MODULE THREE, PART TWO: PANEL DATA ANALYSIS
IN ECONOMIC EDUCATION RESEARCH USING LIMDEP (NLOGIT)

Part Two of Module Three provides a cookbook-type demonstration of the steps required to use LIMDEP (NLOGIT) in panel data analysis. Users of this model need to have completed Module One, Parts One and Two, and Module Three, Part One. That is, from Module One users are assumed to know how to get data into LIMDEP, recode and create variables within LIMDEP, and run and interpret regression results. They are also expected to know how to test linear restrictions on sets of coefficients as done in Module One, Parts One and Two. Module Three, Parts Three and Four demonstrate in STATA and SAS what is done here in LIMDEP.

THE CASE

As described in Module Three, Part One, Becker, Greene and Siegfried (2009) examine the extent to which undergraduate degrees (BA and BS) in economics or Ph.D. degrees (PhD) in economics drive faculty size at those U.S. institutions that offer only a bachelor degree and those that offer both bachelor degrees and PhDs. Here we retrace their analysis for the institutions that offer only the bachelor degree. We provide and demonstrate the LIMDEP (NLOGIT) code necessary to duplicate their results.

DATA FILE

The following panel data are provided in the **comma separated values** (CSV) text file "bachelors.csv", which will automatically open in EXCEL by simply double clicking on it after it has been downloaded to your hard drive. Your EXCEL spreadsheet should look like this:

"College" identifies the bachelor degree-granting institution by a number 1 through 18.

"Year" runs from 1996 through 2006.

"Degrees" is the number of BS or BA degrees awarded in each year by each college.

"DegreBar" is the average number of degrees awarded by each college for the 16-year period.

"Public" equals 1 if the institution is a public college and 2 if it is a private college.

"Faculty" is the number of tenured or tenure-track economics department faculty members.

"Bschol" equals 1 if the college has a business program and 0 if not.

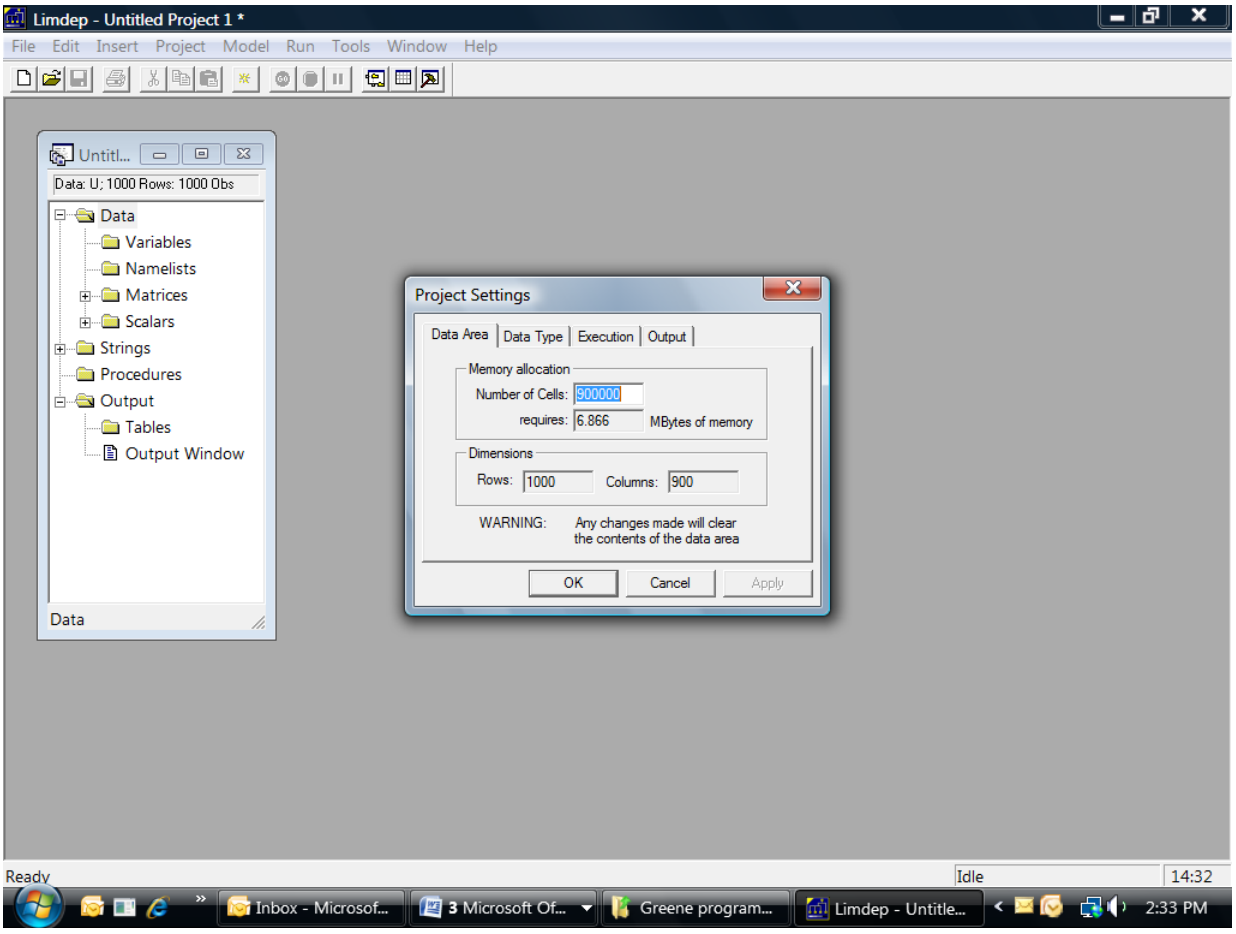
"T" is the time trend running from −7 to 8, corresponding to years from 1996 through 2006.

