{x}_i]}{\partial \mathbf{x}_i} = \Phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\boldsymbol{\beta} + \left(\boldsymbol{\beta}'\mathbf{x}_i\right)\phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\frac{1}{\sigma_i}\boldsymbol{\beta} + \sigma_i\left[-\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\right]\frac{1}{\sigma_i}\boldsymbol{\beta}$$

$$+\left(\boldsymbol{\beta}'\mathbf{x}_i\right)\phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\left(\frac{-1}{\sigma_i}\right)\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\sigma_i\boldsymbol{\alpha} + \phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\sigma_i\boldsymbol{\alpha} + \sigma_i\left[-\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\right]\left(\frac{-1}{\sigma_i}\right)\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma_i}\right)\sigma_i\boldsymbol{\alpha}$$

After collecting the terms, we obtain $\partial E[y_i|\mathbf{x}_i]/\partial\mathbf{x}_i = \Phi(a_i)\boldsymbol{\beta} + \sigma_i\phi(a_i)\boldsymbol{\alpha}$ where $a_i = \boldsymbol{\beta}'\mathbf{x}_i/\sigma_i$. Thus, the marginal effect has two parts. one for $\boldsymbol{\beta}$ and one for α. Now, if a variable appears in σ_i but not in $\mathbf{x}_i$, then only the second term appears while if a variable appears only in $\mathbf{x}_i$ and not in σ_i, then only the first term appears in the marginal effect.

6. The transformed log likelihood function is

$$\log L = \Sigma_{y>0}(-1/2)[\log 2\pi - \log\theta^2 + (\theta y - \mathbf{x}'\boldsymbol{\gamma})^2] + \Sigma_{y=0}\log[1-\Phi(\mathbf{x}'\boldsymbol{\gamma})]$$

It will be convenient to define $a_i = \mathbf{x}_i'\boldsymbol{\gamma}$. Note also that $1 - \Phi(a_i) = \Phi(-a_i)$. The first derivatives and Hessian in the transformed parameters are

$$\frac{\partial \log L}{\partial\theta} = \sum_{y_i>0} (1/\theta) - y_i\left(\theta y_i - a_i\right)$$

$$\frac{\partial \log L}{\partial\boldsymbol{\gamma}} = \sum_{y_i>0} \mathbf{x}_i\left(\theta y_i - a_i\right) + \sum_{y_i=0}\left[\phi(-a_i)/\Phi(-a_i)\right](-\mathbf{x}_i)$$

$$\frac{\partial^2 \log L}{\partial\theta^2} = \sum_{y_i>0} -1/\theta^2 - y_i^2$$

$$\frac{\partial^2 \log L}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'} = \sum_{y_i>0} -\mathbf{x}_i\mathbf{x}_i' + \sum_{y_i=0} -\left[\phi(-a_i)/\Phi(-a_i)\right]\{-a_i + \left[\phi(-a_i)/\Phi(-a_i)\right]\}\mathbf{x}_i\mathbf{x}_i'$$

$$\frac{\partial^2 \log L}{\partial\boldsymbol{\gamma}\partial\theta} = \sum_{y_i>0} -\mathbf{x}_i y_i$$

The second derivatives can be collected in a matrix format:

$$\frac{\partial \log L}{\partial\begin{pmatrix}\boldsymbol{\gamma}\\ \theta\end{pmatrix}\partial\begin{pmatrix}\boldsymbol{\gamma}\\ \theta\end{pmatrix}'} = \sum_{y>0}\left[-\begin{pmatrix}\mathbf{x}_i\\ -y_i\end{pmatrix}\begin{pmatrix}\mathbf{x}_i\\ -y_i\end{pmatrix}' - \begin{pmatrix}0\\ \theta\end{pmatrix}\begin{pmatrix}0\\ \theta\end{pmatrix}'\right] + \sum_{y=0}\delta_i\begin{pmatrix}\mathbf{x}_i\\ 0\end{pmatrix}\begin{pmatrix}\mathbf{x}_i\\ 0\end{pmatrix}'$$

where δ_i is the last scalar term in $\partial^2\log L/\partial\delta\partial\boldsymbol{\gamma}'$. By Theorem 22.2 (see (24-4)), we know that δ_i is negative. Thus, all three parts of the matrix are negative semidefinite. Assuming the data are not linearly dependent and there are more than K observations, the Hessian will have full rank and be negative definite.

