ECONOMETRICS FOR PRACTITIONERS
SERIES I: BASIC CONCEPTS

NOR AZAM ABDUL RAZAK
School of Economics, Finance & Banking
UUM College of Business
Universiti Utara Malaysia

OUTLINE OF WORKSHOP

1. **Introduction**
2. **Multiple Linear Regression Model (MLRM)**
3. **Interpretation of Estimated Coefficients**

References

1. D. Gujarati (2009), Basic Econometrics, 3rd edition
2. J. Wooldridge (2009), Introductory Econometrics, 3rd edition
3. M. Verbeek (2008), A Guide to Econometrics, 2nd edition

UUM COB, Sintok
March 6th – 7th, 2013

1. Introduction

A starting point in the study of econometrics is the classical linear regression model (CLRM).

The term "regression model" refers to a functional relationship between a dependent variable and explanatory variable(s).

The term "linear" refers to a regression model that is linear in parameters (not necessarily linear in explanatory variables).

The term "classical" refers to a set of basic assumptions that must hold in order for ordinary least squares (OLS) estimator to be considered as the best estimator in a regression model.

If one or more of these classical assumptions are violated or relaxed, OLS may no longer be the best (i.e. other estimators may prove to be superior to OLS).

For this reason, it's important for us to determine whether the classical assumptions hold for any regression models that we want to estimate.

We'll start with a multiple linear regression model, where the term "multiple" refers to the number of explanatory variables.

2. Multiple Linear Regression Model

A multiple linear regression model can be written as follows,

(1) $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \ldots + \beta_K x_{Ki} + \varepsilon_i$,

where

i = index of observations (i = 1, 2, …, N);
x_{kit} = k^{th} explanatory variable (k = 2, …, K);
β_1 = intercept term
β_k = coefficient of x_{ki}; and
ε_i = error term (or disturbance term).

Note: i could denote individuals or time; usually t denotes a time index.

Given the model in Eq.(1), the assumptions of the CLRM are as follows:

A1) The regression model is linear in parameters, β's.

This assumption allows us to obtain the analytical solution of OLS estimator.

A2) The regression model is correctly specified (no specification error).

This assumption means that all of the relevant variables have been included and the functional form used is correct.

A3) The expected value of each error term conditional on **x** is zero,

(3) $E(\varepsilon_i \mid \mathbf{x}) = 0$ for all i's

⇒ Each explanatory variable, x_i, is not correlated with ε.

⇒ The expected value of y is not affected by ε.

⇒The marginal effect of x_i on y is not affected by ε.

A4) The variance of each error term is constant,

(4) $Var(\varepsilon_i) = \sigma^2$ for all i's

That is, the individual values of y are spread around their mean value, $E(y_i|x_i)$, with the same variance.

If this assumption is violated, then we'll have heteroskedasticity problem.

A5) The error terms are not correlated with each other,

(5) $Cov(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$

That is, there's no systematic relationship between any two error terms.

If this assumption is violated, then we'll have autocorrelation problem.

A6) There's no exact/perfect collinearity between/among x's.

Consider the regression model,

(6) $E(y_i|x_{2i}, x_{3i}) = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i}$

No exact collinearity means that x_2 cannot be expressed as an exact linear function of x_3 (or vice versa).

E.g. If $x_{2i} = 3 + 2x_{3i}$ or $x_{2i} = 4x_{3i}$, then the two variables are collinear.

Suppose $x_{2i} = 4x_{3i}$. Then, substituting this into Eq.(6) yields

$$E(y_i|x_{2i}, x_{3i}) = \beta_1 + \beta_2(4x_{3i}) + \beta_3 x_{3i}$$
$$= \beta_1 + \underbrace{(4\beta_2 + \beta_3)}_{A} x_{3i} = \beta_1 + Ax_{3i} \quad \text{--------} (7)$$

Eq.(7) is a simple (not multiple) linear regression model. If we estimate Eq.(7) and obtain an estimate for A, there's no way we can get estimates for β_2 and β_3.

Hence, we cannot assess the individual effect of x_2 and x_3 on the conditional mean of y.

In practice, we seldom encounter the case of exact/perfect collinearity. However, there're numerous cases of high yet imperfect collinearity.

A7) The error term follows the normal distribution with zero mean and constant variance,

$$(8) \quad \varepsilon_i \sim N(0, \sigma^2)$$

It's important to note that assumption (A7) is made for the purpose of hypothesis testing, not for the derivation of OLS estimators.

Q: Why bother with these classical assumptions?