"MA_Deg" is a three-year moving average of degrees (unknown for the first two years).

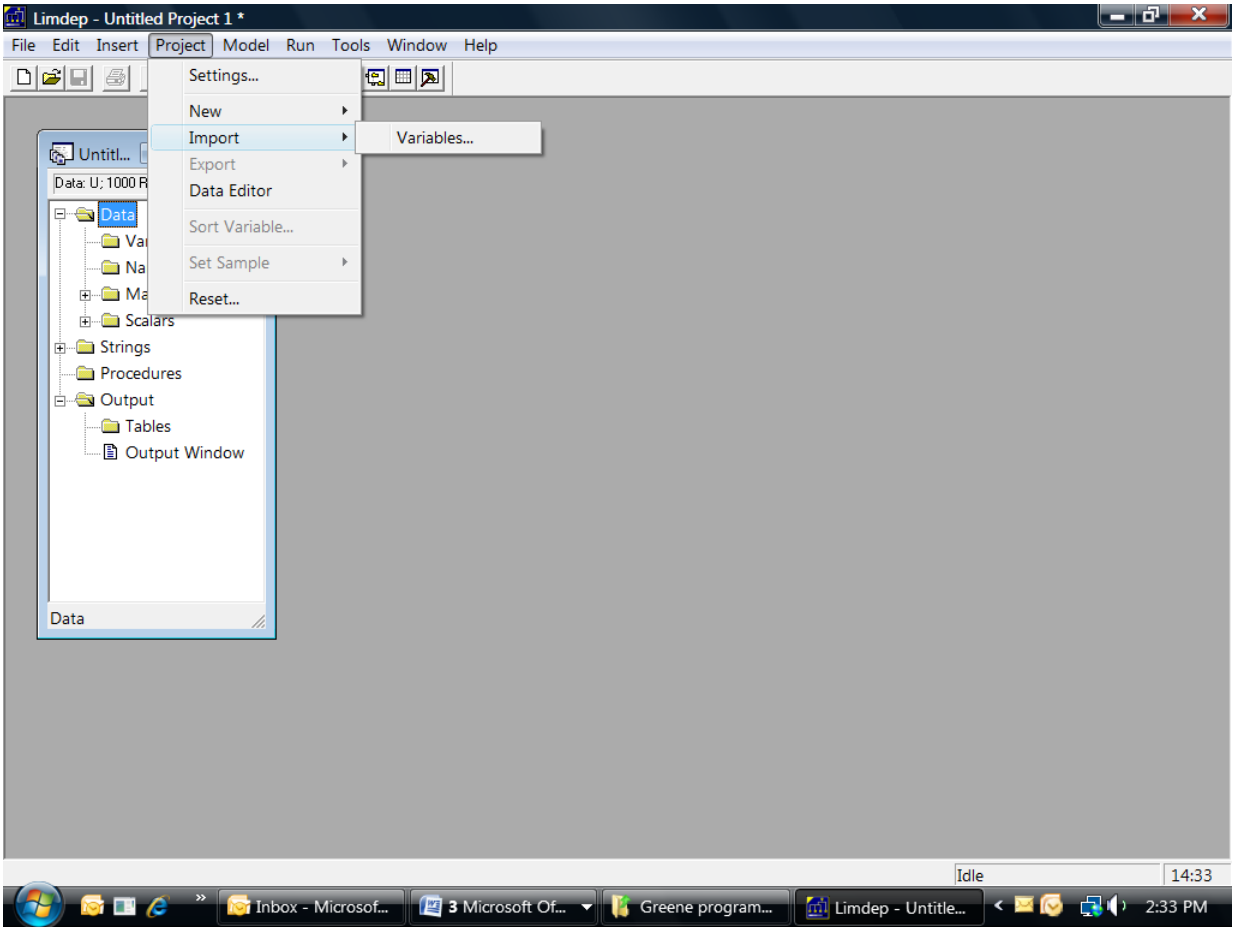
College	Year	Degrees	DegreBar	Public	Faculty	Bschol	T	MA_Deg
1	1991	50	47.375	2	11	1	-7	0
1	1992	32	47.375	2	8	1	-6	0
1	1993	31	47.375	2	10	1	-5	37.667
1	1994	35	47.375	2	9	1	-4	32.667
↓	↓	↓	↓	↓	↓	↓		↓
1	2003	57	47.375	2	7	1	5	56
1	2004	57	47.375	2	10	1	6	55.667
1	2005	57	47.375	2	10	1	7	57
1	2006	51	47.375	2	10	1	8	55
2	1991	16	8.125	2	3	1	-7	0
2	1992	14	8.125	2	3	1	-6	0
2	1993	10	8.125	2	3	1	-5	13.333
↓	↓	↓	↓	↓	↓	↓		↓
2	2004	10	8.125	2	3	1	6	12.667
2	2005	7	8.125	2	3	1	7	11.333
2	2006	6	8.125	2	3	1	8	7.667
3	1991	40	35.5	2	8	1	-7	0
3	1992	31	37.125	2	8	1	-6	0
↓	↓	↓	↓	↓	↓	↓		↓
17	2004	64	39.3125	2	5	0	6	54.667
17	2005	37	39.3125	2	4	0	7	51.333
17	2006	53	39.3125	2	4	0	8	51.333
18	1991	14	8.4375	2	4	0	-7	0
18	1992	10	8.4375	2	4	0	-6	0
18	1993	10	8.4375	2	4	0	-5	11.333
18	1994	7	8.4375	2	3.5	0	-4	9
↓	↓	↓	↓	↓	↓	↓		↓
18	2005	4	8.4375	2	2.5	0	7	7.333
18	2006	7	8.4375	2	3	0	8	6