Applications

1. Tobit model for Redbook data

```
?=============================================================
? Applications in Chapter 24
?=============================================================
? 1.  Tobit, Scaled Tobit, Probit and Truncated Regression.
?     In principle, all are estimating the same paramter.
? For consistency and convenience, we are going to use the
? sample with YRB <= 5 only.
?=============================================================
Sample ; All $
Reject ; YRB > 5 $
Namelist ; X = one,v1,v2,v3,v4,v5$
Tobit ; Lhs = yrb ; Rhs = x ; marginal $
Matrix ; list ;  scaled_b = 1/s * b $
Probit ; Lhs = a ; Rhs = x $
reject ; yrb <= 0 $
Truncation ; Lhs = yrb ; Rhs = x $

+----------------------------------------------+
| Limited Dependent Variable Model - CENSORED |
| Maximum Likelihood Estimates                 |
| Dependent variable                 YRB      |
| Weighting variable                None      |
| Number of observations            6217      |
| Iterations completed                 6      |
| Log likelihood function      -6118.089      |
| Number of parameters                 7      |
| Info. Criterion: AIC =         1.97043      |
|   Finite Sample: AIC =         1.97044      |
| Info. Criterion: BIC =         1.97802      |
| Info. Criterion:HQIC =         1.97306      |
| Threshold values for the model:             |
| Lower=     .0000     Upper=+infinity        |
| LM test [df] for tobit=    622.887[  6]     |
| Normality Test, LM    =    150.850[  2]     |
| ANOVA  based fit measure =    .293201       |
| DECOMP based fit measure =    .438743       |
+----------------------------------------------+
+--------+-------------+---------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+---------------+--------+--------+----------+
---------+Primary Index Equation for Model
 Constant|    4.13828429       .31908252    12.969   .0000
 V1      |    -.80415431       .03782416   -21.260   .0000   4.12272800
 V2      |    -.06923599       .01229186    -5.633   .0000   29.1829661
 V3      |     .10402446       .01325380     7.849   .0000   9.12329098
 V4      |    -.02190617       .03898707     -.562   .5742   1.41499115
 V5      |    -.43110692       .04356398    -9.896   .0000   2.43670581
---------+Disturbance standard deviation
 Sigma   |    2.27697641       .04212836    54.049   .0000
+-----------------------------------------+
| Partial derivatives of expected val. with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs.  |
| Conditional Mean at Sample Point    .3941 |
| Scale Factor for Marginal Effects   .2796 |
+-----------------------------------------+
+--------+-------------+---------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
```

```
+--------+-------------+---------------+--------+--------+----------+
 Constant|    1.15697490       .09110678    12.699   .0000
 V1      |    -.22482418       .01048093   -21.451   .0000   4.12272800
 V2      |    -.01935689       .00342807    -5.647   .0000   29.1829661
 V3      |     .02908299       .00367661     7.910   .0000   9.12329098
 V4      |    -.00612449       .01090115     -.562   .5742   1.41499115
 V5      |    -.12052818       .01207702    -9.980   .0000   2.43670581
 Sigma   |       .000000    ......(Fixed Parameter).......
Matrix SCALED_B has  6 rows and  1 columns.
               1
        +--------------
       1|    1.81745
       2|    -.35317
       3|    -.03041
       4|     .04569
       5|    -.00962
       6|    -.18933
+---------------------------------------------+
| Binomial Probit Model                       |
| Maximum Likelihood Estimates                |
| Dependent variable                     A    |
| Weighting variable                  None    |
| Number of observations              6217    |
| Iterations completed                   5    |
| Log likelihood function       -3310.310    |
| Number of parameters                   6    |
| Info. Criterion: AIC =          1.06685    |
| Info. Criterion: BIC =          1.07335    |
| Restricted log likelihood     -3830.126    |
+---------------------------------------------+
+--------+-------------+---------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+---------------+--------+--------+----------+
---------+Index function for probability
 Constant|    2.03641060       .15678428    12.989   .0000
 V1      |    -.41449474       .01860450   -22.279   .0000   4.12272800
 V2      |    -.03568737       .00593540    -6.013   .0000   29.1829661
 V3      |     .07215336       .00640693    11.262   .0000   9.12329098
 V4      |    -.00241124       .01891503     -.127   .8986   1.41499115
 V5      |    -.21212886       .02089864   -10.150   .0000   2.43670581
+---------------------------------------------+
| Limited Dependent Variable Model - TRUNCATE |
| Maximum Likelihood Estimates                |
| Dependent variable                   YRB    |
| Weighting variable                  None    |
| Number of observations              1904    |
| Iterations completed                   8    |
| Log likelihood function       -2437.473    |
| Number of parameters                   7    |
| Info. Criterion: AIC =          2.56772    |
| Info. Criterion: BIC =          2.58813    |
| Threshold values for the model:             |
| Lower=     .0000     Upper=+infinity         |
| Observations after truncation      1904    |
+---------------------------------------------+
+--------+-------------+---------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+---------------+--------+--------+----------+
---------+Primary Index Equation for Model
 Constant|    5.22651388       .94010948     5.559   .0000
 V1      |    -.45753380       .10715203    -4.270   .0000   3.65388655
 V2      |    -.04779763       .03766086    -1.269   .2044   30.9776786
 V3      |    -.25376184       .04622853    -5.489   .0000   11.6919643
```

```
 V4      |    -.37961397       .12878071    -2.948   .0032   1.81407563
 V5      |    -.22780476       .13328147    -1.709   .0874   2.28308824
---------+Disturbance standard deviation
 Sigma   |    2.38479704       .13327563    17.894   .0000
```

