Econometrics

	Course Calendar
Week	Main Content
Week 7	Extension of Simple Regression: Functional Forms I
Week 8	Extension of Simple Regression: Functional Forms II
Week 9	Extension of Simple Regression: Functional Forms III
Week 10	Multiple Regression -
Week 11	Multiple Regression: The Problem of Inference
Week 12	Multiple Regression: Functional Forms
Week 13	Introduction to Dummy Variables
Week 14	Introduction to Dummy Variables and Regression Methods
Week 15	Regression with Dummy Variables: Hands-on-Exercise
Week 16	Application of Regression

Econometrics
Lecture 11. Multiple Regression Analysis: Estimation and The Problem of Inference

Geetha Rani Prakasam, Ph.D

Professor,

Recap

- **Introduction to Multiple Linear Regression**
- The three-Variable Model: Notation & Assumptions
- **Multicollinearity**
- **Interpretation of Multiple Regression**
- The meaning of partial regression coefficients
- **OLS estimators, Var & SE in the case of three-variable regression**
- Properties of OLS Estimators
- Child Mortality Example

Outline

- Regression on Standardized Variables
- Introduction to Specification Bias
- **R^2 and the Adjusted-R^2**
- **Comparing Two R^2 Values**
- **MR: The Problem of Inference**
- **The Normality Assumption**
- **Hypothesis Testing in MR: General Comments**

Outline

- **Hypothesis Testing in MR**
  - Testing hy. about an individual partial regression coefficient
  - Testing the overall significance and the F test
  - The “Incremental” or “Marginal” contribution of an explanatory variable
- Summary & Conclusion

Regression on Standardized Variables- from Lec 7

- For the bivariate case, the relationship is as follows:

$$\hat{\beta}_2^* = \hat{\beta}_2\left(\frac{S_x}{S_y}\right) \quad (6.3.8)$$

- where S_x = the sample standard deviation of the X regressor and S_y = the sample standard deviation of the regressand.
- Therefore, one can crisscross between the β and beta coefficients if we know the (sample) standard deviation of the regressor and regressand.
- Now we see that this relationship holds true in the multiple regression also.
- Recall that a variable is said to be standardized or in standard deviation units if it is expressed in terms of deviation from its mean and divided by its standard deviation.

Regression on Standardized Variables: Ex

- For our child mortality example, the results are as follows:

$$\widehat{\mathrm{CM}}^{*} = -\,0.2026\,\mathrm{PGNP}_i^{*} - 0.7639\,\mathrm{FLR}_i^{*} \qquad \textbf{(7.6.3)}$$
$$\mathrm{se} = \quad (0.0713) \qquad\qquad (0.0713) \qquad\qquad r^2 = 0.7077$$

- Note: The starred variables are standardized variables.
- Also note that there is no intercept in the model for reasons already discussed in the previous lecture.
- As we can see from this regression, with FLR held constant, a standard deviation increase in PGNP leads, on average, to a 0.2026 standard deviation decrease in CM.

Regression on Standardized Variables: Ex

- Similarly, holding PGNP constant, a standard deviation increase in FLR, on average, leads to a 0.7639 standard deviation decrease in CM.
- Relatively speaking, female literacy has more impact on child mortality than per capita GNP.
- Here we will see the advantage of using standardized variables, for standardization puts all variables on equal footing because all standardized variables have zero means and unit variances.

Introduction to Specification Bias

- Recall that assumption (7.1.6) of the classical linear regression model states that the regression model used in the analysis is "correctly" specified; that is, there is no specification bias or specification error (see Lectures 2 and 3 for some introductory remarks).
- The illustrative example given in the preceding section provides a splendid opportunity not only to drive home the importance of assumption (7.1.6) but also to shed additional light on the meaning of partial regression coefficient and provide a somewhat informal introduction to the topic of specification bias.

Introduction to Specification Bias

- Assume that (7.6.1) is the “true” model explaining the behavior of child mortality in relation to per capita GNP and female literacy rate (FLR).
- But suppose we disregard FLR and estimate the following simple regression:
- $Y_i = \alpha_1 + \alpha_2 X_{2i} + u_{1i}$ (7.7.1)
- where Y = CM and X_2 = PGNP.
- Since (7.6.1) is the true model, estimating (7.7.1) would constitute a specification error; the error here consists in omitting the variable X3, the female literacy rate.
- Notice that we are using different parameter symbols (the alphas) in (7.7.1) to distinguish them from the true parameters (the betas) given in (7.6.1).

