Graduate Econometrics
Lecture 3: The Linear Regression Model cont.

Yonas Alem

Department of Economics
University of Gothenburg

November 18, 2014

- Consider a statistical wage model

$$wage_i = \beta_1 + \beta_2 male + \varepsilon_i \tag{1}$$

- Where $wage_i$ denotes the hourly wage rate of individual i and $male_i = 1$ if i is male and 0 otherwise
- If $E\{\varepsilon_i\} = 0$ and $E\{\varepsilon_i | male_i\} = 0$
  - $E\{wage_i | female_i\} = \beta_1$
  - $E\{wage_i | male_i\} = \beta_1 + \beta_2$
- β_2 is the expected wage differential between an arbitrary male and female

- Using a subsample of the US national Longitudinal Survey (NLS) 1987 data we estimate the above wage model as follows

Table: 2.1: OLS results wage equation

Variable	Estimate	SE
Constant	5.1469	0.0812
male	1.1661	0.1122

Notes: s=3.2174; $R^2 = 0.0317; F = 107.93$

- $\Longrightarrow$ the wage differential between males and females is about \$1.17 with a SE of \$0.11.
- Is this wage differential real or it happened by chance? To answer this, we should undertake a hypothesis testing that $H_0 : \beta_2 = 0$

The Gauss-Markov Assumptions again

- $[A1] : E\{\varepsilon_i\} = 0, \quad i = 1, ..., N$
- $[A2] : E\{x_i\varepsilon_i\} = 0$
- $[A3] : V\{\varepsilon_i\} = \sigma^2, \quad i = 1, ..., N$
- $[A4] : Cov\{\varepsilon_i, \varepsilon_j\} = 0, \quad i \neq j$
- For exact statistical inference from a given sample, of N observations one needs the additional distributional assumption (normality), i.e.,

$$\varepsilon_i \sim NID(0, \sigma^2) \qquad [A5] \tag{2}$$

- $[A1] - [A5] \Longrightarrow b \sim N(\beta, \sigma^2(X'X)^{-1})$ and $b_k \sim N(\beta_k, \sigma^2 c_{kk})$
- Where c_{kk} is the (k, k) element in $(X'X)^{-1}$

- Thus, the variable:

$$z = \frac{b_k - \beta_k}{\sigma\sqrt{c_{kk}}} \tag{3}$$

- has a standard normal distribution, i.e., $z \sim (0, 1)$
- Replace σ by $s \longrightarrow$ the random variable

$$t_k = \frac{b_k - \beta_k}{s\sqrt{c_{kk}}} \tag{4}$$

- is the ratio of a standard normal variable and the square root of an independent squared variable and therefore follows Student's t distribution with $N - K$ d.f
- The t distribution is close to the standard normal distribution and for large $N - K$, the two distributions are identical

A simple t-Test

- Let $H_0 : \beta_k = \beta_k^0$ where β_k^0 is a specific value chosen by the researcher
- If this is true, the **statistic:**

$$t_k = \frac{b_k - \beta_k^0}{se(b_k)} \tag{5}$$

- Has a t distribution with $N - K$ degrees of freedom. If H_0 is not true, the alternative, $H_1 : \beta_k \neq \beta_k^0$ holds
- One rejects H_0 if the probability of observing a value of $|t_k|$ or larger is smaller than a given **significance level** α, often 5% $\Longrightarrow$ the critical values $t_{N-K;\alpha/2}$:

$$P\{|t_k| > t_{N-K;\alpha/2}\} = \alpha \tag{6}$$

A simple t-Test cont.

- The **two-tailed** critical value for $\alpha = 0.05$ is $1.96 \Longrightarrow$ at the 5% level, H_0 will be rejected if

$$|t_k| > 1.96 \tag{7}$$

- Unlike the above, we may want H_1 to be **one-sided**, e.g., the expected wage for a man is larger than that for a woman. In this case,
  - $H_0 : \beta_k \leq \beta_k^0$ and $H_1 : \beta_k > \beta_k^0$
- The critical value is determined from $P\{|t_k| > t_{N-K;\alpha}\} = \alpha$
- One rejects H_0 at the 5% level if $t_k > 1.64$
- Regression packages report the t-statistic for, $H_0 : \beta_k = 0$

$$t_k = \frac{b_k}{se(b_k)} \tag{8}$$

A simple t-Test - Confidence Interval

- A **confidence interval**: the interval of all values for β_k^0 for which $H_0 : \beta_k = \beta_k^0$ is not rejected by the t-tests.