2. Two part Model.

The three estimated models appear above. The test statistic is

```
+-----------------------------------+
| Listed Calculator Results         |
+-----------------------------------+
 TEST2   =    740.610758
```

This is much larger than the chi squared critical value for 5 degrees of freedom. We conclude that the participation equation (probit) is different from the intensity equation (yrb).

Chapter 25

Models for Event Counts and Duration

Exercises

1. a. Conditional variance in the ZIP model. The essential ingredients that are needed for this derivation are

$$E[y^* \mid y^* > 0, \mathbf{x}_i] = \frac{\lambda_i}{1 - \exp(-\lambda_i)} \;=\; E_i^*$$

and

$$Var[y^* \mid y^* > 0, \mathbf{x}_i] = \left(\frac{\lambda_i}{1 - \exp(-\lambda_i)}\right)\left(1 - \frac{\lambda_i}{\exp(\lambda_i) - 1}\right) = E_i^* \left(1 - \frac{\lambda_i}{\exp(\lambda_i) - 1}\right) = E_i^* V_i^*$$

[See, e.g., Winkelmann (2003, pp. 33-34).]. We found the conditional mean in the text to be

$$E[y_i|x_i,w_i] \;=\; \frac{F_i\lambda_i}{1 - \exp(-\lambda_i)} \;=\; F_i\, E_i^*$$

To obtain the variance, we will use the variance decomposition,

$$Var[y_i|x_i,w_i] = E_z[Var[y_i|x_i,z]] + Var_z[E[y_i|x_i,z]].$$

The expectation of the conditional variance is

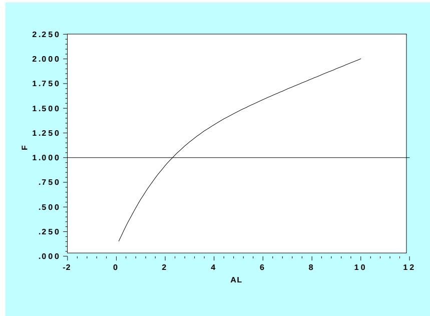
$$E_z[Var[y_i|x_i,z]] \;=\; (1 - F_i)\times 0 + F_i\times\left(\frac{\lambda_i}{1 - \exp(-\lambda_i)}\right)\left(1 - \frac{\lambda_i}{\exp(\lambda_i) - 1}\right) = F_i \times E_i^* \times V_i^*$$