Introduction to Specification Bias

- Now will α_2 provide an unbiased estimate of the true impact of PGNP, which is given by β_2 in model (7.6.1)? In other words, will $E(\hat{\alpha}_2) = \beta_2$, where $\hat{\alpha}_2$ is the estimated value of α_2?
- In other words, will the coefficient of PGNP in (7.7.1) provide an unbiased estimate of the true impact of PGNP on CM, knowing that we have omitted the variable X_3 (FLR) from the model?
- As we would suspect, in general $\hat{\alpha}_2$ will not be an unbiased estimator of the true β_2.
- To give a glimpse of the bias, let us run the regression (7.7.1), which gave the following results.

Introduction to Specification Bias

$$\widehat{\text{CM}}_i = 157.4244 - 0.0114\ \text{PGNP}_i \qquad \textbf{(7.7.2)}$$

$$\text{se} = (9.8455) \quad (0.0032) \qquad r^2 = 0.1662$$

- Observe several things about this regression compared to the “true” multiple regression (7.6.1):
- 1. In absolute terms (i.e., disregarding the sign), the PGNP coefficient has increased from 0.0056 to 0.0114, almost a two-fold increase.
- 2. The standard errors are different.
- 3. The intercept values are different.
- 4. The r^2 values are dramatically different, although it is generally the case that, as the number of regressors in the model increases, the r^2 value increases.

Introduction to Specification Bias

- Now suppose that we regress child mortality on female literacy rate, disregarding the influence of PGNP. You will obtain the following results:

$$\widehat{\mathrm{CM}}_i = 263.8635 - 2.3905\ \mathrm{FLR}_i \qquad \textbf{(7.7.3)}$$

$$\mathrm{se} = (21.2249) \quad (0.2133) \qquad r^2 = 0.6696$$

- Again if we compare the results of this (misspecified) regression with the "true" multiple regression, we will see that the results are different, although the difference here is not as noticeable as in the case of regression (7.7.2).
- The important point to note is that serious consequences can ensue if we misfit a model.

R^2 and the Adjusted-R^2

- R^2 is a non-decreasing function of the number of explanatory variables.
- An additional X variable will not decrease R^2

  R^2= ESS/TSS = 1- RSS/TSS = $1-[\sum \hat{u}^2_i / \sum y^2_i]$ (7.8.1)
- This will make the wrong direction by adding more irrelevant variables into the regression and give an idea for an adjusted-R^2 (R_{bar}) by taking account of degree of freedom
- $R^2_{bar} = 1- [\sum \hat{u}^2_i/(n-k)] / [\sum y^2_i/(n-1)]$, or (7.8.2)

R^2 and the Adjusted-R^2

$R^2_{bar} = 1 - \sigma\hat{}^2 / S^2_Y$ (S^2_Y is sample variance of Y)

K= number of parameters including intercept term

- By substituting (7.8.1) into (7.8.2) we get

  $R^2_{bar} = 1 - (1 - R^2)\,(n-1)/(n-k)$ (7.8.4)

- For $k > 1$, $R^2_{bar} < R^2$ thus when number of X variables increases R^2_{bar} increases less than R^2 and R^2_{bar} can be negative.
- The term adjusted means adjusted for the df associated with the sums of squares entering into (7.8.1):
- $\hat{u}_{2i}$ has $n - k$ df in a model involving k parameters, which include the intercept term, and Y_i has $n - 1$ df.
- For the three-variable case, we know that $\hat{u}_{2i}$ has $n - 3$ df.

R^2 and the Adjusted-R^2

- Comparing Two R^2 Values:

  To compare, the size n and the dependent variable must be the same
- Maximizing adjusted-R^2: Choosing the model that gives the highest R^2_{bar} may be dangerous, for in regression our objective is not for that but for obtaining the dependable estimates of the true population regression coefficients and draw statistical inferences about them
- Should be more concerned about the logical or theoretical relevance of the explanatory variables to the dependent variable and their statistical significance

Comparing Two R^2 Values

- It is crucial to note that in comparing two models on the basis of the coefficient of determination, whether adjusted or not, the sample size n and the dependent variable must be the same; the explanatory variables may take any form.
- Thus for the models

$$\ln Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \qquad \textbf{(7.8.6)}$$

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + u_i \qquad \textbf{(7.8.7)}$$

- the computed R^2 terms cannot be compared. The reason is as follows: By definition, R^2 measures the proportion of the variation in the dependent variable accounted for by the explanatory variable(s).

