ECONOMETRICS I

Course evaluations, etc.

- Quiz/Assignments: 30%
  Midsem: 30%
  Endsem: 40%

- Exams are open-book, open-notes (and not open-internet)

Office hours

- Gaurav Arora (Instructor):
  Tuesday, Friday: 4-5 pm
  By email/appointment

Plagiarism Policy

- Link: https://iiitd.ac.in/academics/resources/academic-dishonesty

- Repercussions

	Misconduct/use of unfair means in assignments/projects	Penalty
1.	First misconduct/use of unfair means during the entire stay in IIIT-Delhi	Zero in the assignment (awarded by faculty) + one letter grade less in the course.
2.	Second misconduct/use of unfair means during entire stay in IIIT-Delhi	Student is assigned an F grade in the course.
3.	Third or further misconduct/use of unfair means during the entire stay in IIIT-Delhi	Student is assigned an F grade in the course and the case is reported to DAC, who may suspend the student for 1 semester to a year.

	Misconduct/use of unfair means in quiz/midsem/endsem	Penalty
1.	First misconduct/use of unfair means during the entire stay in IIIT-Delhi	Student is assigned an F grade in the course.
2.	Second misconduct/use of unfair means during the entire stay in IIIT-Delhi	Student is assigned an F grade in the course and student may be suspended from the program for 1-2 semesters by DAC.
3.	Third misconduct/use of unfair means during the entire stay in IIIT-Delhi	Student is assigned an F grade in the course and student's program may be terminated by DAC.

Source: IIITD New Plagiarism Policy 2017.

Reference book

Introductory Econometrics: A modern approach (4th ed, Cengage India) by **Jeffrey M. Wooldridge**

About the author: Jeffrey M. Wooldridge

Jeffrey M. Wooldridge is University Distinguished Professor of Economics at Michigan State University, where he has taught since 1991. From 1986 to 1991, Dr. Wooldridge was an assistant professor of economics at the Massachusetts Institute of Technology. He received his bachelor of arts, with majors in computer science and economics, from the University of California, Berkeley, in 1982 and received his doctorate in economics in 1986 from the University of California, San Diego. Dr. Wooldridge has published more than three dozen articles in internationally recognized journals, as well as several book chapters. He is also the author of *Econometric Analysis of Cross Section and Panel Data*, second edition. His awards include an Alfred P. Sloan Research Fellowship, the Plura Scripsit award from *Econometric Theory*, the Sir Richard Stone prize from the *Journal of Applied Econometrics*, and three graduate teacher-of-the-year awards from MIT. He is a

What is econometrics?

- An envelope of methods to formally
  - Evaluate (the impact) a government or business policy
    - E.g., impact of a job training program;
    - impact of the plagiarism policy;
    - impact of odd-even policy.

  - Test a simple economic theory
    - E.g., Diversification of risks (i.e., investing in mutual funds, for example) would yield higher returns over time;
    - Minimum legal wages reduces an individual's propensity to participate in crime

  - Estimate a simple economic (or even social) relationship
    - E.g., partisan political events increase the number of posts on Twitter/Facebook;
    - higher minimum wages reduce the crime rate of a city

A (vague) econometric model

$$y_{[\ ]} = \beta_0 + \beta_1 x_{1[\ ]} + \beta_2 x_{2[\ ]} + \ldots + \beta_k x_{k[\ ]} + u_{[\ ]},$$

where

$y_{[\ ]}$ is the dependent/explained/predicted/response variable for unit [],

$x_{k[\ ]}$ is the kth regressand/explainatory variable/control/predictor/regressor/covariate for unit [],

β_0 is the intercept,

β_k is the coefficient of $x_{k[\ ]}$, also termed as the slope variable,

$u_{[\ ]}$ is the disturbance or error term.

Data Structures

- Cross-sectional data

- Time-series data

- Panel or longitudinal data

Why study econometrics?