The variance of the conditional mean is

$$(1 - F_i) \times \left(0 - \frac{F_i\lambda_i}{1 - \exp(-\lambda_i)}\right)^2 + F_i\left(\frac{\lambda_i}{1 - \exp(-\lambda_i)} - \frac{F_i\lambda_i}{1 - \exp(-\lambda_i)}\right)^2 \;=\; F_i(1-F_i)\left(\frac{\lambda_i}{1 - \exp(-\lambda_i)}\right)^2$$

$$=\; F_i(1 - F_i)E_i^{*2}.$$

The unconditional variance is thus, $F_i\, E_i^*\, [V_i^* + (1 - F_i)E_i^*]$. To obtain τ_i we divide by the conditional mean, which is $F_i\, E_i^*$, so $\tau_i = [V_i^* + (1 - F_i)E_i^*]$. Is this greater than E_i^*? Not necessarily. The figure below plots $F_i(1 - F_i)E_i^{*2}$ for $F_i = .9$ and various values of λ from .1 to about 12. There is a large range over which the function is less than one.



b. Partial Effects. The mean is F_i E_i*. We suppose that w_i and x_i are the same for the moment.

$$\partial E_i/\partial x_i = E_i^*\partial F_i/\partial x_i + F_i\ \partial E_i^*/\partial x_i.$$

The first term is $E_i^*\times f_i\times\gamma$. The second term is $F_i\ \partial E_i^*/\partial\lambda_i\ \lambda_i\beta$. The missing element is

$$\partial E_i^*/\partial\lambda_i = \lambda_i/[1-\exp(-\lambda_i)] \times [1 - \exp(-\lambda_i)/[1-\exp(-\lambda_i)].$$

Comnbining terms produces the marginal effects.

2. Let y* denote the unobserved random variable that is distributed as Poisson with probability

$$\text{Prob}(y^* = j|x) = P(j) = \exp(-\lambda)\lambda^j/j!.$$

The observed random variable before the censoring is is y = y*|y*>0. The probabilities are

$$\text{Prob}(y = j|x) = P(j)/[1 - P(0)].$$

Let yc = the censored random variable. Then, yc = y for y = 1,2,3,4. yc = 5 when $y \geq 5$. The probabilities associated with the observed yc are

$$\text{Prob}(yc = 1|x) = \text{Prob}(y = 1|x) = P(1)/[1-P(0)]$$
$$\text{Prob}(yc = 2|x) = \text{Prob}(y = 2|x) = P(2)/[1-P(0)]$$
$$\text{Prob}(yc = 3|x) = \text{Prob}(y = 3|x) = P(3)/[1-P(0)]$$
$$\text{Prob}(yc = 4|x) = \text{Prob}(y = 4|x) = P(4)/[1-P(0)]$$
$$\text{Prob}(yc = 5|x) = \text{Prob}(y = 5|x) + \text{Prob}(y = 6|x) + \text{Prob}(y = 7|x) + ...$$

The last term is an infinite sum. But,

$$\text{Prob}(y = 5|x) + \text{Prob}(y = 6|x) + \text{Prob}(y = 7|x) + ...$$
$$= 1 - \text{Prob}(y = 1|x) - \text{Prob}(y = 2|x) - \text{Prob}(y = 3|x) - \text{Prob}(y = 4|x)$$

Therefore,

$$\text{Prob}(yc = 5|x) = [1 - P(1) - P(2) - P(3) - P(4)]/[1 - P(0)].$$

These are the probabilities used to construct the log likelihood function for the observed values of yc, 1,2,3,4,5.

