Department of Economics

Econometric Analysis of Panel Data

Professor William Greene **Phone: 212.998.0876**
Office: KMC 7-78 **Home page:www.stern.nyu.edu/~wgreene**
Email: wgreene@stern.nyu.edu
URL for course web page:
www.stern.nyu.edu/~wgreene/Econometrics/PanelDataEconometrics.htm

Assignment 1

Part I. Mathematical Statistics

The density f(*y*) for a nonnegative random variable, *y*, is exponential with parameter λ, so

$f(y) = 1/\lambda \exp(-y/\lambda)$, $y \geq 0$, $\lambda > 0$.

For this random variable, the mean is E[*y*] = λ. We make this a regression model by formulating the conditional mean function

$\lambda(x) = \exp(\alpha+\beta x)$.

(This makes it a 'loglinear model.'). Now, the regression function is

$E[y|x]=\exp(\alpha+\beta x)$.

Suppose, further, that *x* is distributed uniformly with density

$f(x) = 1$, $0 \leq x \leq 1$.

Note that with this assumption, the joint density of *y* and *x* is

$f(y,x) = f(y|x)\, f(x) = [1/\exp(\alpha+\beta x)] \exp[-y/\exp(\alpha+\beta x)]$.

1. Derive the parameters of the linear projection,

$P(x) = \delta_0 + \delta_1 x$,

where $\delta_0 = E[y] - \delta_1 E[x]$ and

$\delta_1 = Cov[x,y]/Var[x]$.

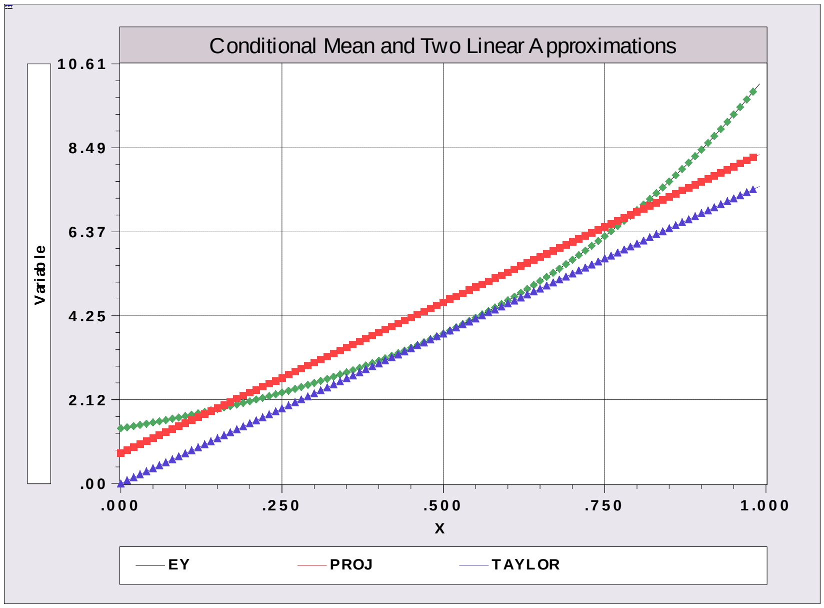
Suppose α = 1/3 and β = 2. What are the values of δ_0 and δ_1?

Hint: $E[y] = E_x E[y|x] = \int_0^1 \exp(\alpha+\beta x)\times 1 dx = \exp(\alpha)\int_0^1 \exp(\beta x)dx$ and

$Cov[y,x] = Cov[x,E[y|x]] = E_x\{x\times E[y|x]\} - E[x]\,E[y] = \exp(\alpha)\int_0^1 x\exp(\beta x)dx - E[x]E[y]$.

Find help at http://en.wikipedia.org/wiki/List_of_integrals_of_exponential_functions

2. Consider the linear Taylor series approximation to the conditional mean function. What are the values of θ_0 and θ_1 in the Taylor series: $E^*[y|x] = \theta_0 + \theta_1 x$ when the expansion point is E[x] = 1/2 and as before, α = 1/3 and β = 2.