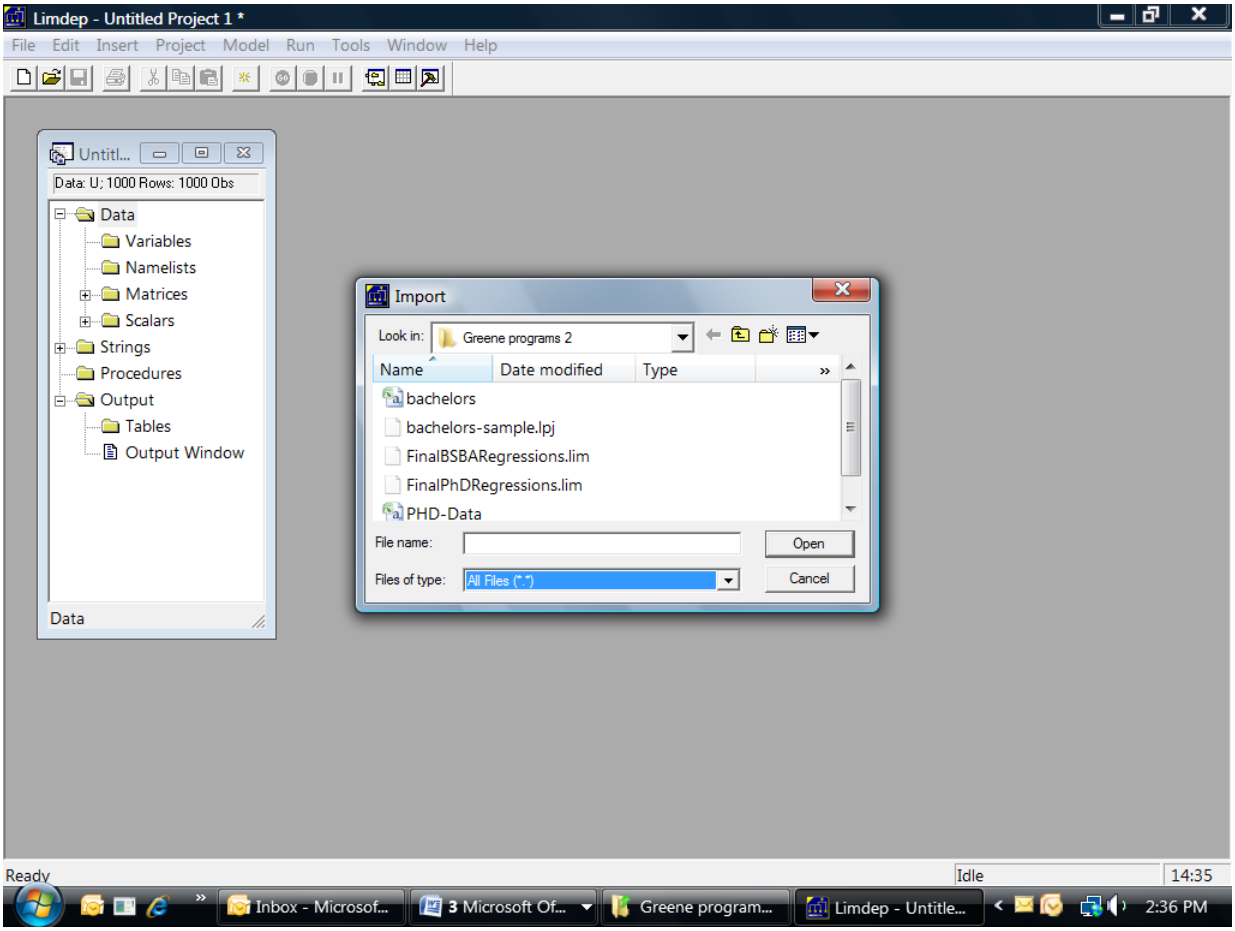
If you opened this CSV file in a word processor or text editing program, it would show that each of the 289 lines (including the headers) corresponds to a row in the EXCEL table, but variable values would be separated by commas and not appear neatly one on top of the other as in EXCEL.