3. The hazard function is easily obtained as h(t) = -dlnS(t)/dt. For the Weibull model, $\ln S(t) = -(\lambda t)^P$ to the hazard function is $(\lambda p)(\lambda t)^{P-1}$. The median survival time occurs where the survival function equals .5. Thus,

$$\exp(-(\lambda t)^P) = .5$$
$$-(\lambda t)^P = \ln .5$$
$$(\lambda t)^P = -\ln .5 = \ln 2$$
$$P^*\ln(\lambda) + P\ \ln t = \ln\ln 2$$
$$P\ \ln t = \ln\ln 2 - P\ \ln\lambda$$
$$\ln t = (1/P)[\ln\ln 2 - P\ \ln\lambda]$$
$$t = \exp[(1/P)[\ln\ln 2 - P\ \ln\lambda].$$

Applications

1.

```
?===================================================
? Application 25.1
?===================================================
Namelist ;x = age,educ,hhninc,hsat $
Poisson ; Lhs = HospVis ; Rhs = One,X
        ; Marginal effects $
Calc    ; Lp = logl $
Regress ; Lhs = HospVis ; Rhs = One,X $
Negbin  ; Lhs = HospVis ; Rhs = One,X
        ; Marginal effects $
Calc    ; Ln = logl $
Calc    ; List ; LRstat = 2*(ln - lp) $

?===================================================
? Application 25.2
?===================================================
Sample  ; All $
Regress ; Lhs = one ; Rhs = one ; Str = ID ; Panel $
Poisson ; Lhs = HospVis ; Rhs = One,X
        ; Marginal effects
        ; Pds = _Groupti $
Poisson ; Lhs = HospVis ; Rhs = One,X
        ; Marginal effects
        ; Pds = _Groupti ; Random $

+---------------------------------------------+
| Poisson Regression                          |
| Maximum Likelihood Estimates                |
| Dependent variable              HOSPVIS     |
| Weighting variable                 None     |
| Number of observations            27326     |
| Iterations completed                  9     |
| Log likelihood function       -12636.40     |
| Number of parameters                  5     |
| Info. Criterion: AIC =           .92523     |
| Info. Criterion: BIC =           .92673     |
| Restricted log likelihood     -13433.21     |
+---------------------------------------------+
+---------------------------------------------+
| Poisson Regression                          |
| Chi- squared =124476.35621  RsqP=   .1947   |
| G  - squared = 20025.66932  RsqD=   .0737   |
| Overdispersion tests: g=mu(i)  :  5.279     |
| Overdispersion tests: g=mu(i)^2:  5.468     |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|     .12613692       .12567036     1.004   .3155
 AGE     |    -.00340754       .00149685    -2.276   .0228   43.5256898
 EDUC    |    -.05295428       .00834958    -6.342   .0000   11.3206310
 HHNINC  |     .39889043       .08982355     4.441   .0000    .35208362
 HSAT    |    -.24901310       .00634000   -39.277   .0000   6.78542607
+-------------------------------------------+
| Partial derivatives of expected val. with |
| respect to the vector of characteristics. |
| Effects are averaged over individuals.    |
| Observations used for means are All Obs.  |
```

```
| Conditional Mean at Sample Point    .1383 |
| Scale Factor for Marginal Effects   .1383 |
+-------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|     .01743926       .02183573      .799   .4245
 AGE     |    -.00047111       .00025979    -1.813   .0698   43.5256898
 EDUC    |    -.00732128       .00149415    -4.900   .0000   11.3206310
 HHNINC  |     .05514924       .01579375     3.492   .0005    .35208362
 HSAT    |    -.03442771       .00220148   -15.638   .0000   6.78542607
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=HOSPVIS  Mean                 =   .1382566     |
|              Standard deviation   =   .8843390     |
| WTS=none     Number of observs.   =      27326     |
| Model size   Parameters           =          5     |
|              Degrees of freedom   =      27321     |
| Residuals    Sum of squares       =   21121.96     |
|              Standard error of e  =   .8792630     |
| Fit          R-squared            =   .1159150E-01 |
|              Adjusted R-squared   =   .1144679E-01 |
| Model test   F[  4, 27321] (prob) =  80.10 (.0000) |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|     .49839670       .04097910    12.162   .0000
 AGE     |    -.00064393       .00048945    -1.316   .1883   43.5256898
 EDUC    |    -.00619390       .00241633    -2.563   .0104   11.3206310
 HHNINC  |     .04936160       .03122845     1.581   .1140    .35208362
 HSAT    |    -.04117251       .00240443   -17.124   .0000   6.78542607
+---------------------------------------------+
| Negative Binomial Regression                |
| Dependent variable               HOSPVIS    |
| Number of observations             27326    |
| Iterations completed                   9    |
| Log likelihood function        -10044.46    |
| Number of parameters                   6    |
| Info. Criterion: AIC =            .73560    |
| Info. Criterion: BIC =            .73740    |
| Restricted log likelihood      -12636.40    |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|     .10394982       .12631220      .823   .4105
 AGE     |    -.00369348       .00143149    -2.580   .0099   43.5256898
 EDUC    |    -.05795593       .00826247    -7.014   .0000   11.3206310
 HHNINC  |     .38542430       .09259876     4.162   .0000    .35208362
 HSAT    |    -.23323713       .00651715   -35.788   .0000   6.78542607
---------+Dispersion parameter for count data model
 Alpha   |    6.70461029       .17537071    38.231   .0000
+-------------------------------------------+
| Partial derivatives of expected val. with |
| respect to the vector of characteristics. |
| Effects are averaged over individuals.    |
| Observations used for means are All Obs.  |
| Conditional Mean at Sample Point    .1367 |
| Scale Factor for Marginal Effects   .1367 |
+-------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
```

```
+--------+--------------+----------------+--------+--------+----------+
 Constant|     .01421398       .02120646      .670   .5027
 AGE     |    -.00050504       .00024071    -2.098   .0359   43.5256898
 EDUC    |    -.00792483       .00146645    -5.404   .0000   11.3206310
 HHNINC  |     .05270247       .01588312     3.318   .0009    .35208362
 HSAT    |    -.03189257       .00226820   -14.061   .0000   6.78542607
+-----------------------------------+
| Listed Calculator Results         |
+-----------------------------------+
 LRSTAT  =   5183.862874
```