1. $f(x) = 1,\ 0 \le x \le 1$. $f(y|x) = [1/\lambda(x)]\exp[-y/\lambda(x)],\ y \ge 0,\ \lambda(x) = \alpha + \beta x,\ \alpha = 1/3,\ \beta = 2$.
$E[x] = ½$, $\text{Var}[x] = 1/12$. (Standard results)
$E[y] = E_x E[y|x] = E[\lambda(x)] = E[\exp(\alpha + \beta x)]$

$$E[y] = \exp(\alpha)E[\exp(\beta x)]$$

$$= \exp(\alpha)\int_0^1 \exp(\beta x)dx$$

$$= \exp(\alpha)\left[\frac{\exp(\beta x)}{\beta}\Big|_0^1\right] = \exp(\alpha)\left[\frac{\exp(\beta)}{\beta} - \frac{1}{\beta}\right] = \frac{\exp(\alpha)}{\beta}[\exp(\beta) - 1] = 4.5832$$

$$Cov(x,y) = Cov(e, E[y\,|\,x]) = E[x\exp(\alpha + \beta x)] - E[x]E[y]$$

$$= \int_0^1 x\exp(\alpha + \beta x)dx - \frac{1}{2}4.5832$$

$$= \exp(\alpha)\int_0^1 x\exp(\beta x)dx - \frac{1}{2}4.5832$$

$$= \exp(\alpha)\left[\frac{\exp(\beta x)}{\beta^2}(\beta x - 1)\,|_0^1\right] - \frac{1}{2}4.5832$$

$$= \exp(\alpha)\left[\frac{\exp(\beta)}{\beta^2}(\beta - 1) - \frac{1}{\beta^2}(-1)\right] - \frac{1}{2}4.5832$$

$$= \frac{\exp(\alpha)}{\beta^2}\left[\exp(\beta)(\beta - 1) + 1\right] - \frac{1}{2}4.5832 = .697808.$$

The slope is $Cov(x,y)/\text{Var}[x] = .697808/(1/12) = 8.37369$
The constant term is $E[y] - \text{slope}\times E[x] = 4.5832 - 8.37369(1/2) = 0.27147$

2. If $E[y|x] = \lambda(x)$, The Taylor series approximation would be
$\lambda(E[x]\} + \partial\lambda(x)/\partial x\ |(x{=}E[x]) \times (x - E[x])$
$= \lambda(E[x])[1 - \beta E[x]$

$E^*(y|x) = \lambda(1/2) + \partial\lambda(x)/\partial x|_{1/2}\ [x - ½]$
$\quad = \exp(1/3 + 2(1/2)) - (½)\beta\lambda(1/2) + \beta\lambda(1/2)x$

The slope is 2exp(1/3 + 2(1/2)) = 2exp(4/3) = 7.587.
The constant is $\lambda(1/2)[1 - ½\beta] = 0$.

Part II. Linear Regression Analysis

Data for this exercise are on the course website – please use the "Cornwell and Rupert Returns to Schooling Data." We begin with the linear regression model (using the variable names in the data set)

(*) $LWAGE_{it} = \beta_1 + \beta_2 OCC_{it} + \beta_3 SMSA_{it} + \beta_4 MS_{it} + \beta_5 FEM_i + \beta_6 ED_i + \beta_7 EXP_{it} + \varepsilon_{it}$

The dependent variable is log wage. The RHS variables are defined in the data set. Although this is a panel data set, we are going to ignore that aspect and "pool" the data.

1. Compute the linear least squares regression results and report the coefficients, standard errors, 't-ratios,' R^2, adjusted R^2, residual standard deviation, and *F* statistic for testing the joint significance of all the variables in the equation.

2. Test the hypothesis that neither education (ED) nor experience (EXP) is a significant determinant of the expected log wage. Use an *F* (Wald), likelihood ratio (assuming normality of ε), and a Lagrange multiplier (also assuming normality) test. In each case, document in minute detail exactly how you are computing your results and what conclusion you reach.

3. The model contains a dummy variable for sex, FEM = 1 for female, 0 for male. What is the value of the coefficient on FEM in your estimated model? How do you interpret this value? I.e., what is the economic meaning of the value you computed for this coefficient? Test the hypothesis that this coefficient equals zero.
GENDER DIFFERENCE (PROPORTIONAL). -39%. HYPOTHESIS THAT IT EQUALS ZERO IS REJECTED BASED ON T RATIO OF -15.164.

```
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=LWAGE    Mean                 =    6.676346     |
|              Standard deviation   =    .4615122     |
| WTS=none     Number of observs.   =        4165     |
| Model size   Parameters           =           7     |
|              Degrees of freedom   =        4158     |
| Residuals    Sum of squares       =    556.3030     |
|              Standard error of e  =    .3657745     |
| Fit          R-squared            =    .3727592     |
|              Adjusted R-squared   =    .3718541     |
| Model test   F[  6,  4158] (prob) = 411.84 (.0000)  |
| Diagnostic   Log likelihood       =  -1717.476      |
|              Restricted(b=0)      =  -2688.806      |
|              Chi-sq [   6]  (prob) =1942.66 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -2.009797      |
|              Akaike Info. Criter. =  -2.009797      |
|              Bayes Info. Criter.  =  -1.999151      |
+----------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Constant	5.66098***	.04685914	120.808	.0000	
OCC	-.11220***	.01464317	-7.662	.0000	.5111645
SMSA	.15504***	.01233744	12.567	.0000	.6537815