$$-t_{N-K;\alpha/2} < \frac{b_k - \beta_k}{se(b_k)} < t_{N-K;\alpha/2}, \tag{9}$$

or

$$b_k - t_{N-K;\alpha/2} se(b_k) < \beta_k < b_k + t_{N-K;\alpha/2} se(b_k), \tag{10}$$

- Using the standard normal approximation, a 95% confidence interval (setting $\alpha = 0.05$ for β_k is given by the interval

$$[b_k - 1.96se(b_k), b_k + 1.96se(b_k)] \tag{11}$$

- In repeated sampling, 95% of these intervals will contain the true value β_k
- Test $H_0 : \beta_2 = 0$ for the wage regression reported in table 2.1

Joint Test of Significance: *F-Test*

- Tests whether the increase in R^2 moving from a restricted model to the more general model is significant:

$$F = \frac{(S_0 - S_1)/J}{S_1/(N-K)} \tag{12}$$

where

- S_1 is the residual sum of squares of the model, i.e., $S_1 = \sum_i e_i^2$ and
- S_0 is the residual sum of squares of the restricted model
- Under H_0 : (which tests that J of the K coefficients are equal to zero) F has and F distribution with J (the number of regressors omitted) and $N-K$ d.f, denoted as F^J_{N-K}

Joint Test of Significance: *F-Test* cont.

- In terms of R^2:

$$F = \frac{(R_1^2 - R_0^2/J}{(1 - R_1^2)/(N - K)} \tag{13}$$

- Where R_1^2 and R_0^2 refer to those of the unrestricted and the restricted model respectively
- A special case of this $F - test$ is the : standard test automatically supplied by regression a package
- Tests the null hypothesis that the partial slope coefficients are equal to zero, i.e, tests $H_0 : \beta_2 = \beta_3 = ... = \beta_K = 0$. The test statistic is given by

$$F = \frac{(S_0 - S_1)/(K - 1)}{S_1/(N - K)} \tag{14}$$

Joint Test of Significance: *F-Test* cont.

- By construction, R^2 of $S_0 = 0 \Longrightarrow$

$$F = \frac{R^2/(K-1)}{(1-R^2)/(N-K)} \tag{15}$$

- Has an F distribution with $(K-1)\&(N-K)$ degrees of freedom, denoted as F^{K-1}_{N-K}
- Given the wage equation

$$wage_i = \beta_1 + \beta_2 male_i + \beta_3 school_i + \beta_4 exper_i + \varepsilon_i \tag{16}$$

  - Test $H_0 : \beta_k = 0$ for each explanatory variable
  - Test $H_0 : \beta_2 = \beta_3 = ... = \beta_K = 0$, and
  - Test $H_0 : \beta_3 = \beta_4 = 0$

Joint Test of Significance: *F-Test* cont.

Table: Estimation results : regress

Variable	**Coefficient**	**(Std. Err.)**
male	1.344	(0.108)
school	0.639	(0.033)
exper	0.125	(0.024)
Intercept	-3.380	(0.465)
N	3294	
R^2	0.133	
$F_{(3,3290)}$	167.63	

Size, Power and p-Values

- Two types of errors can be made while testing a hypothesis
  - **Type I error:** One rejects H_0, while it is actually true (can be controlled by choice of α
  - **Type II error:** H_0 is not rejected while H_1 is true (its probability depends on the true parameter values)
- The probability of rejecting H_0 when it is not true is known as the **power** of the test
- **p-value**: the marginal significance level for which H_0 would still be rejected
  - If $p - value < \alpha$, H_0 is rejected
  - Reported with the regression table in many softwares

- Under the Gauss-Markov conditions, $E(\hat{\beta}_{OLS}) = \beta$
- What if assumption [A2] is violated? In this case, $E(\hat{\beta}_{OLS}) \neq \beta \Longrightarrow \hat{\beta}_{OLS}$ would be biased
- **Asymptotic theory:** what happens if, hypothetically, the sample size grows infinitely large?
- **Consistency:** under the Gauss-Markov conditions:

$$lim_{N \longrightarrow \infty} P\{|b_k - \beta_k| > \delta\} = 0 \; \forall \delta > 0. \tag{17}$$