Comparing Two R^2 Values

- Therefore, in (7.8.6) R^2 measures the proportion of the variation in lnY explained by X_2 and X_3, whereas in (7.8.7) it measures the proportion of the variation in Y, and the two are not the same thing:
- As noted in Lectures 7 to 10, a change in lnY gives a relative or proportional change in Y, whereas a change in Y gives an absolute change.
- Therefore, var $\hat{Y}_i$/var(Y_i) is not equal to var (ln Y_i) ^/var (ln Y_i); that is, the two coefficients of determination are not the same.
- So, we may not be able to compare the R^2's of two models when the regressand is not in the same form.
- To answer this question, let us first consider a numerical example.

Ex. 7.2 Coffee Consumption in The US, 1970–1980

- Consider the data in Table 7.1. The data pertain to consumption of cups of coffee per day (Y) and real retail price of coffee (X) in the United States for years 1970–1980.
- Applying OLS to the data, we obtain the following regression results:

$$
\begin{aligned}
\hat{Y}_t &= 2.6911 - 0.4795X_t \\
\text{se} &= (0.1216) \quad (0.1140) \qquad \text{RSS} = 0.1491;\ r^2 = 0.6628
\end{aligned}
\tag{7.8.8}
$$

- The results make economic sense:
- As the price of coffee increases, on average, coffee consumption goes down by about half a cup per day.

Ex. 7.2 Coffee Consumption in The US, 1970–1980

- We can readily verify that the slope coefficient is statistically significant.
- The r^2 value of about 0.66 means that the price of coffee explains about 66 percent of the variation in coffee consumption.
- From the same data, the following double log, or constant elasticity, model can be estimated:

$$\widehat{\ln Y_t} = 0.7774 - 0.2530 \ln X_t$$

$$\text{se} = (0.0152) \quad (0.0494) \qquad \text{RSS} = 0.0226;\ r^2 = 0.7448 \tag{7.8.9}$$

- Since this is a double log model, the slope coefficient gives a direct estimate of the price elasticity coefficient.
- In the present instance, it tells us that if the price of coffee per pound goes up by 1 percent, on average, per day coffee consumption goes down by about 0.25 percent.

TABLE 7.1: U.S. Coffee Consumption (Y) in Relation to Average Real Retail Price (X),* 1970–1980

	Y	X
Year	Cups per person per day	$ per lb
1970	2.57	0.77
1971	2.5	0.74
1972	2.35	0.72
1973	2.3	0.73
1974	2.25	0.76
1975	2.2	0.75
1976	2.11	1.08
1977	1.94	1.81
1978	1.97	1.39
1979	2.06	1.2
1980	2.02	1.17

*Note: The nominal price was divided by the Consumer Price Index (CPI) for food and beverages, 1967 = 100. Source: The data for Y are from Summary of National Coffee Drinking Study, Data Group, Elkins Park, Penn., 1981; and the data on nominal X (i.e., X in current prices) are from Nielsen Food Index, A. C. Nielsen, New York, 1981.

Basic Econometrics, Damodar Gujarati, Page, 220

Ex. 7.2 Coffee Consumption in The US, 1970–1980

- Remember that in the linear model (7.8.8) the slope coefficient only gives the rate of change of coffee consumption with respect to price. (How do we estimate the price elasticity for the linear model?).
- The r^2 value of about 0.74 means that about 74 percent of the variation in the log of coffee demand is explained by the variation in the log of coffee price.
- Since the r^2 value of the linear model of 0.6628 is smaller than the r^2 value of 0.7448 of the log–linear model, you might be tempted to choose the latter model because of its high r^2 value.
- But for reasons already noted, we cannot do so. But if we do want to compare the two r^2 values, you may proceed as follows:

Ex. 7.2 Coffee Consumption in The US, 1970–1980

- 1. Obtain ln Yt from (7.8.9) for each observation; that is, obtain the estimated log value of each observation from this model.
- Take the antilog of these values and then compute r^2 between these antilog values and actual Y_t in the manner indicated by Eq. (3.5.14).
- This r^2 value is comparable to the r^2 value of the linear model (7.8.8).
- 2. Alternatively, assuming all Y values are positive, take logarithms of the Y values, lnY.
- Obtain the estimated Y values, $\hat{Y}_t$, from the linear model (7.8.8), take the logarithms of these estimated Y values (i.e., $\ln \hat{Y}_t$) and compute the r^2 between $(\ln Y_t)$ and $(\ln \hat{Y}_t)$ in the manner indicated in Eq. (3.5.14).
- This r^2 value is comparable to the r^2 value obtained from (7.8.9).