The answer to this question is given by the so-called Gauss-Markov theorem:

Given the classical assumptions (A1) – (A6), it can be shown that the OLS estimator for β_k (where k = 1, 2, …, K) is the best, linear unbiased estimator (or BLUE).

The term "linear" means that the OLS estimator is a linear function of Y.

The term "unbiased" means that, in repeated sampling, the OLS estimator is identical to the true population parameter.

$$(9) \quad E(\hat{\beta}_j) = \beta_j \quad for\ j = 0,1,2,\ldots,k$$

The term "best" means that the OLS estimator has the minimum variance among the class of linear, unbiased estimators.

$$(10) \quad Var(\hat{\beta}_j) < Var(\tilde{\beta}_j) \quad for\ j = 0,1,2,\ldots,k$$

If we add assumption (A7), then it can be shown that the OLS estimator is also normally distributed with certain mean and variance.

$$(11) \quad \hat{\beta}_j \sim N(.) \quad for\ j = 0,1,2,\ldots,k$$

Accordingly, hypothesis testing based on normal distribution can be readily applied.

3. Interpretation of Estimated Coefficients

a) The Standard Linear Model

In the standard linear model, each of the parameters is interpreted as the marginal effect of a given explanatory variable, x_{ij}, on the expected value of y_i.

To illustrate, consider the model:

$$(12) \quad y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_K x_{iK} + \varepsilon_i$$

If we assume that x_i's are exogenous (i.e. $E(\varepsilon_i|\mathbf{x}_i)=0$), then the expected value of y_i conditional on $\mathbf{x}_i$ is given by a linear combination of x_i's only:

$$(13)\quad E(y_i|\mathbf{x}_i)=\mathbf{x}_i'\boldsymbol{\beta}=\beta_1+\beta_2 x_{i2}+\ldots+\beta_K x_{iK}$$

Given Eq.(13), the impact of a change in a given explanatory variable, x_{ij}, on the expected value of y_i is given by its corresponding coefficient:

$$(14)\quad \frac{\partial E(y_i|\mathbf{x}_i)}{\partial x_{ij}}=\beta_j$$

Loosely speaking, the partial derivative in Eq.(14) can be interpreted as a change in the expected (or average) value of y_i if x_{ij} changes by *one unit*, holding other variables constant.[1]

In economics, this partial derivative is more commonly referred to as the *marginal effect* of x_{ij} on the expected value of y_i, holding other variables constant.

It's important to realize that the phrase *holding other variables constant*, more popularly known as *ceteris paribus*, implies the *net* marginal effect of x_{ij} on y_i.

Sometimes, this ceteris paribus assumption is hard to maintain as two (or more) explanatory variables tend to move together, either positively or negatively.

To illustrate, consider the wage regression model again:

$$(15)\quad y_i=\beta_1+\beta_2 x_{i2}+\beta_3 x_{i3}+\ldots+\varepsilon_i$$

If x_2 = age and x_3 = experience, then the ceteris paribus assumption is unlikely to hold as older people tend to have more experience.

Consequently, it's not possible to assess the *net* marginal effect of x_{i2} on y_i, and vice versa.

Sometimes, this ceteris paribus assumption is deliberately violated as the same explanatory variable appears in various powers.

To illustrate, consider the wage regression model again:

$$(16)\quad y_i=\beta_1+\beta_2 x_{i2}+\beta_3 x_{i2}^2+\ldots+\varepsilon_i$$

Again, if x_2 = age, then x_2^2 = age-squared. It's obvious that the ceteris paribus assumption is violated as changes in age is automatically accompanied by changes in age-squared.

Fortunately, the violation of this assumption is beneficial because it allows us to determine whether the marginal effect of age on wage is nonlinear.

---

[1] Strictly speaking, the partial derivative in Eq.(3) is interpreted as a change in the expected (or average) value of y_i if x_{ij} changes by *an infinitesimal amount*, holding other variables constant. Although the term *an infinitesimal amount* is more accurate, we'll make use of the term *one unit* in this course, for simplicity.

Formally, the marginal effect of x_{i2} on y_i is given by

$$(17)\quad \frac{\partial E(y_i|x_i)}{\partial x_{i2}} = \beta_2 + 2\beta_3 x_{i2}$$

Eq.(17) indicates that the marginal effect is a function of the value of age; unless this value is specified, we can't calculate the marginal effect (ME).

Given the value of age, the ME of x_{i2} on y_i may be positive or negative. If the ME is positive/negative, then wage is an increasing/decreasing function of age.