- Asymptotically, the probability that the OLS estimator deviates more than δ from the true parameter value is zero. Or

$$plim \;\; b = \beta \tag{18}$$

- It is **consistent!**
- But if A2 is not true, this will not be true

- High correlation between explanatory variables may be problematic
- Example: age and experience in the wage equation
- Meaning: the matrix $(X'X)$ would be close to being not invertible
- It would be difficult to identify the effects of these variables
- Possible results.
  - Unreliable estimates (high standard errors)
  - Unexpected sign and magnitude
- **Multicollinearity:** when an approximate linear relationship among the explanatory variables leads to unreliable regression estimates

- Consider the the variance of a single coefficient β_k

$$V\{b_k\} = \frac{\sigma^2}{1 - R_k^2} \frac{1}{N} [\frac{1}{N} \sum_{i=1}^{N} (x_{ik} - \bar{x}_k)^2]^{-1}, \quad k = 2, ..., K, \quad (19)$$

where

- R_k^2 denotes the squared multiple correlation coefficient between x_{ik} and the other explanatory variables
  - I.e., the R^2 from regressing x_{ik} upon the remaining regressors and a constant
- If $R_k^2 \longrightarrow 1, V\{b_k\}$ would be large
- If there is enough variation in x_{ik}, N is large, and σ^2 is sufficiently small, the large value of R_k^2 need not cause a problem

The VIF test

- The **variance inflation factor (VIF)** is sometimes used to detect multicollinearity

$$VIF(b_k) = \frac{1}{1 - R_k^2} \tag{20}$$

- Indicates the factor by which $V\{b_k\}$ is inflated compared with the hypothetical situation when there is no correlation between x_{ik} and any other explanatory variables
- As a rule of thumb, $VIF \geq 10$ is "too high"
- The test is warranted if we suspect multicollinearity
- Ignore multicollinearity if the focus variable is not correlated with other variables
- In the case of **exact multicollinearity**, drop one of the variables

The VIF test cont.

- Check what would happen if we estimate the following wage equation

$$wage_i = \beta_1 + \beta_2 male + \beta_3 female + \varepsilon_i \tag{21}$$

- An **outlier**: an observation that deviates markedly from the rest of the sample, for e.g., due to
  - Measurement error in the data (correct or discard)
  - Chance (less obvious what to do)
- Good to check summary statistics of variables before estimation
- Use robust estimation methods, e.g., the least square absolute deviation (LAD) model
- Estimates the conditional median (rather than the conditional mean)
  - No closed-form solution to minimizing the LAD deviations
- **Winsorize** the data (adjusting the tails of the distribution of each variable)
- The use of **trimmed least squares** (omitting the most extreme 5% observations)

- This is a frequently encountered problem in empirical work, especially with micro-economic data
- Software packages skip missing data in estimation
- Not a problem if the missing data is random
- Is a problem if missing data is not random! may lead to **sample selection bias**
- For e.g., if we only observe wages above a certain threshold and have missing values otherwise, the OLS estimator in the wage will suffer from selection bias
- Replace missing data by some number, e.g., zero, sample average, or use a dummy variable for the missing data
  - It may however bias results
- **Imputation:** replacing the data by sample mean, randomly generated values etc.. could also be problematic

- Given the linear model

$$y_i = x_i'\beta + \varepsilon_i \qquad and \qquad E\{\varepsilon_i|x_i\} = 0 \tag{22}$$

- The expected change in y_i if x_{ik} changes with one unite but all other variables in x_i do not change (**ceteris paribus**) is given by

$$\frac{\partial E\{y_i|x_i\}}{\partial x_{ik}} = \beta_k \tag{23}$$

- What would be the marginal effect of age in the following wage equation?

$$wage_i = \beta_1 + \beta_2 age_i + \beta_3 age^2 + \varepsilon_i \tag{24}$$

in this case,

$$\frac{\partial E\{y_i|x_i\}}{\partial age_i} = \beta_2 + 2\beta_3 age_i \tag{25}$$

- Checking the varying effect of age on wage based on gender

$$wage_i = \beta_1 + \beta_2 age_i + \beta_3 age_i * male_i + \varepsilon_i \tag{26}$$

$$\frac{\partial E\{y_i|x_i\}}{\partial age_i} = \beta_2 + \beta_3 male_i \tag{27}$$