```
|MS      |      .09569***        .02133490     4.485   .0000     .8144058|
|FEM     |     -.39478***        .02603413   -15.164   .0000     .1126050|
|ED      |      .05688***        .00267743    21.244   .0000    12.845378|
|EXP     |      .01044***        .00054206    19.256   .0000    19.853782|
+--------+------------------------------------------------------------+
| Note: ***, **, * = Significance at 1%, 5%, 10% level.                |
+----------------------------------------------------------------------+
--> MATRIX ; List ; Wald = b2'*<v2>*b2 $

Matrix WALD     has  1 rows and  1 columns.
               1
        +--------------
       1|  668.77622
--> CALC ; Logl1 = Logl $
--> Regress ; Lhs = Lwage ; Rhs = One,OCC,SMSA,MS,FEM ; Res = e0 $

+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=LWAGE    Mean                 =   6.676346     |
|              Standard deviation   =   .4615122     |
| WTS=none     Number of observs.   =       4165     |
| Model size   Parameters           =          5     |
|              Degrees of freedom   =       4160     |
| Residuals    Sum of squares       =   645.7792     |
|              Standard error of e  =   .3939992     |
| Fit          R-squared            =   .2718733     |
|              Adjusted R-squared   =   .2711731     |
| Model test   F[  4,  4160] (prob) = 388.32 (.0000) |
| Diagnostic   Log likelihood       =  -2028.070     |
|              Restricted(b=0)      =  -2688.806     |
|              Chi-sq [  4]  (prob) =1321.47 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -1.861613     |
|              Akaike Info. Criter. =  -1.861613     |
|              Bayes Info. Criter.  =  -1.854009     |
| Autocorrel   Durbin-Watson Stat.  =   .7730622     |
|              Rho = cor[e,e(-1)]   =   .6134689     |
| Model was estimated Feb 10, 2009 at 06:49:39AM     |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
|Constant|    6.61346***        .02495768   264.987   .0000            |
|OCC     |    -.28539***        .01249653   -22.838   .0000     .5111645|
|SMSA    |     .19827***        .01316629    15.059   .0000     .6537815|
|MS      |     .15351***        .02274471     6.749   .0000     .8144058|
|FEM     |    -.40738***        .02799278   -14.553   .0000     .1126050|
+--------+------------------------------------------------------------+
| Note: ***, **, * = Significance at 1%, 5%, 10% level.                |
+----------------------------------------------------------------------+

--> Create ; e02 = e0*e0 $
--> CALC ; Logl0 = Logl $
--> Matrix ; List ; LMtest = e0'XPart2 * <XPart2'[E02]XPart2> * Xpart2'e0 $

Matrix LMTEST   has  1 rows and  1 columns.
               1
        +--------------
       1|  411.91215
--> CALC ; list ; LR = 2*(Logl1 - Logl0) $
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 LR      =    621.187281
```

Part III. Structural Change

The implication of the specification of FEM in the model in Part II is that the extent of the difference between men and women is captured in a shift of the regression function (based on a change in the intercept alone). Consider, instead, the hypothesis that different regression functions apply to men and women. Fit the model separately for men and women, then use a Chow test to test the null hypothesis that the same equation applies to men and women. (Note, for purposes of this exercise, your model will not contain the FEM variable.) The model is

$$(**) \quad LWAGE_{it} = \beta_1 + \beta_2 OCC_{it} + \beta_3 SMSA_{it} + \beta_4 MS_{it} + \beta_5 ED_i + \beta_6 EXP_{it} + \varepsilon_{it}$$

Completely document your analysis. Include in your results a table that shows the results of the three regressions, male, female and pooled, so that the reader can easily see the comparison of the estimated coefficients. What is the result of the test?

Looking ahead to our work in panel data modeling, repeat this analysis for the 7 years of data in the sample. That is, compute the regression in (**) using the full pooled data set, then again for each of the 7 years. (There are 595 observations for each of the 7 years.) Using a Chow (F) test, test the null hypothesis that the same model applies to all 7 years. To investigate whether a structural change might be explained by a simple shift of the function, fit the model

$$(***) \quad LWAGE_{it} = \beta_1 + \beta_2 OCC_{it} + \beta_3 SMSA_{it} + \beta_4 MS_{it} + \beta_5 ED_i + \beta_6 EXP_{it} + \gamma_1 T_{2,t} + \ldots \ \gamma_6 T_{6,t} + \varepsilon_{it}$$

where $T_{2,t}, \ldots T_{6,t}$ are 6 dummy variables for the 6 years, omitting the first. Test the null hypothesis that the 6 dummy variable coefficients all equal zero and report all results. Intepret your findings.