- Why should I study econometrics when we already have ML, Data Science and Statistics courses on-campus?

- After all, multiple regression analysis is the mainstay of all these disciplines!

- The journey from *what* to **why**?

- *Correlation* to **Causality**

- *Experimental Data* to **Observational Data**

- *Goodness of fit to* **inference-based model**

Structuring an *empirical* <u>analysis</u>

- An empirical analysis uses data to test a theory or estimate a relationship.

- Step 1: Careful formulation of a research question. (seems trivial, it isn't)
  - Econometric methods can be used to answer a wide array of questions

- Step 2: In some cases, building an economic model is necessary.
  - E.g., econometric analysis of demand

  In some cases, applying economic reasoning or intuition, or even informal logic works.
  - E.g., impact of job training program on worker productivity

Example 1

EXAMPLE 1.1   ECONOMIC MODEL OF CRIME

In a seminal article, Nobel Prize winner Gary Becker postulated a utility maximization framework to describe an individual's participation in crime. Certain crimes have clear economic rewards, but most criminal behaviors have costs. The opportunity costs of crime prevent the criminal from participating in other activities such as legal employment. In addition, there are costs associated with the possibility of being caught and then, if convicted, the costs associated with incarceration. From Becker's perspective, the decision to undertake illegal activity is one of resource allocation, with the benefits and costs of competing activities taken into account.

Under general assumptions, we can derive an equation describing the amount of time spent in criminal activity as a function of various factors. We might represent such a function as

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7), \qquad [1.1]$$

where

y = hours spent in criminal activities,
x_1 = "wage" for an hour spent in criminal activity,
x_2 = hourly wage in legal employment,
x_3 = income other than from crime or employment,
x_4 = probability of getting caught,
x_5 = probability of being convicted if caught,
x_6 = expected sentence if convicted, and
x_7 = age.

Other factors generally affect a person's decision to participate in crime, but the list above is representative of what might result from a formal economic analysis. As is common in economic theory, we have not been specific about the function $f(\cdot)$ in (1.1). This function depends on an underlying utility function, which is rarely known. Nevertheless, we can use economic theory—or introspection—to predict the effect that each variable would have on criminal activity. This is the basis for an econometric analysis of individual criminal activity.

The ambiguities inherent in the economic model of crime are resolved by specifying a particular econometric model:

$$crime = \beta_0 + \beta_1 wage_m + \beta_2 othinc + \beta_3 freqarr + \beta_4 freqconv + \beta_5 avgsen + \beta_6 age + u,$$ [1.3]

where

crime	= some measure of the frequency of criminal activity,
$wage_m$	= the wage that can be earned in legal employment,
othinc	= the income from other sources (assets, inheritance, and so on),
freqarr	= the frequency of arrests for prior infractions (to approximate the probability of arrest),
freqconv	= the frequency of conviction, and
avgsen	= the average sentence length after conviction.

The choice of these variables is determined by the economic theory as well as data considerations. The term *u* contains unobserved factors, such as the wage for criminal activity, moral character, family background, and errors in measuring things like criminal

Example 2

EXAMPLE 1.2 **JOB TRAINING AND WORKER PRODUCTIVITY**

Consider the problem posed at the beginning of Section 1.1. A labor economist would like to examine the effects of job training on worker productivity. In this case, there is little need for formal economic theory. Basic economic understanding is sufficient for realizing that factors such as education, experience, and training affect worker productivity. Also, economists are well aware that workers are paid commensurate with their productivity. This simple reasoning leads to a model such as

$$wage = f(educ, exper, training), \qquad [1.2]$$

where

$$wage = \text{hourly wage,}$$
$$educ = \text{years of formal education,}$$
$$exper = \text{years of workforce experience, and}$$
$$training = \text{weeks spent in job training.}$$

Again, other factors generally affect the wage rate, but equation (1.2) captures the essence of the problem.

Causatily versus Correlation

- Notion of Ceteris Paribus or *all else held constant*