- β_2 for females and $\beta_2 + \beta_3$ for males
- Note however that the conditional expectation does not necessarily imply that the β parameter always imply the causal effect of x_i on y_i

- An **elasticity** measures the relative change in the dependent due to a relative change in one of the x_i variables
- Can be estimated directly using (a **log linear model**) log values in both sides of the regression (excluding dummy variables)

$$log\ y_i = (log\ x_i)'\gamma + v_i \tag{28}$$

- Where $log\ x_i$ represents the vector with elements $(1, log\ x_{i2}, ..., log\ x_{iK})'$ and assuming that $E\{v_i|log\ x_i\} = 0$

$$\frac{\partial E\{y_i|x_i\}}{\partial x_{ik}} * \frac{x_{ik}}{E\{y_i|x_i\}} \approx \frac{\partial E\{log\ y_i|log\ x_i\}}{\partial log\ x_{ik}} = \gamma_k \tag{29}$$

- The elasticity of the linear model $(y_i = x_i'\beta + \varepsilon_i)$ however is given by

$$\frac{\partial E\{y_i|x_i\}}{\partial x_{ik}} * \frac{x_{ik}}{E\{y_i|x_i\}} = \frac{x_{ik}}{x_i'\beta}\beta_k \tag{30}$$

- Elasticities in this model vary with x_i (are not constant)
- It is sometimes useful to log variables to reduce the problem of heteroskedasticity
- If x_{ik} is a dummy variable (or another variable which will not take -ve values), we can not take its logarithm and we include the original variable in the model
- Interpret β_k as the relative change in y_i owing to an absolute change of one unit in x_{ik}. It is referred to as a **semi-elasticity**
- If the coefficient of the dummy variable for male is $\beta_k = 0.1$, how would it be interpreted?

Misspecification

- Consider the following two models:

$$y_i = x_i'\beta + z_i'\gamma + \varepsilon_i \tag{31}$$

$$y_i = x_i'\beta + v_i \tag{32}$$

- Both describe $E(y_i|x_i, z_i)$ and the second model is *nested* in the first model or it assumes that $\gamma_i = 0 \Longrightarrow z_i$ is irrelevant
- What is the consequence of estimating model 2 while model 1 is the correct model?
- The OLS estimator for β based on the second model denoted as b_2, is given by

$$b_2 = (\sum_{i=1}^{N} x_i x_i')^{-1} \sum_{i=1}^{N} x_i y_i \tag{33}$$

Misspecification cont.

- Substitute the first model in the formula for b_2

$$b_2 = \beta + (\sum_{i=1}^{N} x_i x_i')^{-1} \sum_{i=1}^{N} x_i z_i' \gamma + (\sum_{i=1}^{N} x_i x_i')^{-1} \sum_{i=1}^{N} x_i \varepsilon_i \qquad (34)$$

- The expected value of the last term would be zero
- The second term on the right-hand side, however, corresponds to a bias (asymptotic bias) in the OLS estimator because of estimating the incorrect model
  - This is what is called **omitted variable bias**
- But there would not be any bias if:
  1. $\gamma = 0$
  2. x_i and z_i are **orthogonal**, i.e., $E\{x_i z_i'\} = 0$

Misspecification cont.

- What if we estimate the first model while the actual model is the second?
  - This implies inclusion of irrelevant variables
- The β in the model with irrelevant variables will have higher variance
  - It would be less reliable
- Including as many variables as possible is therefore not a good strategy!

Selecting Regressors

- The set of regressors must be chosen based on economic theory
  - Human capital theory is the basis for the wage equation
  - Consumer theory $\Longrightarrow$ demand functions
- Selecting regressors based on a sequence of tests of significance is not a good practice
  - This practice is called data **snooping** or **data mining**
- May be difficult to avoid the practice in applied work but it is important to be aware of it
- Some practices of selecting regressions:
  - Specific to general specification search - bad practice
  - General-to-specific modeling (The LSE methodology): would result in the correct specification in the long-run
- Researchers in general start somewhere "in the middle"

Selecting Regressors cont.