Ex. 7.2 Coffee Consumption in The US, 1970–1980

- For our coffee example, we present the necessary raw data to compute the comparable r^2's in Table 7.2.
- To compare the r^2 value of the linear model (7.8.8) with that of (7.8.9), we first obtain log of ($\hat{Y}_t$) [given in column (6) of Table 7.2], then we obtain the log of actual Y values [given in column (5) of the table], and then compute r^2 between these two sets of values using Eq. (3.5.14).
- The result is an r^2 value of 0.7318, which is now comparable with the r^2 value of the log–linear model of 0.7448.
- Now the difference between the two r^2 values is very small.

Ex. 7.2 Coffee Consumption in The US, 1970–1980

- On the other hand, if we want to compare the r^2 value of the log–linear model with the linear model, we obtain ln Yt for each observation from (7.8.9) [given in column (3) of the table], obtain their antilog values [given in column (4) of the table], and finally compute r^2 between these antilog values and the actual Y values, using formula (3.5.14).
- This will give an r^2 value of 0.7187, which is slightly higher than that obtained from the linear model (7.8.8), namely, 0.6628.
- Using either method, it seems that the log–linear model gives a slightly better fit.

TABLE 7.2 RAW DATA FOR COMPARING TWO R^2 VALUES

Year	Y_t (1)	$\hat{Y}_t$ (2)	$\widehat{\ln Y_t}$ (3)	Antilog of $\widehat{\ln Y_t}$ (4)	$\ln Y_t$ (5)	$\ln(\hat{Y}_t)$ (6)
1970	2.57	2.321887	0.843555	2.324616	0.943906	0.842380
1971	2.50	2.336272	0.853611	2.348111	0.916291	0.848557
1972	2.35	2.345863	0.860544	2.364447	0.854415	0.852653
1973	2.30	2.341068	0.857054	2.356209	0.832909	0.850607
1974	2.25	2.326682	0.846863	2.332318	0.810930	0.844443
1975	2.20	2.331477	0.850214	2.340149	0.788457	0.846502
1976	2.11	2.173233	0.757943	2.133882	0.746688	0.776216
1977	1.94	1.823176	0.627279	1.872508	0.662688	0.600580
1978	1.97	2.024579	0.694089	2.001884	0.678034	0.705362
1979	2.06	2.115689	0.731282	2.077742	0.722706	0.749381
1980	2.02	2.130075	0.737688	2.091096	0.703098	0.756157

Notes: Column (1): Actual Y values from Table 7.1 Column (2): Estimated Y values from the linear model (7.8.8) Column (3): Estimated log Y values from the double-log model (7.8.9) Column (4): Antilog of values in column (3) Column (5): Log values of Y in column (1) Column (6): Log values of Yˆt in column (2)

Multiple Regression:

The Problem of Inference

MR: The Problem of Inference

- It extends interval estimation and hypothesis testing developed to MR.
- In many ways the concepts developed in Lecture 6 can be applied straightforwardly to the multiple regression model.
- Few additional features are unique to such models, and it is these features that will be discussed here.

8.1 The Normality Assumption

- It is assumed that the ui follow the normal distribution with zero mean and constant variance, σ^2 .
- We continue to make the same assumption for multiple regression models.
- With the normality assumption and following the discussion of Lectures 2 and 3, we find that the OLS estimators of the partial regression coefficients are best linear unbiased estimators (BLUE).
- Moreover, the estimators ˆ β_2, ˆβ_3, and ˆβ_1 are themselves normally distributed with means equal to true β_2, β_3, and β_1 and the variances as noted in Lecture 10.

8.1 The Normality Assumption

- Also, $(n - 3)\hat{\sigma}^2/\sigma^2$ follows the χ^2 distribution with n – 3 df, and the three OLS estimators are distributed independently of $\hat{\sigma}^2$.
- One can show that, upon replacing σ^2 by its unbiased estimator $\hat{\sigma}^2$ in the computation of the SE, each of the following variables
- $t = (\hat{\beta}_1 - \beta_1)/ (se (\hat{\beta}_1)$ (8.1.1)
- $t = (\hat{\beta}_2 - \beta_2)/ (se (\hat{\beta}_2)$ (8.1.2)
- $t = (\hat{\beta}_3 - \beta_3)/ (se (\hat{\beta}_3)$ (8.1.3)

follows the t distribution with n – 3 df.