2.

```
+--------------------------------------------+
| Panel Model with Group Effects             |
| Dependent variable              HOSPVIS    |
| Weighting variable                 None    |
| Number of observations            27326    |
| Log likelihood function       -4198.145    |
| Number of parameters                  4    |
| Info. Criterion: AIC =           .30756    |
| Info. Criterion: BIC =           .30876    |
| Unbalanced panel has    7293 individuals.  |
| Missing or sumY=0, Skipped  5640 groups.   |
| Poisson      Regression -- Fixed Effects   |
+--------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 AGE     |    -.00020613       .00705126     -.029   .9767   43.5256898
 EDUC    |    -.04033708       .09220144     -.437   .6618   11.3206310
 HHNINC  |     .49927712       .18484588     2.701   .0069    .35208362
 HSAT    |    -.16686419       .01027579   -16.239   .0000   6.78542607
+------------------------------------------+
| Partial derivatives of expected val. with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs.  |
| Conditional Mean at Sample Point    .1383 |
| Scale Factor for Marginal Effects   .1383 |
+------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 AGE     |   -.284995D-04      .00097488     -.029   .9767   1.00000000
 EDUC    |    -.00557687       .01274746     -.437   .6618   43.5256898
 HHNINC  |     .06902836       .02555616     2.701   .0069   11.3206310
 HSAT    |    -.02307008       .00142070   -16.239   .0000    .35208362
+--------------------------------------------+
| Panel Model with Group Effects             |
| Dependent variable              HOSPVIS    |
| Number of observations            27326    |
| Log likelihood function       -10200.91    |
| Number of parameters                  6    |
| Info. Criterion: AIC =           .74705    |
| Info. Criterion: BIC =           .74885    |
| Unbalanced panel has    7293 individuals.  |
| Poisson      Regression -- Random Effects  |
+--------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    -.22178663       .13617622    -1.629   .1034
 AGE     |    -.00170639       .00145901    -1.170   .2422   43.5256898
```

```
 EDUC    |    -.05399730       .01001912    -5.389   .0000   11.3206310
 HHNINC  |     .40499179       .06938275     5.837   .0000    .35208362
 HSAT    |    -.20075292       .00400154   -50.169   .0000   6.78542607
 Alpha   |    3.59227655       .11685254    30.742   .0000
+--------------------------------------------+
| Partial derivatives of expected val. with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs.  |
| Conditional Mean at Sample Point    .1383 |
| Scale Factor for Marginal Effects   .1383 |
+--------------------------------------------+
+--------+-------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+----------------+--------+--------+----------+
 Constant|    -.03066347       .01882726    -1.629   .1034
 AGE     |    -.00023592       .00020172    -1.170   .2422   43.5256898
 EDUC    |    -.00746548       .00138521    -5.389   .0000   11.3206310
 HHNINC  |     .05599279       .00959262     5.837   .0000    .35208362
 HSAT    |    -.02775542       .00055324   -50.169   .0000   6.78542607
```