Now, if we differentiate this ME, we'll obtain

$$(18)\quad \frac{\partial^2 E(y_i|x_i)}{\partial x_{i2}^2} = 2\beta_3$$

Eq.(18) indicates the rate of change in the ME. If ME > 0 and $\beta_3 > 0$, then wage increases at an increasing rate with respect to age (i.e. as age rises, wage rises at an increasing rate).

If ME > 0 and $\beta_3 < 0$, then wage increases at a decreasing rate with respect to age (i.e. as age rises, wage rises at a decreasing rate). This is more common.

<u>Note</u>: If the estimate of β_3 is insignificant, then its sign conveys little meaning. In this case, if the estimate of β_2 is significant, then we conclude that the ME of age on wage is linear.

b) Alternative Functional Forms

Sometimes we're not interested in finding the marginal effect of x_{ij} on the expected value of y_i (i.e. absolute Δy_i due to a given absolute Δx_{ij}).

Instead, we're interested in finding the elasticity of y_i with respect to x_{ij} (i.e. relative Δy_i due to a given relative Δx_{ij}).

$$(19)\quad E = \frac{\Delta y/y}{\Delta x/x} = \frac{\Delta y}{y}.\frac{x}{\Delta x} = \frac{\Delta y}{\Delta x}.\frac{x}{y}$$

<u>Note</u>: $\Delta y/\Delta x$ is essentially the marginal effect of x on y; hence, the concepts of elasticity and marginal effect are closely related.

It turns out that this measure of elasticity can be obtained by specifying a log-linear model (i.e. a model where both y_i and x_i appear in logarithms):

$$(20)\quad \ln y_i = (\ln \mathbf{x}_i)'\beta + \varepsilon_i = \beta_1 + \beta_2 \ln x_{i2} + \ldots + \varepsilon_i$$

In this model, the coefficient of the log of a given explanatory variable can be shown to be a measure of elasticity:

(21) $\beta_j = \frac{\partial y_i}{\partial x_{ij}} \cdot \frac{x_{ij}}{y_i}$

To verify this, let's define $y_i^* \equiv \ln y_i$ *and* $x_{ij}^* \equiv \ln x_{ij}$. Then, Eq.(20) can be rewritten as

(22) $y_i^* = (x_i^*)' \beta + \varepsilon_i = \beta_1 + \beta_2 x_{i2}^* + \ldots + \varepsilon_i$

Hence, β_1 can be derived by partially differentiating y_i^* with respect to x_{ij}^*:

(23) $\frac{\partial y_i^*}{\partial x_{ij}^*} = \beta_j$

Since $y_i^* \equiv \ln y_i$ *and* $x_{ij}^* \equiv \ln x_{ij}$, then $\partial y_i^* / \partial x_{ij}^*$ can be solved by the chain rule to be

(24) $\frac{\partial y_i^*}{\partial x_{ij}^*} = \frac{\partial y_i^*}{\partial y_i} \cdot \frac{\partial y_i}{\partial x_{ij}} \cdot \frac{\partial x_{ij}}{\partial x_{ij}^*}$

Since $y_i^* \equiv \ln y_i$, *then* $\partial y_i^* / \partial y_i = 1/y_i$. Also, since $x_{ij}^* \equiv \ln x_{ij}$, *then* $\partial x_{ij}^* / \partial x_{ij} = 1/x_{ij}$. Hence, Eq.(24) becomes

(25) $\frac{\partial y_i^*}{\partial x_{ij}^*} = \frac{1}{y_i} \cdot \frac{\partial y_i}{\partial x_{ij}} . x_{ij} = \frac{\partial y_i}{\partial x_{ij}} \cdot \frac{x_{ij}}{y_i}$

Eq.(25) confirms that β_1 is a measure of elasticity.

Ex/ Consider the log-linear version of the wage regression model,

(26) $\ln y_i = \beta_1 + \beta_2 \ln x_{2i} + \beta_3 \ln x_{3i} + \varepsilon_i$

where y is the hourly wage rate, x_2 is education (measured by years of schooling), and x_3 is experience (measured by years of working experience).

Using the same data for a sample of N = 3294 individuals in the U.S., we estimate Eq.(26) by OLS and obtain the following results:

(26)' $\hat{\ln} y_i = \underset{(-10.18))}{-1.905} + \underset{(17.31)}{1.172} \ln x_{2i} + \underset{(9.51)}{0.309} \ln x_{3i}$

Q: How do we interpret the results?