- In presenting regression results, it is not good to hide insignificant variables
  - Should be reported and discussed
- It is recommended to keep an intercept term in the regression model even if it is insignificant
- There is a trade-off between goodness of fit and simplicity of the model
- Formal tests to check to check this trade-off: Akaike's Information Criterion (AIC), Schwarz Bayesian Information Criterion (BIC), and the F-test

Selecting Regressors cont.

$$AIC = log\frac{1}{N}\sum_{i=1}^{N} e_i^2 + \frac{2K}{N} \tag{35}$$

$$BIC = log\frac{1}{N}\sum_{i=1}^{N} e_i^2 + \frac{K}{N}logN \tag{36}$$

- Models with a lower *AIC* or *BIC* are typically preferred
- Both criteria add a penalty that increases with the number of regressors
- The penalty is larger for *BIC*, it tends to favour more parsimonious models than *AIC*
- These two criteria are used mostly when
  1. Alternative models are not nested
  2. Economic theory provides no guidance on selecting the right model

Nonlinearities

- The interpretation that $E\{y_i|x_i\} = x_i'\beta \Longrightarrow$ no other functions of x_i are relevant in explaining the expected value of y_i
- Nonlinearity
  - The model may be non-linear in its explanatory variables (e.g., age^2 in the wage equation): the model would still be linear in the parameters and can be estimated by OLS
  - The model is non-linear in its parameters $\Longrightarrow$ $E\{y_i|x_i\} = g(x_i, \beta)$ where $g(.)$ is a regression function nonlinear in β e.gs.,

$$g(x_i, \beta) = \beta_1 + \beta_2 x_i^{\beta_3} \tag{37}$$

$$g(x_i, \beta) = \beta_1 x_{i1}^{\beta_2} x_{i2}^{\beta_3} \tag{38}$$

Nonlinearities

- The latter corresponds to a Cobb-Douglas production function with two inputs
- Taking logs (assuming $\beta_1 > 0$) would give a model which is linear in the parameters
- The first case could be estimated using a **nonlinear least squares** estimation which minimizes the following objective function

$$S(\tilde{\beta}) = \sum_{i=1}^{N} (y_i - g(x_i, \tilde{\beta}))^2 \tag{39}$$

- It is not however possible to solve for $\tilde{\beta}$ that minimizes $(\tilde{\beta})$ analytically
  - A numerical method should be used to obtain the non-linear least square estimator

Testing the functional form

- For the linear model given by

$$E\{y_i|x_i\} = x_i'\beta \tag{40}$$

- One could test the functional form by checking whether additional nonlinear terms in x_i are significant using **t-tests**, **F-tests** or more generally, **Wald tests**
- Rambsey (1969) proposed a test based upon the diea that, under H_0, nonlinear functions of $\hat{y}_i = x_i'b$ should not help in explaining y_i
- Consider the following auxiliary regression

$$y_i = x_i'\beta + \alpha_2\hat{y}_i^2 + \alpha_3\hat{y}_i^3 + ... + \alpha_Q\hat{y}_i^Q + v_i \tag{41}$$

Testing the functional form cont.

- The method tests whether the powers of $\hat{y}_i$ have nonzero coefficients in the auxiliary regression
- One could use a standard $F-test$ for $Q-1$ restrictions in $H_0 : \alpha_2 = ... = \alpha_Q = 0$ or a more general Wald test (with asymptotic χ^2 distribution with $Q-1$ d.f)
- Called Regression Equation Specification Error Tests **(RESET tests)**
- The test is usually performed for $Q = 2$ only

Testing a structural break

- A **structural break** a difference in regression coefficients for two or more sub-samples
  - Difference in coefficients between male and female or married and unmarried workers
- Consider an alternative specification consisting of two groups, indicated by $g_i = 0$ and $g_i = 1$ respectively
- The general specification could be given by.

$$y_i = x_i'\beta + g_i x_i'\gamma + \varepsilon_i \tag{42}$$

- Where the K-dimensional vector $g_i x_i$ contains all explanatory variables interacted with the indicator variable g_i

Testing a structural break

- The equation says that the coefficient vector for group 0 is β, whereas for group 1 it is $\beta + \gamma$, so $H_0 : \gamma = 0$ and the model reduces to the restricted version
- $H_0 : \gamma = 0$ can be tested by the F-test we saw earlier and called the **Chow test**

$$F = \frac{(S_R - S_{UR})/K}{(S_{UR})/(N - 2K)}) \tag{43}$$

- Where K is the number of regressors in the restricted model (including the intercept) and $S_R \& S_{UR}$ denote the residual sum of squares of the restricted and unrestricted models respectively