As discussed in Module One, Part Two, older versions of LIMDEP (NLOGIT) have a data matrix default restriction of no more than 222 rows (records per variable), 900 columns (number of variables) and 200,000 cells. LIMDEP 9 and NLOGIT 4.0 automatically adjust the data constraints but in older versions the number of cells must be increased to accommodate work with our data set. After opening LIMDEP, the number of working cells can be increased by clicking the Project button on the top ribbon, going to Settings, and changing the number of cells. Going from the default 200,000 cells to 900,000 cells (1,000 Rows and 900 columns) is more than sufficient for this panel data set.



We could write a "READ" command to bring this text data file into LIMDEP but like EXCEL it can be imported into LIMDEP directly by clicking the Project button on the top ribbon, going to Import, and then clicking on Variables, from which the bachelors.cvs file can be located wherever it is stored (in our case in the "Greene programs 2" folder). Hitting the Open button will bring the data set into LIMDEP, which can be checked by clicking the "Activate Data Editor" button, which is second from the right on the tool bar or go to Data Editor in the Window's menu, as described and demonstrated in Module One, Part Two.





In addition to a visual inspection of the data via the "Activate Data Editor," we use the "dstat" command to check the descriptive statistics. First, however, we need to remove the two years (1991 and 1992) for which no data are available for the degree moving average measure. This is done with the "Reject" command. In our "File:New Text/Command Document" (which was described in Module One, Part Two), we have

```
reject ; year < 1993 $
dstat;rhs=*$
```

which upon highlighting and pressing "Go" yields

```
--> reject ; year < 1993 $
--> dstat;rhs=*$
Descriptive Statistics
All results based on nonmissing observations.
================================================================================
Variable     Mean       Std.Dev.     Minimum      Maximum        Cases Missing
================================================================================
All observations in current sample
--------+-------------------------------------------------------------------------
COLLEGE |   9.50000      5.19845      1.00000      18.0000          252       0
YEAR    |   1999.50      4.03915      1993.00      2006.00          252       0
DEGREES |   23.1111      19.2264      .000000      81.0000          252       0
DEGREBAR|   23.6528      18.0143      2.00000      62.4375          252       0
PUBLIC  |   1.77778      .416567      1.00000      2.00000          252       0
FACULTY |   6.51786      3.13677      2.00000      14.0000          252       0
BSCHOOL |   .388889      .488468      .000000      1.00000          252       0
T       |   1.50000      4.03915     -5.00000      8.00000          252       0
MA_DEG  |   23.1931      18.5540      1.33333      80.0000          252       0
```

CONSTANT COEFFICIENT REGRESSION

The constant coefficient panel data model for the faculty size data-generating process for bachelor degree-granting undergraduate departments is given by

$$Faculty\ size_{it} = \beta_1 + \beta_2 T_t + \beta_3 BA\&S_{it} + \beta_4 MEANBA\&S_i + \beta_5 PUBLIC_i + \beta_6 Bschl + \beta_7 MA\_Deg_{it} + \varepsilon_{it}$$

where the error term ε_{it} is independent and identically distributed (*iid*) across institutions and over time and $E(\varepsilon_{it}^2|\mathbf{x}_{it}) = \sigma^2$, for $I = 18$ colleges and $T = 14$ years (−5 through 8) for 252 complete records. The LIMDEP OLS regression command that needs to be entered into the command document (again, following the procedure for opening the command document window shown in Module One, Part Two), including the standard error adjustment for clustering is

```
reject ; year < 1993 $
regress
;lhs=faculty;rhs=one,t,degrees,degrebar,public,bschool,MA_deg
;cluster=14$
```