1.172 ⇒ If education increases by 1%, then wage rate is expected to increase by about 1.17%, ceteris paribus.

0.309 ⇒ If experience increases by 1%, then wage rate is expected to increase by about 0.31%, ceteris paribus.

Sometimes we're interested in finding the quasi-elasticity of y_i with respect to x_{ij} (i.e. relative Δy_i due to a given absolute Δx_{ij}).

$$(27)\quad E_Q = \frac{\Delta y / y}{\Delta x} = \frac{\Delta y}{y}.\frac{1}{\Delta x} = \frac{\Delta y}{\Delta x}.\frac{1}{y}$$

This measure of quasi-elasticity can be obtained by specifying a log-lin model (i.e. a model where y_i appears in logs but x_i appears in levels):

$$(28)\quad \ln y_i = x_i^{'}\beta + \varepsilon_i = \beta_1 + \beta_2 x_{i2} + \ldots + \varepsilon_i$$

In this model, the coefficient of a given explanatory variable can be shown to be a measure of quasi-elasticity:

$$(29)\quad \beta_j = \frac{\partial y_i}{\partial x_{ij}}.\frac{1}{y_i}$$

To verify this, let's define $y_i^* \equiv \ln y_i$. Then, Eq.(28) can be rewritten as

$$(30)\quad y_i^* = x_i^{'}\beta + \varepsilon_i = \beta_1 + \beta_2 x_{i2} + \ldots + \varepsilon_i$$

Hence, β_j can be derived by partially differentiating y_i^* with respect to x_{ij}:

$$(31)\quad \frac{\partial y_i^*}{\partial x_{ij}} = \beta_j$$

Since $y_i^* \equiv \ln y_i$, then $\partial y_i^* / \partial x_{ij}$ can be solved by the chain rule to be

$$(32)\quad \frac{\partial y_i^*}{\partial x_{ij}} = \frac{\partial y_i^*}{\partial y_i}.\frac{\partial y_i}{\partial x_{ij}}$$

Since $y_i^* \equiv \ln y_i$, *then* $\partial y_i^* / \partial y_i = 1/y_i$. Hence, Eq.(32) becomes

$$(33)\quad \frac{\partial y_i^*}{\partial x_{ij}} = \frac{1}{y_i}.\frac{\partial y_i}{\partial x_{ij}} = \frac{\partial y_i}{\partial x_{ij}}.\frac{1}{y_i}$$

Eq.(33) confirms that β_j is a measure of quasi-elasticity.

Ex/ Consider the log-lin version of the wage regression model,

$$(34)\quad \ln y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i$$

where all variables are as defined before. Using the same data, we estimate Eq.(23) by OLS and obtain the following results:

(34)' $\hat{\ln y_i} = \underset{(-1.08)}{-0.096} + \underset{(18.36)}{0.116} x_{2i} + \underset{(9.01)}{0.041} x_{3i}$

Q: How do we interpret the results?

0.116 ⇒ If education increases by 1 year, then wage rate is expected to increase by about 11.6%, ceteris paribus.

0.041 ⇒ If experience increases by 1 year, then wage rate is expected to increase by about 4.1%, ceteris paribus.

Sometimes we're interested in finding the reciprocal of quasi-elasticity of y_i with respect to x_{ij} (i.e. an absolute Δy_i due to a given relative Δx_{ij}).

(35) $E_R = \dfrac{\Delta y}{\Delta x / x} = \dfrac{\Delta y}{1} \cdot \dfrac{x}{\Delta x} = \dfrac{\Delta y}{\Delta x} . x$

This reciprocal measure of quasi-elasticity can be obtained by specifying a lin-log model (i.e. a model where y_i appears in levels but x_i appears in logs):

(36) $y_i = (\ln x_i)' \beta + \varepsilon_i = \beta_1 + \beta_2 \ln x_{i2} + \ldots + \varepsilon_i$

In this model, the coefficient of the log of a given explanatory variable can be shown to be the reciprocal measure of quasi-elasticity:

(37) $\beta_j = \dfrac{\partial y_i}{\partial x_{ij}} . x_i$

To verify this, we simply differentiate y_i with respect to x_{ij}, and rearrange terms:

$$\frac{\partial y_i}{\partial x_{ij}} = \beta_j . \frac{1}{x_{ij}} \Rightarrow \frac{\partial y_i}{\partial x_{ij}} . x_{ij} = \beta_j \quad --- \quad (38)$$

Eq.(38) confirms that β_j is the reciprocal measure of quasi-elasticity.

