ECO 372: Introduction to Econometrics

Spring 2025

Lecture 1: Introduction- What is Econometrics?

Sakib Bin Amin, Ph.D.

Associate Professor in Economics

What is Econometrics?

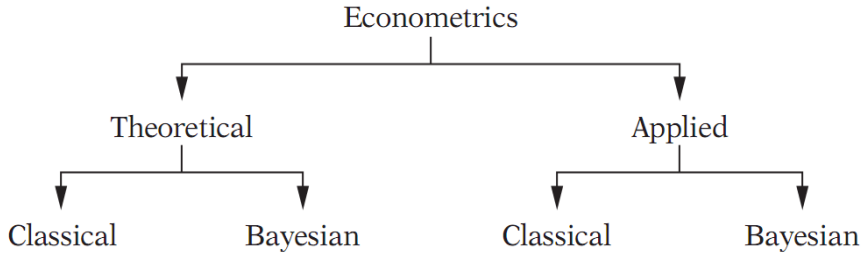
- Literally interpreted, econometrics means "economic measurement."
- Although measurement is an important part of econometrics, the scope of econometrics is much broader.
- Ask half dozen econometricians what econometrics is, and you could get a half dozen different answers.
  - *Econometrics is the science of testing economic theories*
  - *Econometrics is the set of tools used for forecasting future values of economic variables, such as GDP*
  - *Econometrics is the process of fitting mathematical economic models to real-world data*
  - *It is the science and art of using historical data to make numerical, or quantitative, policy recommendations in government and business*
- In fact, all these answers are right! econometrics is the science and art of using economic theory and statistical techniques to analyse economic data.

What Distinguishes Econometrics?

- Econometrics is much more than just statistics using economic data, although it is of course very closely related to statistics.
- Econometrics has special focus on prediction. In many respects the goal of econometrics is to help agents (consumers, firms, investors, policymakers, etc.) make better decisions, and good forecasts are key inputs to good decisions.
  - *Think about negative or inverse relationship between the price and quantity demanded of a commodity. How much the quantity will go up or down as a result of a certain change in the price.*
- Econometrics confronts the special issues and features that arise routinely in economic data, such as trends, seasonality and cycles.
  - *Think about the demand and price distortions in some commodity markets only during Ramdan. Why this type of particular change is not observed in rest of the months? Can we predict and forecast that pattern? What government intervention needed to reduce such distorions.*
- Econometrics confronts the special problems arising due to its largely non-experimental nature
  - *Model misspecication, structural changes, etc.*

Types of Econometrics

- Econometrics can be divided into theoretical and applied components.
- In each category, one can approach the subject in the **Classical** or **Bayesian** tradition.
- Theoretical econometrics is concerned with the development of appropriate methods for measuring economic relationships specified by econometric models.
- Theoretical econometrics must spell out the assumptions of this method, its properties, and what happens to these properties when one or more of the assumptions of the method are not fulfilled
- In applied econometrics we use the tools of theoretical econometrics to study some special field(s) of economics and business.
- A Bayesian Regression model's output is obtained from a probability distribution, as opposed to traditional regression techniques, which simply derive the output from a single number of each attribute.



Approach to Solution: An Econometrician's Perspective

Although there are several schools of thought on econometric methodology, the basic approach is somewhat same

Broadly speaking, traditional econometric methodology proceeds along the following:

—Statement of theory or hypothesis or any economic question

—Specification of the mathematical model of the theory

—Specification of the statistical, or econometric, model

—Obtaining the data

—Estimation of the parameters of the econometric model

—Hypothesis testing (if needed)

—Forecasting or prediction (if needed)

—Using the model for control or policy purposes

Approach to Solution: A Simple Example

- The Question: Does schooling increase earnings?

$$Earnings = f(Schooling)$$

- But how do we measure it? let's use wage
- Simple linear equation:

$$Wage = a + b\ (Schooling)$$

- We have seen this before. Consider: y = 3 + 4 x
- This is just a linear equation (slope-Intercept form)!
- There might be some unobservable components, which can not be modelled with data due to difficulty in measurement. So, add an extra term, namely error term
  - $Wage = a + b\ (Schooling) + Error$
- With this, we can estimate the coefficient (i.e., b) to answer the question.

Data: Sources and Types

- **Experimental Data:** Experimental data come from experiments designed to evaluate a treatment or policy or to investigate a causal effect.
  - *Finding out the treatment effect of providing free books and food for poor school going children in two villages.*
  - *You will have two sets of datasets for control group (that receives no treatment) and treatment group (that receives treatment.*
  - *Randomisation is conditioned here.*
- **Observational Data:** Data obtained by observing actual behavior outside an experimental setting are called observational data.
  - *Examples: GDP, Export, Import. Household earning, Education, Commodity Prices , etc.*
  - *Randomisation is not conditioned here.*

Whether the data are experimental or observational, data sets come in three main types: **cross-sectional data**, **time series data**, and **panel data**.