Upon highlighting and hitting the "Go" button the Output file shows the following results

```
--> reject ; year < 1993 $
--> regress;lhs=faculty;rhs=one,t,degrees,degrebar,public,bschool,MA_deg
    ;cluster=14$
```

```
+---------------------------------------------------+
| Ordinary    least squares regression              |
| LHS=FACULTY  Mean                 =   6.517857    |
|              Standard deviation   =   3.136769    |
|              Number of observs.   =        252    |
| Model size   Parameters           =          7    |
|              Degrees of freedom   =        245    |
| Residuals    Sum of squares       =   868.4410    |
|              Standard error of e  =   1.882726    |
| Fit          R-squared            =   .6483574    |
|              Adjusted R-squared   =   .6397458    |
| Model test   F[  6,   245] (prob) =  75.29 (.0000) |
| Diagnostic   Log likelihood       =  -513.4686    |
|              Restricted(b=0)      =  -645.1562    |
|              Chi-sq [  6]  (prob) = 263.38 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =   1.292840    |
|              Akaike Info. Criter. =   1.292826    |
|              Bayes Info. Criter.  =   1.390866    |
| Autocorrel   Durbin-Watson Stat.  =   .3295926    |
|              Rho = cor[e,e(-1)]   =   .8352037    |
| Model was estimated Jul 16, 2009 at 04:21:28PM    |
+---------------------------------------------------+
+-------------------------------------------------------------------+
| Covariance matrix for the model is adjusted for data clustering.  |
| Sample of    252 observations contained     18 clusters defined by |
|     14 observations (fixed number) in each cluster.               |
| Sample of    252 observations contained      1 strata defined by  |
|    252 observations (fixed number) in each stratum.               |
+-------------------------------------------------------------------+
+--------+-------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |t-ratio |P[|T|>t]| Mean of X|
+--------+-------------+----------------+--------+--------+----------+
|Constant|    10.1397***       .91063       11.135   .0000            |
|T       |    -.02809          .02227       -1.261   .2083     1.50000|
|DEGREES |    -.01636          .01866        -.877   .3814     23.1111|
|DEGREBAR|     .10832***       .03378        3.206   .0015     23.6528|
|PUBLIC  |   -3.86239***       .56950       -6.782   .0000     1.77778|
|BSCHOOL |     .58112          .94253         .617   .5381      .38889|
|MA_DEG  |     .03780**        .01810        2.089   .0377     23.1931|
+--------+------------------------------------------------------------+
| Note: ***, **, * = Significance at 1%, 5%, 10% level.              |
+--------------------------------------------------------------------+
```

Contemporaneous degrees have little to do with current faculty size but both overall number of degrees awarded (the school means) and the moving average of degrees (MA_DEG) have significant effects. It takes an increase of 26 or 27 bachelor degrees in the moving average to expect just one more faculty position. Whether it is a public or a private college is highly significant. Moving from a public to a private college lowers predicted faculty size by nearly four members for otherwise comparable institutions. There is an insignificant erosion of tenured and tenure-track faculty size over time. Finally, while economics departments in colleges with a business school tend to have a larger permanent faculty, ceteris paribus, the effect is small and insignificant.

FIXED-EFFECTS REGRESSION

To estimate the fixed-effects model we can either insert seventeen (0,1) covariates to capture the unique effect of each of the 18 colleges (where each of the 17 dummy coefficients are measured relative to the constant term) or the insert of 18 dummy variables with no overall constant term in the OLS regression. The results for the other coefficients and for R^2 will be identical either way, while the the constant terms, since they measure the difference of each college from the 18^{th} in the first case, or the difference of all 18 from zero in the second, will differ. This difference is inconsequential for the regression of interest. Which way the model is estimated is purely a matter of convenience and preference.

An important implication of the fixed effects specification is that no time invariant variables can be included in the equation because they would be perfectly correlated with the respective college dummies. Thus, the overall school mean number of degrees, the public or private dummy variables, and business school dummy variables must all be excluded from the fixed effects model.

A LIMDEP (NLOGIT) program to be run from the Test/Command Document, including the commands to create the dummy variables then run the regression is shown below. (An alternative, more compact way to create the dummies and run the regression is shown in the Appendix.)

```
reject ; year < 1993 $

create
;Col1=college=1
;Col2=college=2
;Col3=college=3
;Col4=college=4
;Col5=college=5
;Col6=college=6$
create
;Col7=college=7
;Col8=college=8
;Col9=college=9
;Col10=college=10
;Col11=college=11
;Col12=college=12$
create
;Col13=college=13
;Col14=college=14
;Col15=college=15
;Col16=college=16
;Col17=college=17
;Col18=college=18$

regress;lhs=faculty;rhs=one,t,degrees,MA_deg,
Col1,Col2,Col3,Col4,Col5,Col6,Col7,Col8,Col9,
Col10,Col11,Col12,Col13,Col14,Col15,Col16,Col17; cluster=14$
```