3. Ship Accidents

```
Create ; logmth = log(months) $
Name ; X=logmth,one,ta,tb,tc,td,t6064,t6569,t7074,o6074$
Reject ; acc < 0 $
Pois ; lhs = acc ; Rhs = x $
Pois ; lhs = acc ; Rhs = x ; Rst = 1,9_b $
Negb ; lhs = acc ; Rhs = x ; Rst = 1,9_b,alpha $
+----------------------------------------------+
| Poisson Regression                           |
| Dependent variable                  ACC      |
| Number of observations               34      |
| Log likelihood function       -67.99930      |
| Number of parameters                 10      |
| Info. Criterion: AIC =          4.58819      |
| Info. Criterion: BIC =          5.03712      |
| Restricted log likelihood     -356.2029      |
+----------------------------------------------+
+----------------------------------------------+
| Poisson Regression                           |
| Chi- squared =    39.70580  RsqP=   .9491    |
| G  - squared =    38.13211  RsqD=   .9380    |
| Overdispersion tests: g=mu(i)  :   .853      |
| Overdispersion tests: g=mu(i)^2:  -.760      |
+----------------------------------------------+
+--------+-------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+----------------+--------+--------+----------+
 LOGMTH  |     .90617018       .10174566     8.906   .0000   7.04925451
 Constant|   -4.61752968       .72938865    -6.331   .0000
 TA      |    -.26966656       .24189066    -1.115   .2649    .20588235
 TB      |    -.62826604       .32582681    -1.928   .0538    .20588235
 TC      |   -1.03179604       .34039236    -3.031   .0024    .20588235
 TD      |    -.40106977       .30540945    -1.313   .1891    .20588235
 T6064   |    -.36146212       .24726698    -1.462   .1438    .23529412
 T6569   |     .30035782       .21325393     1.408   .1590    .29411765
 T7074   |     .39874282       .20053445     1.988   .0468    .29411765
 O6074   |    -.36986273       .11821010    -3.129   .0018    .41176471
+----------------------------------------------+
| Poisson Regression                           |
| Maximum Likelihood Estimates                 |
| Dependent variable                  ACC      |
```

```
| Number of observations               34     |
| Log likelihood function       -68.41456     |
| Number of parameters                  9     |
| Info. Criterion: AIC =          4.55380     |
| Info. Criterion: BIC =          4.95783     |
| Restricted log likelihood     -356.2029     |
+---------------------------------------------+
+---------------------------------------------+
| Poisson Regression                          |
| Chi- squared =    42.44145  RsqP=   .9456   |
| G  - squared =    38.96262  RsqD=   .9366   |
| Overdispersion tests: g=mu(i)  :   .934     |
| Overdispersion tests: g=mu(i)^2:  -.613     |
+---------------------------------------------+
+--------+-------------+---------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+---------------+--------+--------+----------+
 LOGMTH  |    1.00000000    ......(Fixed Parameter).......
 Constant|   -5.25351861       .24642858   -21.319   .0000
 TA      |    -.32052881       .23575203    -1.360   .1740    .20588235
 TB      |    -.86524026       .19852119    -4.358   .0000    .20588235
 TC      |   -1.00929327       .33950071    -2.973   .0030    .20588235
 TD      |    -.39483795       .30680184    -1.287   .1981    .20588235
 T6064   |    -.44497064       .23323916    -1.908   .0564    .23529412
 T6569   |     .25087485       .20875483     1.202   .2295    .29411765
 T7074   |     .37248476       .19930193     1.869   .0616    .29411765
 O6074   |    -.38385913       .11826046    -3.246   .0012    .41176471
```