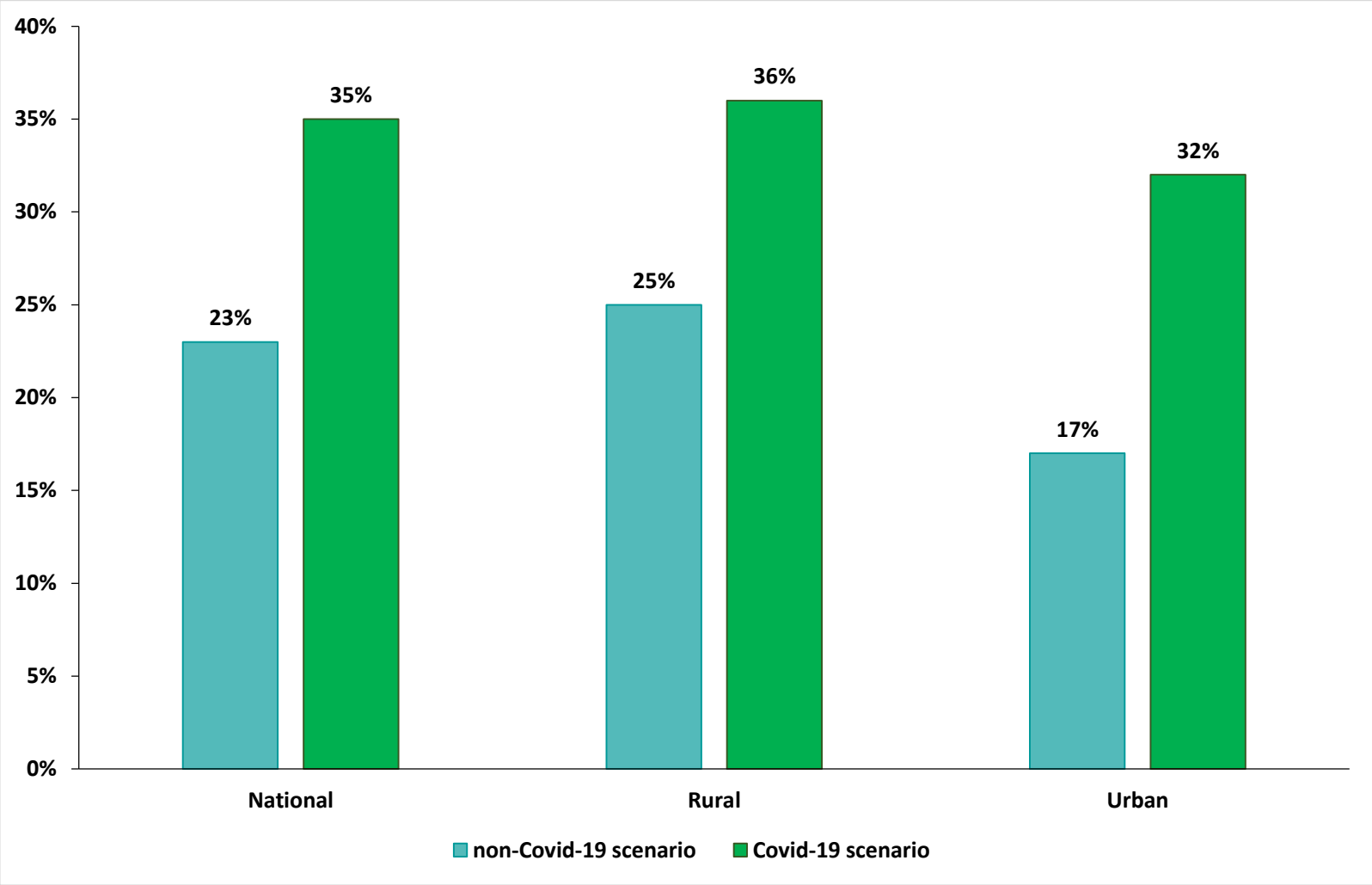
Data: Sources and Types

Cross-Sectional Data:

- Dara are collected across several units at a single point or period of time
- Units: economic agents, e.g. individuals, households, investors, firms, economic sectors, cities, countries, etc.
- In general: the order of observations has no meaning.
- Popular to use index *i*.
- Optimal: the data are a random sample of the underlying population
- Cross-Sectional data allow to explain differences between individual units