Ex/ Consider the lin-log version of the wage regression model,

(39) $y_i = \beta_1 + \beta_2 \ln x_{2i} + \beta_3 \ln x_{3i} + \varepsilon_i$

where all variables are as defined before. Using the same data, we estimate Eq.(39) by OLS and obtain the following results:

(39)' $\hat{y}_i = \underset{(-11.45)}{-11.321} + \underset{(16.76)}{5.999} \ln x_{2i} + \underset{(6.93)}{1.189} \ln x_{3i}$

Q: How do we interpret the results?

$5.999 \approx 6.0 \Rightarrow$ If education increases by 1%, then wage rate is expected to increase by about \$0.06/hour, ceteris paribus.

$1.189 \approx 1.2 \Rightarrow$ If experience increases by 1%, then wage rate is expected to increase by about \$0.012/hour, ceteris paribus.

c) Dummy Variables

Sometimes, we work with dummy, rather than the usual quantitative, variables. In this case, there's a need for a special interpretation of the dummy parameters.

Why? It's essentially because the parameter of a given dummy cannot be given the usual marginal effect interpretation.

To illustrate, let's revisit the wage regression model discussed in Chapter 2:

$$(40)\quad y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$$

where y is the wage rate, x_2 is the gender dummy (= 1 for male and 0 for female), x_3 is education, and x_4 is experience.

Since x_2 is a dummy variable, it doesn't make sense to interpret β_2 as "the expected change in the wage rate if gender increases by 1 unit."

Instead of making this absurd marginal effect interpretation, what we can do is evaluate the expected value of y_i for males and females:

$$(41)\quad E(y_i|x_{2i} = 1) = \beta_1 + \beta_2 + \beta_3 x_{3i} + \beta_4 x_{4i}$$

$$(42)\quad E(y_i|x_{2i} = 0) = \beta_1 + \beta_3 x_{3i} + \beta_4 x_{4i}$$

If we subtract Eq.(42) from Eq.(41), we'll obtain

$$(43)\quad E(y_i|x_{2i} = 1) - E(y_i|x_{2i} = 0) = \beta_2$$

Eq.(43) says that the difference in the expected/average wage rate between males and females is β_2, ceteris paribus.

Ex/ The OLS estimation results were given as follows:

$$(40)'\quad \underset{t}{\hat{y}_i} = \underset{(-7.269)}{-3.380} + \underset{(12.485)}{1.344}x_{2i} + \underset{(19.478)}{0.639}x_{3i} + \underset{(5.253)}{0.125}x_{4i}$$

Q: How do we interpret the results?

1.344 $\Rightarrow$ The difference in the average wage rate between males and females is about \$1.344/hour, ceteris paribus.

In other words, men are paid more than women by about \$1.344/hour on average, ceteris paribus.

0.639 $\Rightarrow$ If education increases by 1 year, then wage rate is expected to increase by about \$0.64/hour, ceteris paribus.

In other words, the marginal effect of education on wage is \$0.64/hour, regardless of gender and experience.

0.125 $\Rightarrow$ If experience increases by 1 year, then wage rate is expected to increase by about \$0.125/hour, ceteris paribus.

In other words, the marginal effect of experience on wage is \$0.125/hour, regardless of gender and education.

Q: What if we conjecture that the marginal effect of education (or experience) on wage is different between men and women?

This conjecture can be tested by introducing an interactive term between a) gender and education, and b) gender and experience:

(44) $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 (x_{2i} * x_{3i}) + \beta_6 (x_{2i} * x_{4i}) + \varepsilon_i$

As before, we evaluate the expected value of y_i for males and females:

(45) $E(y_i | x_{2i} = 1) = (\beta_1 + \beta_2) + (\beta_3 + \beta_5) x_{3i} + (\beta_4 + \beta_6) x_{4i}$

(46) $E(y_i | x_{2i} = 0) = \beta_1 + \beta_3 x_{3i} + \beta_4 x_{4i}$

Eqs.(45) and (46) suggest the following:

a) The ME of education on wage is (β_3 + β_5) for males, and β_3 for females. Hence, β_5 is the difference in the ME of education on wage across gender.

b) The ME of experience on wage is (β_4 + β_6) for males, and β_4 for females. Hence, β_6 is the difference in the ME of experience on wage across gender.