There is no evidence of overdispersion. The tests from the Poisson model are both insignificant, and the estimate of α in the negative binomial model is essentially zero.

```
+---------------------------------------------+
| Negative Binomial Regression                |
| Dependent variable                  ACC     |
| Weighting variable                 None     |
| Number of observations               34     |
| Log likelihood function       -68.42007     |
| Number of parameters                 10     |
| Info. Criterion: AIC =          4.61295     |
|   Finite Sample: AIC =          4.89428     |
| Info. Criterion: BIC =          5.06188     |
| Info. Criterion:HQIC =          4.76604     |
| NegBin form 2; Psi(i) = theta               |
+---------------------------------------------+
+--------+-------------+---------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+-------------+---------------+--------+--------+----------+
 LOGMTH  |    1.00000000    ......(Fixed Parameter).......
 Constant|   -5.25074235       .26830333   -19.570   .0000
 TA      |    -.32296435       .39695609     -.814   .4159    .20588235
 TB      |    -.86731524       .20092395    -4.317   .0000    .20588235
 TC      |   -1.01171406       .24980570    -4.050   .0001    .20588235
 TD      |    -.39875463       .23889734    -1.669   .0951    .20588235
 T6064   |    -.44585250       .31679943    -1.407   .1593    .23529412
 T6569   |     .25060358       .27552926      .910   .3631    .29411765
 T7074   |     .37073607       .25504806     1.454   .1461    .29411765
 O6074   |    -.38364155       .15800844    -2.428   .0152    .41176471
---------+Dispersion parameter for count data model
 Alpha   |    .648724D-04      .02406424      .003   .9978
```

4. Strikes. There are 9 years of data. The number of strikes is 8,6,11,3,3,2,19,2,9. The Poisson regression is shown below. It does appear that the number of strikes is significantly related to the PROD variable. However, with only 9 observations, use of the asymptotic distribution for the test is probably overly optimistic. The result is probably borderline.

```
+---------------------------------------------+
| Poisson Regression                          |
| Dependent variable            _GROUPTI      |
| Weighting variable                None      |
| Number of observations               9      |
| Log likelihood function      -28.99317      |
| Number of parameters                 2      |
| Info. Criterion: AIC =         6.88737      |
| Info. Criterion: BIC =         6.93120      |
| Restricted log likelihood    -31.19884      |
+---------------------------------------------+
+---------------------------------------------+
| Poisson Regression                          |
| Chi- squared =    25.08061  RsqP=   .2317   |
| G  - squared =    26.13767  RsqD=   .1444   |
| Overdispersion tests: g=mu(i)  :  1.954     |
| Overdispersion tests: g=mu(i)^2:  2.618     |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
 Constant|    1.90854253       .12998621    14.683   .0000
 PROD    |    5.16576744      2.51306610     2.056   .0398   -.00302000
```