```
--> NAMELIST ; XPart3 = One,OCC,SMSA,MS,ED,EXP $
--> REGRESS ; Lhs = Lwage ; Rhs = XPart3 $

+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=LWAGE    Mean                 =   6.676346     |
|              Standard deviation   =   .4615122     |
| WTS=none     Number of observs.   =       4165     |
| Model size   Parameters           =          6     |
|              Degrees of freedom   =       4159     |
| Residuals    Sum of squares       =   587.0679     |
|              Standard error of e  =   .3757074     |
| Fit          R-squared            =   .3380713     |
|              Adjusted R-squared   =   .3372755     |
| Model test   F[  5,  4159] (prob) = 424.83 (.0000) |
| Diagnostic   Log likelihood       =  -1829.572     |
|              Restricted(b=0)      =  -2688.806     |
|              Chi-sq [  5]  (prob) =1718.47 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -1.956450     |
|              Akaike Info. Criter. =  -1.956450     |
|              Bayes Info. Criter.  =  -1.947325     |
| Autocorrel   Durbin-Watson Stat.  =   .7555981     |
|              Rho = cor[e,e(-1)]   =   .6222010     |
| Model was estimated Feb 10, 2009 at 06:56:48AM     |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
```

```
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
|Constant|    5.40421***       .04487911   120.417   .0000            |
|OCC     |    -.09787***       .01500945    -6.521   .0000    .5111645|
|SMSA    |     .14884***       .01266551    11.752   .0000    .6537815|
|MS      |     .32660***       .01534942    21.277   .0000    .8144058|
|ED      |     .05882***       .00274699    21.412   .0000   12.845378|
|EXP     |     .01024***       .00055661    18.398   .0000   19.853782|
+--------+------------------------------------------------------------+
| Note: ***, **, * = Significance at 1%, 5%, 10% level.               |
+---------------------------------------------------------------------+
```

```
--> CALC ; SSPool = sumsqdev $
--> REGRESS ; For[FEM=0] ; Lhs = Lwage ; Rhs = XPart3 $
*****************************************************
* Setting up an iteration over the values of FEM      *
* The model command will be executed for   1 values   *
* of this variable.  In the current sample of    4165 *
* observations, the following counts were found:      *
* Subsample   Observations    Subsample  Observations *
* FEM      =   0      3696    FEM     =****  ******** *
*-----------------------------------------------------*
* Actual subsamples may be smaller if missing values  *
* are being bypassed.  Subsamples with 0 observations *
* will be bypassed.                                   *
*****************************************************

**************************************************************************
*       Subsample analyzed for this command is FEM      =        0       *
**************************************************************************
```

```
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=LWAGE    Mean                 =   6.729774     |
|              Standard deviation   =   .4382202     |
| WTS=none     Number of observs.   =       3696     |
| Model size   Parameters           =          6     |
|              Degrees of freedom   =       3690     |
| Residuals    Sum of squares       =   503.8896     |
|              Standard error of e  =   .3695341     |
| Fit          R-squared            =   .2898728     |
|              Adjusted R-squared   =   .2889105     |
| Model test   F[  5,  3690] (prob) = 301.25 (.0000) |
| Diagnostic   Log likelihood       =  -1561.981     |
|              Restricted(b=0)      =  -2194.572     |
|              Chi-sq [  5]  (prob) =1265.18 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -1.989402     |
|              Akaike Info. Criter. =  -1.989402     |
|              Bayes Info. Criter.  =  -1.979313     |
| Autocorrel   Durbin-Watson Stat.  =   .7604481     |
|              Rho = cor[e,e(-1)]   =   .6197759     |
| Model was estimated Feb 10, 2009 at 06:56:48AM     |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
|Constant|    5.63563***       .05029666   112.048   .0000            |
|OCC     |    -.08984***       .01576766    -5.697   .0000    .5262446|
|SMSA    |     .15905***       .01299522    12.239   .0000    .6360931|
|MS      |     .09638***       .02199861     4.381   .0000    .9145022|
|ED      |     .05707***       .00286496    19.920   .0000   12.846591|
|EXP     |     .01083***       .00058000    18.677   .0000   20.214015|
+--------+------------------------------------------------------------+
| Note: ***, **, * = Significance at 1%, 5%, 10% level.               |
+---------------------------------------------------------------------+
```

```
--> CALC ; SSMen = sumsqdev $
--> REGRESS ; For[FEM=1] ; Lhs = Lwage ; Rhs = XPart3 $
*****************************************************
* Setting up an iteration over the values of FEM      *
* The model command will be executed for   1 values   *
* of this variable.  In the current sample of    4165 *
* observations, the following counts were found:      *
* Subsample   Observations    Subsample  Observations *
```