Ex/ If we estimate Eq.(44) by OLS, we'll obtain the following results:

(44)' $\hat{y}_i = -\underset{(-3.86)}{2.593} - \underset{(-0.03)}{0.030} x_{2i} + \underset{(11.22)}{0.561} x_{3i} + \underset{(4.00)}{0.142} x_{4i} + \underset{(1.99)}{0.132} (x_{2i} * x_{3i}) - \underset{(-0.45)}{0.021} (x_{2i} * x_{4i})$

Q: How do we interpret the results?

For the sake of exposition, let us ignore the statistical significance of the estimated coefficients for a moment in making the following interpretations.

Having said this, note that the ME of education on wage is 0.561 + 0.132 = 0.693 for males and the ME of experience on wage is 0.142 – 0.021 = 0.121 for males.

0.561 $\Rightarrow$ If female education increases by 1 year, then wage rate is expected to increase by about \$0.56/hour, ceteris paribus.

0.693 $\Rightarrow$ If male education increases by 1 year, then wage rate is expected to increase by about \$0.69/hour, ceteris paribus.

Hence, the marginal effect of education on wage is *higher* for males by about \$0.13/hour.

0.142 $\Rightarrow$ If female experience increases by 1 year, then wage rate is expected to increase by about \$0.14/hour, ceteris paribus.

0.121 $\Rightarrow$ If male experience increases by 1 year, then wage rate is expected to increase by about \$0.12/hour, ceteris paribus.

Hence, the marginal effect of experience on wage is *lower* for males by about \$0.021/hour.

Strictly speaking, however, we shouldn't be making an interpretation about male experience because the estimated coefficient, –0.021, is insignificant.

Now let's convert the linear version of the wage regression model in Eq.(40) into a log-linear one,

(47) $\ln y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 \ln x_{3i} + \beta_4 \ln x_{4i} + \varepsilon_i$

Since x_2 takes the value of 1 and 0 only, it's not possible to take log of this gender dummy (log of 1 is 0 and log of 0 is undefined).

Q: How do we interpret its coefficient?

On the surface, it seems that we can interpret its coefficient in the same manner we did with parameters in the log-lin model (i.e. relative Δy_i due to absolute Δx_{ij}).

However, since it doesn't make sense to differentiate $\ln y_i$ with respect to x_{i2}, we need to come up with an alternative, yet sensible interpretation.

As usual, let's evaluate the expected value of y_i for males and females. If we ignore x_3 and x_4, we'll have

(48) $E(\ln y_i | x_{2i} = 1) = \beta_1 + \beta_2$

(49) $E(\ln y_i | x_{2i} = 0) = \beta_1$

If we ignore the expected value operator and exponentiate each of the above equations, we'll have

(50) $w_1 \equiv \exp(\ln y_i | x_{2i} = 1) = \exp(\beta_1 + \beta_2) = \exp(\beta_1).\exp(\beta_2)$

(51) $w_0 \equiv \exp(\ln y_i | x_{2i} = 0) = \exp(\beta_1)$

where w_1 is male wages and w_0 is female wages.

If we subtract Eq.(51) from Eq.(50), we'll obtain the difference between male and and female wages:

(52) $$w_1 - w_0 = \exp(\beta_1).\exp(\beta_2) - \exp(\beta_1)$$
$$= \exp(\beta_1).[\exp(\beta_2) - 1] = w_0.[\exp(\beta_2) - 1]$$

If we divide both sides of Eq.(52) by w_0, we'll obtain the percentage difference between male and female wages:

(53) $$\frac{w_1 - w_0}{w_0} = \exp(\beta_2) - 1$$

Ex/ If we estimate Eq.(47) by OLS, we'll obtain the following results:

(47)' $$\hat{\ln y_i} = \underset{(-11.70)}{-2.160} + \underset{(11.80)}{0.242} x_{2i} + \underset{(18.84)}{1.257} \ln x_{3i} + \underset{(8.45)}{0.270} \ln x_{4i}$$

Q: How do we interpret the result on gender dummy?

$\hat{\beta}_2 = 0.242 \Rightarrow \exp(0.242) - 1 = 0.2738$

0.2738 $\Rightarrow$ The difference between male and female wages is about 27.38%.

In practice, researchers seldom calculate $\exp(\beta_2) - 1$. Instead, they appeal to a mathematical property which stipulates that $\exp(a) \approx (1 + a)$ if a is close to 0.

In our case, $a = 0.242$; thus, $\exp(0.242) \approx 1.242$, which means that the difference between male and female wages is about 24.2%(which is very close to 27.38%).

Note that this interpretation is not the same as saying "if gender rises by 1 unit, then wage rate is expected to rise by about 24.2%," which is nonsense.