The resulting regression information appearing in the output window is

```
+----------------------------------------------------------------------+
| Covariance matrix for the model is adjusted for data clustering.     |
| Sample of    252 observations contained     18 clusters defined by   |
|     14 observations (fixed number) in each cluster.                  |
+----------------------------------------------------------------------+

----------------------------------------------------------------------
Ordinary     least squares regression .............
LHS=FACULTY  Mean                 =        6.51786
             Standard deviation   =        3.13677
             Number of observs.   =            252
Model size   Parameters           =             21
             Degrees of freedom   =            231
Residuals    Sum of squares       =      146.63709
             Standard error of e  =         .79674
Fit          R-squared            =         .94062
             Adjusted R-squared   =         .93548
Model test   F[ 20,   231] (prob) =   183.0(.0000)
Diagnostic   Log likelihood       =     -289.34751
             Restricted(b=0)      =     -645.15625
             Chi-sq [ 20]  (prob) =   711.6(.0000)
Info criter. LogAmemiya Prd. Crt. =        -.37441
             Akaike Info. Criter. =        -.37480
             Bayes Info. Criter.  =        -.08068
Model was estimated on Sep 23, 2009 at 06:44:38 PM
--------+-------------------------------------------------------------
Variable| Coefficient    Standard Error  t-ratio  P[|T|>t]   Mean of X
--------+-------------------------------------------------------------
Constant|    2.69636***       .15109       17.846   .0000
       T|    -.02853          .02245       -1.271   .2051      1.50000
 DEGREES|    -.01608          .01521       -1.058   .2913      23.1111
  MA_DEG|     .03985***       .01485        2.683   .0078      23.1931
    COL1|    5.77747***       .76816        7.521   .0000       .05556
    COL2|     .15299***       .01343       11.392   .0000       .05556
    COL3|    4.29759***       .55420        7.755   .0000       .05556
    COL4|    6.28973***       .65533        9.598   .0000       .05556
    COL5|    4.91094***       .56987        8.618   .0000       .05556
    COL6|    5.02016***       .02561      196.041   .0000       .05556
    COL7|    1.21384***       .01321       91.876   .0000       .05556
    COL8|     .77797***       .06785       11.466   .0000       .05556
    COL9|    3.16474***       .06270       50.478   .0000       .05556
   COL10|    2.86345***       .15540       18.427   .0000       .05556
   COL11|    5.15181***       .02403      214.385   .0000       .05556
   COL12|    -.06802***       .02153       -3.160   .0018       .05556
   COL13|    3.98895***      1.01415        3.933   .0001       .05556
   COL14|    -.63196***       .11986       -5.272   .0000       .05556
   COL15|    8.25859***       .47255       17.477   .0000       .05556
   COL16|    8.00970***       .55461       14.442   .0000       .05556
   COL17|     .43544          .59258         .735   .4632       .05556
--------+-------------------------------------------------------------
Note: ***, **, * = Significance at 1%, 5%, 10% level.
----------------------------------------------------------------------
```

Once again, contemporaneous degrees is not a driving force in faculty size. An F test is not needed to assess if at least one of the 17 colleges differ from college 18. With the exception of college 17, each of the other colleges are significantly different. The moving average of degrees is again significant.

The preceding approach, of computing all the dummy variables and building them into the regression, is likely to become unduly cumbersome if the number of colleges (units) is very large. Most contemporary software, including LIMDEP will do this computation automatically without explicitly computing the dummy variables and including them in the equation. As an alternative to specifying all the dummies in the regression command, the same results can be obtained with the simpler "FixedEffects" command:

```
regress;lhs=faculty;rhs=one,t,degrees,MA_deg
;Panel;Str=College
;FixedEffects;Robust$

---------------------------------------------------------------------
Least Squares with Group Dummy Variables.........
Ordinary     least squares regression ............
LHS=FACULTY  Mean                 =        6.51786
             Standard deviation   =        3.13677
             Number of observs.   =            252
Model size   Parameters           =             21
             Degrees of freedom   =            231
Residuals    Sum of squares       =      146.63709
             Standard error of e  =         .79674
Fit          R-squared            =         .94062
             Adjusted R-squared   =         .93548
Model test   F[ 20,   231] (prob) =   183.0(.0000)
Diagnostic   Log likelihood       =     -289.34751
             Restricted(b=0)      =     -645.15625
             Chi-sq [ 20]  (prob) =   711.6(.0000)
Info criter. LogAmemiya Prd. Crt. =        -.37441
             Akaike Info. Criter. =        -.37480
             Bayes Info. Criter.  =        -.08068
Model was estimated on Sep 23, 2009 at 06:44:38 PM
Estd. Autocorrelation of e(i,t)   =        .293724
Robust cluster corrected covariance matrix used
Panel:Groups Empty      0,     Valid data       18
             Smallest  14,     Largest          14
             Average group size in panel     14.00
--------+------------------------------------------------------------
Variable| Coefficient    Standard Error  t-ratio  P[|T|>t]   Mean of X
--------+------------------------------------------------------------
       T|    -.02853          .02245       -1.271   .2050      1.50000
 DEGREES|    -.01608          .01521       -1.058   .2912      23.1111
  MA_DEG|     .03985***       .01485        2.683   .0078      23.1931
--------+------------------------------------------------------------
Note: ***, **, * = Significance at 1%, 5%, 10% level.
---------------------------------------------------------------------
```