```
* FEM      =   1       469    FEM     =****  ******** *
*----------------------------------------------------*
* Actual subsamples may be smaller if missing values  *
* are being bypassed.  Subsamples with 0 observations *
* will be bypassed.                                   *
*******************************************************

**************************************************************************
*       Subsample analyzed for this command is FEM      =        1       *
**************************************************************************


+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=LWAGE    Mean                =   6.255308      |
|              Standard deviation  =   .4227426      |
| WTS=none     Number of observs.  =        469      |
| Model size   Parameters          =          6      |
|              Degrees of freedom  =        463      |
| Residuals    Sum of squares      =   47.90453      |
|              Standard error of e =   .3216605      |
| Fit          R-squared           =   .4272321      |
|              Adjusted R-squared  =   .4210467      |
| Model test   F[  5,   463] (prob) =  69.07 (.0000) |
| Diagnostic   Log likelihood      =  -130.4956      |
|              Restricted(b=0)     =  -261.1765      |
|              Chi-sq [  5]  (prob) = 261.36 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -2.255805     |
|              Akaike Info. Criter. =  -2.255806     |
|              Bayes Info. Criter.  =  -2.202707     |
| Autocorrel   Durbin-Watson Stat.  =   .8105792     |
|              Rho = cor[e,e(-1)]   =   .5947104     |
| Model was estimated Feb 10, 2009 at 06:56:48AM     |
+----------------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |t-ratio |P[|T|>t]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
|Constant|    5.26320***       .11046622    47.645   .0000            |
|OCC     |    -.25591***       .03915105    -6.536   .0000    .3923241|
|SMSA    |     .12504***       .03921038     3.189   .0015    .7931770|
|MS      |    -.00121          .09661168     -.013   .9900    .0255864|
|ED      |     .06611***       .00754249     8.765   .0000   12.835821|
|EXP     |     .00851***       .00154633     5.503   .0000   17.014925|
+--------+------------------------------------------------------------+
| Note: ***, **, * = Significance at 1%, 5%, 10% level.               |
+---------------------------------------------------------------------+
```

--> CALC ; SSWomen = Sumsqdev $
--> CALC ; K = Col(XPart3) $
--> CALC ; List ; Chow = ((SSPool - (SSMen+SSWomen))/ K) /
((SSMen + SSWomen) / (N - 2*K)) $

```
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 CHOW    =     44.247081
```

--> calc;LIST;FTB(.95,K,(N-2*K))$

```
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 Result  =      2.100770
```

```
CREATE ; T = Trn(-7,0)$
CALC ; SST = 0 $
Procedure
REGRESS ; For[T=j] ; Lhs = Lwage ; Rhs = Xpart3 ; quietly $
CALC ; SST = SST + Sumsqdev $
ENDPROC
Exec ; J = 1,7 $
CALC ; List ; Chow = ((SSPool - SST)/(6*K)) / (SST/(N-7*K)) $
--> CALC ; List ; Chow = ((SSPool - SST)/(6*K)) / (SST/(N-7*K)) $
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 CHOW    =     34.789262
--> CALC ; List ; FTB(.95,(6*K),(n-7*K)) $
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 Result  =      1.419543

CALC ; List ; FTB(.95,(6*K),(n-7*K)) $
CREATE ; T1 = t=1 ; T2 = t=2 ; T3 = t=3 ; T4 = t=4 ; T5 = t=5 ; T6 = T=6 $
REGRESS ; Lhs = Lwage ; Rhs = XPart3 $
CALC ; SS0 = Sumsqdev $
REGRESS ; Lhs = Lwage ; Rhs = XPart3,t1,t2,t3,t4,t5,t6 $
CALC ; List ; FStat = ((SS0 - Sumsqdev)/6) / (sumsqdev/(n-k-6)) $
--> CALC ; List ; FStat = ((SS0 - Sumsqdev)/6) / (sumsqdev/(n-k-6)) $
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 FSTAT   =    203.701855
```

Part IV. A Nonlinear Regression

1. The model (*) above omits a well known phenomenon with respect to the association of wages and experience – earnings often do not increase uniformly with experience, but rather increase more rapidly in the earlier years of employment than in the later years. Test this theory by adding EXP^2 to your model. The equation is

$$(****)\ LWAGE_{it} = \beta_1 + \beta_2 OCC_{it} + \beta_3 SMSA_{it} + \beta_4 MS_{it} + \beta_5 FEM_i + \beta_6 ED_i + \gamma EXP_{it} + \delta EXP_{it}^2 + \varepsilon_{it}$$

Refit the model by least squares and discuss your results. Use the entire sample. Does squared experience help to explain the variation in log wages? Test the null hypothesis that it does not. What do you find?

2. Partial Effect. As part of your analysis, derive and statistically analyze the partial effect of experience,

$$\theta(EXP_{it}) = \partial E[LWAGE|\mathbf{x}]/\partial EXP = \gamma + 2\delta EXP_{it}$$

at the sample mean value of EXP_{it}. Compute an asymptotic standard error for the estimator of θ then test the hypothesis that θ equals zero.

3. Examining the Regression. Obtain the sample mean values of OCC, SMSA, MS ED. Then, using your estimated coefficients, compute

$$\hat{\alpha} = \hat{\beta}_1 + \hat{\beta}_2 \overline{OCC} + \hat{\beta}_3 \overline{SMSA} + \hat{\beta}_4 \overline{MS} + \hat{\beta}_6 \overline{ED}\,.$$

We are interested in what the regression model implies about the trajectory of wages as a function of experience. Thus, we want to plot

$$\widehat{WAGE}(Exp \mid Female) = \exp\left(\hat{\alpha} + \hat{\beta}_5 + \hat{\gamma}Exp + \hat{\delta}Exp^2\right)$$

and

$$\widehat{WAGE}(Exp \mid Male) = \exp\left(\hat{\alpha} + \hat{\gamma}Exp + \hat{\delta}Exp^2\right)$$

What do you find? Interpret the figure.

HINT: Here are NLOGIT commands that you can use to do this computation