8.1 The Normality Assumption

- Note that the df are now $n - 3$ because in computing $\hat{u}^2_i$ and hence $\hat{\sigma}^2$ we need to estimate the three partial regression coefficients, which therefore put three restrictions on the residual sum of squares (RSS) (following this logic in the four-variable case there will be $n - 4$ df, and so on).
- Therefore, the t distribution can be used to establish confidence intervals as well as test statistical hypotheses about the true population partial regression coefficients.
- Similarly, the χ^2 distribution can be used to test hypotheses about the true σ^2.

8.2 EX. 8.1: Child Mortality Example Revisited

- In Lecture 10, we regressed child mortality (CM) on per capita GNP (PGNP) and the female literacy rate (FLR) for a sample of 64 countries.
- The regression results given in eq. (7.6.2) are reproduced below with some additional information:

$$
\begin{array}{rlll}
\widehat{\mathrm{CM}}_i = & 263.6416 & - \quad 0.0056\ \mathrm{PGNP}_i & - \quad 2.2316\ \mathrm{FLR}_i \\
\mathrm{se} = & (11.5932) & (0.0019) & (0.2099) \\
t = & (22.7411) & (-2.8187) & (-10.6293) \\
p \text{ value} = & (0.0000)^{*} & (0.0065) & (0.0000)^{*} \\
 & & R^2 = 0.7077 & \bar{R}^2 = 0.6981
\end{array}
\qquad \textbf{(8.2.1)}
$$

- Where * denotes extremely low value.

8.2 EX. 8.1: Child Mortality Example Revisited

- In Eq. (8.2.1) we have followed the format:-
- where the figures in the first set of parentheses are the estimated standard errors,
- those in the second set are the t values under the null hypothesis that the relevant population coefficient has a value of zero, and
- those in the third are the estimated p values.
- Also given are R^2 and adjusted R^2 values.
- We have already interpreted this regression in Example 7.1.
- What about the statistical significance of the observed results?

8.2 EX. 8.1: Child Mortality Example Revisited

- Consider, for example, the coefficient of PGNP of −0.0056. Is this coefficient statistically significant, that is, statistically different from zero?
- Likewise, is the coefficient of FLR of −2.2316 statistically significant?
- Are both coefficients statistically significant?
- To answer this and related questions, let us first consider the kinds of hypothesis testing that one may encounter in the context of a multiple regression model.

8-3. Hypothesis Testing in MR: General Comments

- In MR, hypothesis testing assumes several interesting forms, such as the following:

1. Testing hypotheses about an individual partial regression coefficient
2. Testing the overall significance of the estimated multiple regression model, that is, finding out if all the partial slope coefficients are simultaneously equal to zero
3. Testing that two or more coefficients are equal to one another
4. Testing that the partial regression coefficients satisfy certain restrictions
5. Testing the stability of the estimated regression model over time or in different cross-sectional units
6. Testing the functional form of regression models

8-4. Hypothesis testing about individual partial regression coefficients

- With the assumption that $u_i \sim N(0,\sigma^2)$ we can use t-test to test a hypothesis about any individual partial regression coefficient.
- H_0: $\beta_2 = 0$
- H_1: $\beta_2 \# 0$
- If the computed t value > critical t value at the chosen level of significance, we may reject the null hypothesis; otherwise, we may not reject it.
- If we invoke the assumption that ui ∼ N(0, σ2), then, as noted earlier, we can use the t test to test a hypothesis about any individual partial regression coefficient.

8-4. Hypothesis testing about individual partial regression coefficients: Ex

- Consider the child mortality regression, Eq. (8.1.4). Let us postulate that $H_0 : \beta_2 = 0$ and $H_1 : \beta_2 \# 0$
- The null hy. states that, with X_3 (female literacy rate) held constant, X_2 (PGNP) has no (linear) influence on Y (child mortality).
- To test the null hy., we use the t test given in Eq. (8.1.2).
- If the computed t value exceeds the critical t value at the chosen level of significance, we may reject the null hy.; otherwise, we may not reject it.
- In our example, using Eq. (8.1.2) and noting that $\beta_2 = 0$ under the null hypothesis, we obtain

8-4. Hypothesis testing about individual partial regression coefficients: Ex

- $t = -0.0056 / 0.0020 = -2.8187$ (8.4.1)
- as shown in Eq. (8.1.2).
- Df = 64-3=61.
- Using these df, and assume α, the level of significance (i.e., the probability of committing a Type I error) of 5 percent, the critical t value is 2.0 for a two-tail test (look up $t_{\alpha/2}$ for 60 df) or 1.671 for a one-tail test (look up t_α for 60 df).
- Since the computed t value of 2.8187 (in absolute terms) exceeds the critical t value of 2, we can reject the null hypothesis that PGNP has no effect on CM.