RANDOM-EFFECTS REGRESSION

Finally, consider the random-effects model in which we employ Mundlak's (1978) approach to estimating panel data. The Mundlak model posits that the fixed effects in the equation, β_{1i} , can be projected upon the group means of the time-varying variables, so that

$$\beta_{1i} = \beta_1 + \delta' \bar{x}_i + w_i$$

where $\bar{x}_i$ is the set of group (school) means of the time-varying variables and w_i is a (now) random effect that is uncorrelated with the variables and disturbances in the model. Logically, adding the means to the equations picks up the correlation between the school effects and the other variables. We could not incorporate the mean number of degrees awarded in the fixed-effects model (because it was time invariant) but this variable plays a critical role in the Mundlak approach to panel data modeling and estimation.

The random effects model for BA and BS degree-granting undergraduate departments is

$$FACULTY\ size_{it} = \beta_1 + \beta_2 T_t + \beta_3 BA\&S_{it} + \beta_4 MEANBA\&S_i + \beta_5 MOVAVBA\&BS + \beta_6 PUBLIC_i + \beta_7 Bschl + \varepsilon_{it} + u_i$$

where error term ε is *iid* over time, $E(\varepsilon_{it}^2|\mathbf{x}_{it}) = \sigma^2$ for $I = 18$ and $T_i = 14$ and $E[u_i^2] = \theta^2$ for $I = 18$. The LIMDEP program to be run from the Text/Command Document (with 1991 and 1992 data suppressed) is

```
regress
;lhs=faculty
;rhs=one,t,degrees,degrebar,public,bschool,MA_deg
;pds=14
;panel
;random
;robust$
```

The resulting regression information appearing in the output window is

```
--> regress
    ;lhs=faculty;rhs=one,t,degrees,degrebar,public,bschool,MA_deg
    ;pds=14;panel;random;robust$

---------------------------------------------------------------------
OLS Without Group Dummy Variables.................
Ordinary     least squares regression .............
LHS=FACULTY  Mean                 =        6.51786
             Standard deviation   =        3.13677
             Number of observs.   =            252
Model size   Parameters           =              7
             Degrees of freedom   =            245
Residuals    Sum of squares       =      868.44104
             Standard error of e  =        1.88273
Fit          R-squared            =         .64836
             Adjusted R-squared   =         .63975
Model test   F[  6,   245] (prob) =    75.3(.0000)
Diagnostic   Log likelihood       =     -513.46861
             Restricted(b=0)      =     -645.15625
             Chi-sq [  6]  (prob) =   263.4(.0000)
Info criter. LogAmemiya Prd. Crt. =        1.29284
             Akaike Info. Criter. =        1.29283
             Bayes Info. Criter.  =        1.39087
Model was estimated on Sep 23, 2009 at 07:17:22 PM
Panel Data Analysis of FACULTY           [ONE way]
              Unconditional ANOVA (No regressors)
Source         Variation  Deg. Free.   Mean Square
Between       2312.22321         17.     136.01313
Residual       157.44643        234.        .67285
Total         2469.66964        251.       9.83932
--------+------------------------------------------------------------
Variable| Coefficient    Standard Error  t-ratio  P[|T|>t]   Mean of X
--------+------------------------------------------------------------
       T|    -.02809          .03030        -.927   .3549      1.50000
 DEGREES|    -.01636          .02334        -.701   .4839      23.1111
DEGREBAR|     .10832***       .02047        5.293   .0000      23.6528
  PUBLIC|   -3.86239***       .29652      -13.026   .0000      1.77778
 BSCHOOL|     .58112**        .25115        2.314   .0215       .38889
  MA_DEG|     .03780          .02907        1.300   .1947      23.1931
Constant|    10.1397***       .52427       19.341   .0000
--------+------------------------------------------------------------
Note: ***, **, * = Significance at 1%, 5%, 10% level.
---------------------------------------------------------------------
```

```
+----------------------------------------------------+
| Panel:Groups   Empty       0,   Valid data      18 |
|                Smallest   14,   Largest         14 |
|                Average group size            14.00 |
| There are  3 vars. with no within group variation. |
| DEGREBAR PUBLIC   BSCHOOL                          |
+----------------------------------------------------+

--------------------------------------------------------------------------
Random Effects Model: v(i,t)    = e(i,t) + u(i)
Estimates:  Var[e]              =       .643145
            Var[u]              =      2.901512
            Corr[v(i,t),v(i,s)] =       .818559
Lagrange Multiplier Test vs. Model (3) =1096.30
( 1 degrees of freedom, prob. value =  .000000)
(High values of LM favor FEM/REM over CR model)
Baltagi-Li form of LM Statistic =       1096.30
            Sum of Squares           868.488173
            R-squared                   .648338
Robust cluster corrected covariance matrix used
--------+-----------------------------------------------------------------
Variable| Coefficient    Standard Error  b/St.Er. P[|Z|>z]   Mean of X
--------+-----------------------------------------------------------------
       T|    -.02853          .02146       -1.329   .1838      1.50000
 DEGREES|    -.01609          .01793        -.897   .3696      23.1111
DEGREBAR|     .10610***       .03228        3.287   .0010      23.6528
  PUBLIC|   -3.86365***       .54685       -7.065   .0000      1.77778
 BSCHOOL|     .58176          .90497         .643   .5203       .38889
  MA_DEG|     .03981**         .01728        2.305   .0212      23.1931
Constant|    10.1419***       .87456       11.597   .0000
--------+-----------------------------------------------------------------
Note: ***, **, * = Significance at 1%, 5%, 10% level.
--------------------------------------------------------------------------
```