```
regress;lhs=lwage;rhs=one,occ,smsa,ms,fem,ed,exp,expsq$
calc;occbar=xbr(occ);smsabar=xbr(smsa);msbar=xbr(ms);edbar=xbr(ed)$
calc;ahat=b(1)+b(2)*occbar+b(3)*smsabar+b(4)*msbar+b(6)*edbar$
samp;1-51$
create;fitexp=trn(1,1)$
create;wagef=exp(ahat + b(5)  + b(7)*fitexp+b(8)*fitexp^2)$
create;wagem=exp(ahat +         b(7)*fitexp+b(8)*fitexp^2)$
plot;lhs=fitexp;rhs=wagef,wagem;fil
;title =Wage Trajectories for Men and Women Based on Experience
;Vaxis =Projected Weekly Earnings
;Grid$
```

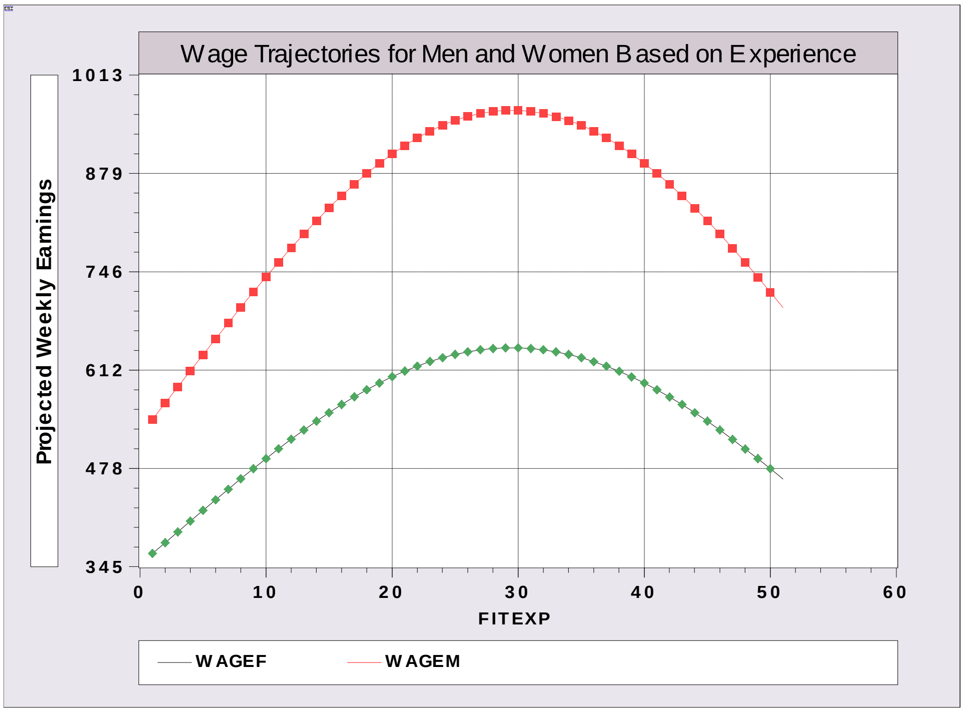
```
+----------------------------------------------------+
| Ordinary    least squares regression               |
| LHS=LWAGE    Mean                 =   6.676346     |
|              Standard deviation   =   .4615122     |
| WTS=none     Number of observs.   =       4165     |
| Model size   Parameters           =          8     |
|              Degrees of freedom   =       4157     |
| Residuals    Sum of squares       =   529.2850     |
|              Standard error of e  =   .3568246     |
| Fit          R-squared            =   .4032224     |
|              Adjusted R-squared   =   .4022174     |
| Model test   F[  7,  4157] (prob) = 401.25 (.0000) |
| Diagnostic   Log likelihood       =  -1613.797     |
|              Restricted(b=0)      =  -2688.806     |
|              Chi-sq [  7]  (prob) =2150.02 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =  -2.059103     |
|              Akaike Info. Criter. =  -2.059103     |
|              Bayes Info. Criter.  =  -2.046936     |
| Autocorrel   Durbin-Watson Stat.  =   .7637078     |
|              Rho = cor[e,e(-1)]   =   .6181461     |
| Model was estimated Feb 10, 2009 at 07:00:17AM     |
+----------------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Constant	5.41799***	.04866099	111.342	.0000	
OCC	-.10684***	.01428962	-7.477	.0000	.5111645
SMSA	.16053***	.01204145	13.331	.0000	.6537815
MS	.07571***	.02085803	3.630	.0003	.8144058
FEM	-.40722***	.02541147	-16.025	.0000	.1126050
ED	.05683***	.00261192	21.758	.0000	12.845378
EXP	.04152***	.00219803	18.888	.0000	19.853782
EXPSQ	-.00070***	.483146D-04	-14.567	.0000	514.40504