Explaining Individual Wages

- Data: a part of the European Community Household Panel
- Consists of 1472 randomly chosen individuals from Belgium surveyed in 1994
- Check summary statistics
- Interpret regression results

Explaining Individual Wages

```
. sum  wage educ exper
```

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	1472	11.05062	4.450513	2.190978	47.57552
educ	1472	3.378397	1.204522	1	5
exper	1472	17.21739	10.16667	0	47

```
. reg  wage male educ exper
```

Source	SS	df	MS		
Model	10651.6554	3	3550.55181	Number of obs =	1472
Residual	18484.5373	1468	12.5916467	F(3, 1468) =	281.98
				Prob > F =	0.0000
				R-squared =	0.3656
				Adj R-squared =	0.3643
Total	29136.1928	1471	19.8070651	Root MSE =	3.5485

wage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
male	1.346144	.1927364	6.98	0.000	.9680761	1.724212
educ	1.98609	.0806396	24.63	0.000	1.827909	2.144271
exper	.1922751	.0095831	20.06	0.000	.1734771	.2110731
_cons	.2136922	.386895	0.55	0.581	-.5452338	.9726183

Explaining Individual Wages cont

```
. reg  wage  male educ exper expersq
```

Source	SS	df	MS
Model	11023.4381	4	2755.85953
Residual	18112.7546	1467	12.3467993
Total	29136.1928	1471	19.8070651

Number of obs	= 1472
F(4, 1467)	= 223.20
Prob > F	= 0.0000
R-squared	= 0.3783
Adj R-squared	= 0.3766
Root MSE	= 3.5138

wage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
male	1.333693	.1908668	6.99	0.000	.9592925	1.708094
educ	1.988127	.0798526	24.90	0.000	1.831489	2.144764
exper	.3579993	.0316566	11.31	0.000	.2959024	.4200963
expersq	-.0043692	.0007962	-5.49	0.000	-.005931	-.0028073
_cons	-.8924851	.4329127	-2.06	0.039	-1.741679	-.0432912

Explaining Individual Wages cont

- Experience has a non-linear effect on wages (the effect increases, reaches max. and then starts to decline)
- The marginal impact of one more year of experience: $0.358 - 0.0044 \times 2 \times exper_i$
- Estimated wage difference between a persons with 31 and 30 years experience: $0.358 - 0.0044(31^2 - 30^2) = 0.091$
- If the data is heteroskedastic $\Longrightarrow$ the statistical tests may not be valid
- We will see in detail on how to detect test and address heteroskedasticity in the next lecture
- One solution: change the functional form and use a log linear model
- The interpretations change!

Explaining Individual Wages cont

```
. reg  lnwage male lneduc lnexper lnexpersq
```

Source	SS	df	MS		
Model	72.5485883	4	18.1371471	Number of obs	= 1467
Residual	119.903447	1462	.082013301	F(4, 1462)	= 221.15
				Prob > F	= 0.0000
Total	192.452035	1466	.131276968	R-squared	= 0.3770
				Adj R-squared	= 0.3753
				Root MSE	= .28638

lnwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
male	.1175712	.0156158	7.53	0.000	.0869393	.148203
lneduc	.4420035	.0182017	24.28	0.000	.4062993	.4777076
lnexper	.4823204	.1522105	3.17	0.002	.1837461	.7808948
lnexpersq	-.1072235	.0661947	-1.62	0.105	-.2370701	.0226232
_cons	1.016871	.0794516	12.80	0.000	.8610193	1.172722

Explaining Individual Wages cont

- Does including experience and its square improve the model?
- Perform an F-test using the R^2 of unrestricted and restricted models

$$logw_i = \beta_1 + \beta_2 male_i + \beta_3 logeduc_i + \beta_4 logexper_i + \beta_5 logexper_i^2 + \varepsilon_i \tag{44}$$

$$logw_i = \beta_1 + \beta_2 male_i + \beta_3 logeduc_i + \varepsilon_i \tag{45}$$

$$F = \frac{(0.3770 - 0.1787)/2}{(1 - 0.3770)/(1472 - 5)} = 233.7 \tag{46}$$

- $\Longrightarrow$ strong rejection of $H_0 : \beta_4 = \beta_5 = 0$!