The marginal effect of an additional economics major is again insignificant but slightly negative within the sample. Both the short-term moving average number and long-term average number of bachelor degrees are significant. A long-term increase of about 10 students earning degrees in economics is required to predict that one more tenured or tenure-track faculty member is in a department. Ceteris paribus, economics departments at private institutions are smaller than comparable departments at public schools by a large and significant number of four members. Whether there is a business school present is insignificant. There is no meaningful trend in faculty size.

CONCLUDING REMARKS

The goal of this hands-on component of this third of four modules is to enable economic education researchers to make use of panel data for the estimation of constant coefficient, fixed-effects and random-effects panel data models in LIMDEP (NLOGIT). It was not intended to explain all of the statistical and econometric nuances associated with panel data analysis. For this an intermediate level econometrics textbook (such as Jeffrey Wooldridge, *Introductory Econometrics*) or advanced econometrics textbook (such as William Greene, *Econometric Analysis*) should be consulted.

APPENDIX:

ALTERNATIVES FOR CREATING COLLEGE DUMMY VARIABLES AND RUNNING REGRESSIONS IN LIMDEP (NLOGIT)

There are two alternative ways to create the college dummy variables for use in the regression.

First is the "`create ; expand(college)$`" command, where COLLEGE is expanded as _COLLEG_, with the following resulting output:

```
COLLEGE  was expanded as _COLLEG_.
Largest value =  18.  18 New variables were created.
Category
  1  New variable = COLLEG01    Frequency=      14
  2  New variable = COLLEG02    Frequency=      14
  3  New variable = COLLEG03    Frequency=      14
  4  New variable = COLLEG04    Frequency=      14
  5  New variable = COLLEG05    Frequency=      14
  6  New variable = COLLEG06    Frequency=      14
  7  New variable = COLLEG07    Frequency=      14
  8  New variable = COLLEG08    Frequency=      14
  9  New variable = COLLEG09    Frequency=      14
 10  New variable = COLLEG10    Frequency=      14
 11  New variable = COLLEG11    Frequency=      14
 12  New variable = COLLEG12    Frequency=      14
 13  New variable = COLLEG13    Frequency=      14
 14  New variable = COLLEG14    Frequency=      14
 15  New variable = COLLEG15    Frequency=      14
 16  New variable = COLLEG16    Frequency=      14
 17  New variable = COLLEG17    Frequency=      14
 18  New variable = COLLEG18    Frequency=      14
Note, this is a complete set of dummy variables.  If
you use this set in a regression, drop the constant.
```

The second method for creating dummies and running the regression is an even more condensed, and as yet an undocumented feature in the LIMDEP (NLOGIT) manual:

```
regress;lhs=faculty;rhs=one,t,degrees,MA_deg,expand(college)
       ;cluster=college$
```

which produces the same results as the fixed effects regression command we used at the beginning of this duscussion.

REFERENCES

Becker, William, William Greene and John Siegfried (2009). "Does Teaching Load Affect Faculty Size? " Working Paper (July).

Greene, William (2008). *Econometric Analysis*. 6th Edition, New Jersey: Prentice Hall.

Mundlak, Yair (1978). "On the Pooling of Time Series and Cross Section Data," *Econometrica.* Vol. 46. No. 1 (January): 69-85.

Wooldridge, Jeffrey (2009). *Introductory Econometrics*. 4th Edition, Mason OH: South-Western.