Note: nnnnn.D-xx or D+xx => multiply by 10 to -xx or +xx.
Note: ***, **, * = Significance at 1%, 5%, 10% level.

```
--> CALC ; EXpBar = Xbr(EXP) $
--> Calc ; list ; Theta = b(7)+2*b(8)*Expbar $
```

```
+-------------------------------------+
| Listed Calculator Results           |
+-------------------------------------+
 THETA    =       .013570
--> Calc ; list ; stheta = sqr(Varb(7,7) + 4*expbar*expbar*varb(8,8) + 4*Expb...
+-------------------------------------+
| Listed Calculator Results           |
+-------------------------------------+
 STHETA   =       .000571
--> Calc ; list ; ttheta = theta/stheta $
+-------------------------------------+
| Listed Calculator Results           |
+-------------------------------------+
 TTHETA   =     23.772215
```



Part V. Nonlinear Function of Parameters

Based on the regression model in part IV,

(****) $LWAGE_{it} = \beta_1 + \beta_2 OCC_{it} + \beta_3 SMSA_{it} + \beta_4 MS_{it} + \beta_5 FEM_i + \beta_6 ED_i +$
$EXP_{it} + \delta EXP_{it}^2 + \varepsilon_{it}$

In Part IV, you plotted the trajectory of WAGE (not log WAGE) using the mean values of the variables in the model. Note that the figure shows a parabola with a maximum at about 29 years, We are interested in exploring the computation of the peak earning year.

TIP. For purposes of this exercise, you will find it convenient to use

$EXP100 = EXP/100$

and

$EXP100SQ = EXP100^2$

for your regression. This scaling will make the standard errors that you need to use much more convenient but will, of course, not change the model.

(*****) $LWAGE_{it} = \beta_1 + \beta_2 OCC_{it} + \beta_3 SMSA_{it} + \beta_4 MS_{it} + \beta_5 FEM_i + \beta_6 ED_i +$
$\gamma EXP100_{it} + \delta EXP100_{it}^2 + \varepsilon_{it}$

Refit the model, using EXP100 and EXP100SQ. Our prediction of WAGE is

$$\widehat{WAGE}(Exp \mid Female) = \exp\left(\hat{\alpha} + \hat{\beta}_5 + \hat{\gamma}Exp100 + \hat{\delta}Exp100^2\right)$$

1. Prove that the maximum of this function occurs at $EXP100^* = -\hat{\gamma}/2\hat{\delta}$. We are interested in estimating and forming a confidence interval for EXP100*.

2. Compute the value of EXP100* using the results of your regression.

3. Use the delta method to obtain estimated asymptotic standard error for EXP100*.

HINTS: You can use the following NLOGIT commands to do the regression.

```
sample;all$
create;exp100=exp/100 ; exp100sq=exp100^2$
regress;lhs=lwage;rhs=one,occ,smsa,ms,fem,ed,exp100,exp100sq$
```

The rest of the computations can be done with a hand calculator or with the CALC command. After you compute the regression, go into the project window, open the Matrices list, then double click on VARB to show the asymptotic covariance matrix.

```
--> calc ; list ; max = -b(7)/(2*b(8)) $
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 MAX     =       .294943
--> calc ; g1 = -1/(2*b(8)) ; g2 = -max/b(8) $
--> calc ; smax = sqr(g1*g1*varb(7,7)+g2*g2*varb(8,8)+2*g1*g2*varb(7,8))$
--> calc ; list ; max/smax $
+------------------------------------+
| Listed Calculator Results          |
+------------------------------------+
 Result  =     46.619582
--> wald ; start = b ; var = varb ; labels = 8_c ; fn1=-c7/(2*c8)$

+-----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic             =   2173.38527     |
| Prob. from Chi-squared[ 1] =       .00000     |
+-----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
|Fncn(1) |     .29494***       .00632660    46.620   .0000 |
+--------+--------------------------------------------------+
```