8-4. Hypothesis testing about individual partial regression coefficients: Ex

- To put it more positively, with the FLR held constant, per capita GNP has a significant (negative) effect on CM, as one would expect a priori.
- In practice, we can use the p value given in Eq. (8.1.4), which in the present case is 0.0065.
- The interpretation of this p value (i.e., the exact level of significance) is that if the null hypothesis were true, the probability of obtaining a t value of as much as 2.8187 or greater (in absolute terms) is only 0.0065 or 0.65 percent, which is a small probability, much smaller than the value of $\alpha = 5\%$.

8-5. Testing the overall significance of a multiple regression: The F-Test

- For $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots\ldots\ldots + \beta_k X_{ki} + u_i$
- To test the hypothesis $H_0: \beta_2 = \beta_3 = \ldots = \beta_k = 0$ (all slope coefficients are simultaneously zero) versus H_1: Not all slope coefficients are simultaneously zero.
- A test of such a hypothesis is called a test of the **overall significance of the observed** or estimated regression line, that is, whether Y is linearly related to both X2 and X3.
- However, this joint hypothesis can be tested by the **analysis of variance (ANOVA) technique**

ANOVA Table for the Child Mortality Ex.

Source of Variation	SSS	df	MSS
Due to regression	257,362.4	2	128,681.2
Due to residuals	106,315.6	61	1742.88
Total	363,678	63	

F = 128,681.2 / 1742.88
= 73.8325

The p value of obtaining an F value of as much as 73.8325 or greater is almost zero, leading to the rejection of the hypothesis that together PGNP & FLR have no effect on child mortality.

8-5. Testing the overall significance of a MR: **The F-Test**

- $F=(ESS/df)/(RSS/df)=(ESS/(k-1))/(RSS/(n-k))$ (8.5.7)
- (k = total number of parameters to be estimated including intercept)
- If $F > F_{critical} = F_{\alpha}(k-1,n-k)$, reject H_0
- Otherwise we do not reject it.
- Alternatively, if the p-value of F obtained from (8.5.7) is sufficiently low, one can reject H_0.
- An important relationship between R^2 and F:

  $F=(ESS/(k-1))/(RSS/(n-k))$ or

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \qquad (8.5.1)$$

Summary and Conclusions

- Standardized variables with a CM example
- Specification Bias with a CM example
- R^2 and adjusted R^2 to take care of the number of parameters estimated in the model
- Although R^2 and adjusted R^2 are overall measures of how the chosen model fits a given set of data, their importance should not be overplayed.
- What is critical is the underlying theoretical expectations about the model in terms of a priori signs of the coefficients of the variables entering the model and, as it is shown in the following lecture, their statistical significance.

Summary and Conclusions

- 1. This lecture extended and refined the ideas of interval estimation and hypothesis testing first introduced in Lecture 6 in the context of the two-variable linear regression model.
- 2. In a multiple regression, testing the individual significance of a partial regression coefficient (using the t test) and testing the overall significance of the regression (i.e., H_0: all partial slope coefficients are zero or $R^2 = 0$) are not the same thing.

Summary and Conclusions

3. In particular, the finding that one or more partial regression coefficients are statistically insignificant on the basis of the individual t test does not mean that all partial regression coefficients are also (collectively) statistically insignificant. The latter hypothesis can be tested only by the F test
4. The F test can test a variety of hypotheses, such as whether (1) an individual regression coefficient is statistically significant, (2) all partial slope coefficients are zero, (3) two or more coefficients are statistically equal, (4) the coefficients satisfy some linear restrictions, and (5) there is structural stability of the regression model.

Reference

Basic Econometrics by Damodar Gujarati, **Chapter 7:MULTIPLE REGRESSION ANALYSIS: The Problem of Estimation**

Basic Econometrics by Damodar Gujarati, **Chapter 8:MULTIPLE REGRESSION ANALYSIS: The Problem of Inference**

What next?

- Multiple Regression: Inference-Hypothesis testing
- Different functional forms and applications