4. We will now use the method of Krinsky and Robb as an alternative to the delta method. The method proceeds as follows. We are interested in analyzing a nonlinear function of $\hat{\gamma}$ and $\hat{\delta}$. We have estimated the 2×2 asymptotic covariance matrix for this pair of estimators; call it Σ. Our estimators are asymptotically normally distributed with mean (γ,δ) and asymptotic covariance matrix Σ. What we will do is draw a large random sample from this population, (γ_r,δ_r), r = 1,...,R, then compute from this sample, a sample of values $EXP_r = -\gamma_r/(2\delta_r)$. We will then use the empirical standard deviation from the sample of draws as our estimator of the asymptotic standard deviation of the estimator of EXP.

How to draw a random sample from this population: We need a sample of draws of the form

$$\gamma_r = \hat{\gamma} + w_{\gamma,r} \text{ and } \delta_r = \hat{\delta} + w_{\delta,r} \text{ then } EXP_r^* = -\gamma_r/(2\delta_r)$$

where $(w_{\gamma,r},w_{\delta,r})$ have bivariate normal distribution with mean vector (0,0) and covariance matrix Σ. Here is how to do that. We will use the Cholesky decomposition of Σ. L is a lower triangular matrix such that $LL' = \Sigma$. Let $v_{1,r}$ and $v_{2,r}$ be samples of independent draws from the standard normal distribution. Then,

$$w_{\gamma,r} = L_{11}v_{1,r} \quad \text{and} \quad w_{\delta,r} = L_{21}v_{1,r} + L_{22}v_{2,r}.$$

a. Let σ_{11} = the asymptotic variance of $\hat{\gamma}$, σ_{22} = the asymptotic variance of $\hat{\delta}$ and let σ_{12} = the asymptotic covariance.
Show that $L_{11} = \text{sqr}(\sigma_{11})$, $L_{21} = \sigma_{12}/L_{11}$ and $L_{22} = \text{sqr}(\sigma_{22} - \sigma_{12}^2/\sigma_{11})$.
b. Compute the random sample of draws on v1 and v2.
c. Compute the random sample of draws on γ and δ.
d. Compute the random sample of draws on EXP*
e. Compute the standard error for your estimate of EXP* using the sample standard deviation.
f. Compare your result to the results using the delta method in part 3.

HINT: This set of NLOGIT commands does the computation after the regression.

```
calc   ;sgg=varb(7,7);sgd=varb(8,7);sdd=varb(8,8)$
calc   ;list; L11=sqr(sgg) ; L21=sgd/sqr(sgg) ; L22=sqr(sdd-sgd^2/sgg)$
create ;u1=rnn(0,1);u2=rnn(0,1)$
create ;gr=b(7) + L11*u1 ; dr=b(8) + L21*u1 + L22*u2 $
create ;expr = -gr/(2*dr)$
dstat  ;rhs=expr$
? This command uses the delta method after the regression
wald   ;start=b;var=varb;labels=8_b ; fn1=-b7/(2*b8)$
```

```
--> create ;u1=rnn(0,1);u2=rnn(0,1)$
--> create ;gr=b(7) + L11*u1 ; dr=b(8) + L21*u1 + L22*u2 $
--> create ;expr = -gr/(2*dr)$
--> dstat  ;rhs=expr$
Descriptive Statistics
All results based on nonmissing observations.
==============================================================================
Variable     Mean       Std.Dev.     Minimum      Maximum        Cases Missing
==============================================================================
All observations in current sample
--------+---------------------------------------------------------------------
EXPR    |  .295373      .635183E-02  .277729      .327965         4165       0
--> wald   ;start=b;var=varb;labels=8_b ; fn1=-b7/(2*b8)$

+-----------------------------------------------+
| WALD procedure. Estimates and standard errors |
| for nonlinear functions and joint test of     |
| nonlinear restrictions.                       |
| Wald Statistic             =   2173.38527     |
| Prob. from Chi-squared[ 1] =       .00000     |
+-----------------------------------------------+
+--------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+--------+--------------+----------------+--------+--------+
|Fncn(1) |     .29494***       .00632660    46.620   .0000 |
+--------+------------------------------------------------+
| Note: ***, **, * = Significance at 1%, 5%, 10% level.    |
+---------------------------------------------------------+
```