Data: Sources and Types

Poverty Scenario in Bangladesh in 2020



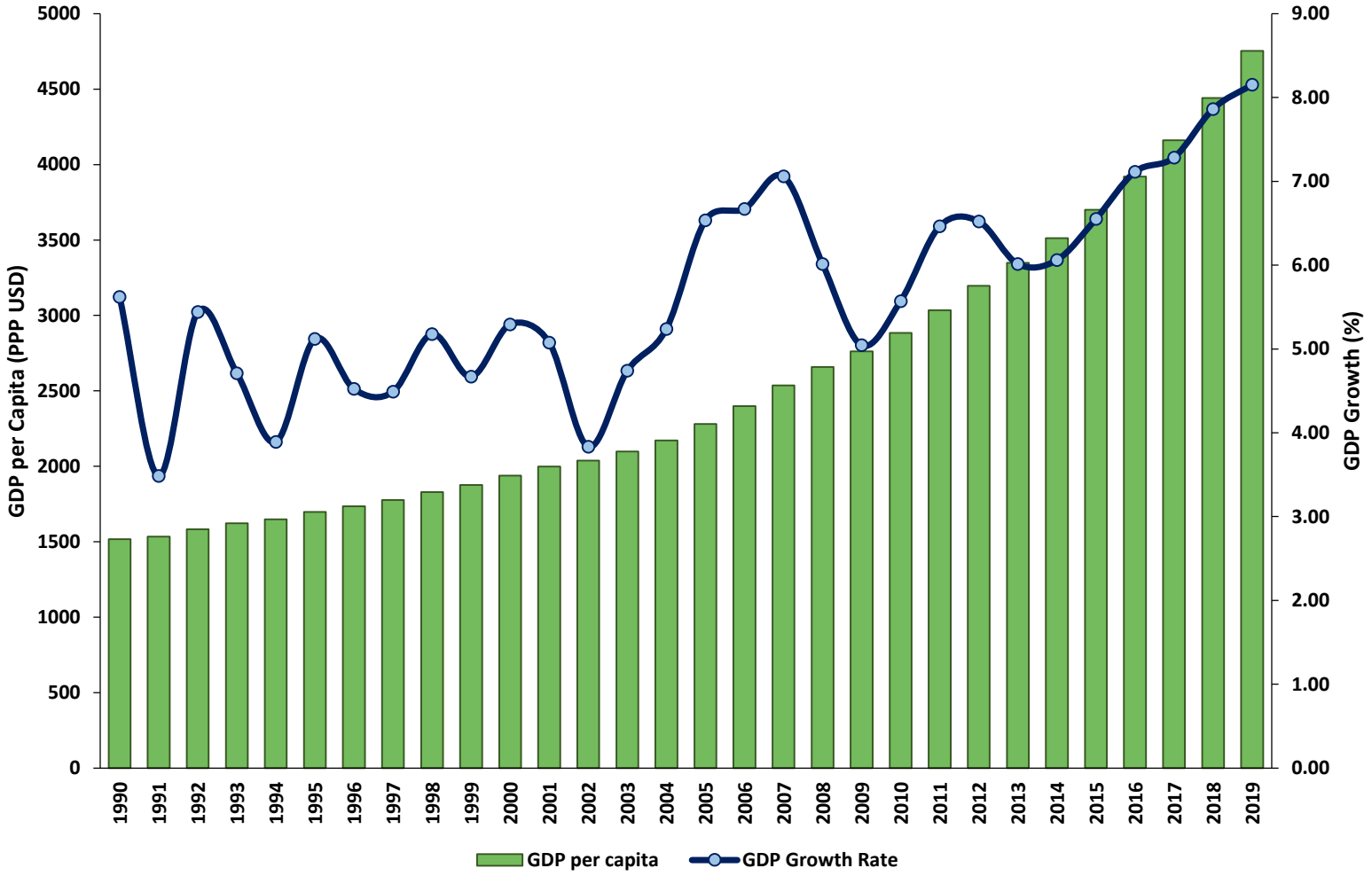
Data: Sources and Types

Time Series Data:

- Data are sampled across differing points/periods of time.
- Popular to use index *t*.
- Sampling frequency is important:
  - *variable versus fixed*
  - *fixed: annually, quarterly, monthly, weekly, daily, intra-daily*
  - *variable: ticker data, duration data (e.g. unemployment spells)*
- Time series data allow the analysis of dynamic effects.
- Univariate versus multivariate time series data

Data: Sources and Types

GDP Growth and Per Capita Income in Bangladesh



Data: Sources and Types

Panel Data:

- Also called longitudinal data, are data for multiple entities in which each entity is observed at two or more time periods.
- Individual units remain identical in each cross-sectional sample.
- Use of double index: i, t where i = $1; \ldots ; N$ and $t = 1; \ldots ; T$.
- Typical problem: missing values - for some units and periods there are no data
- Panel data can be used to learn about economic relationships from the experiences of the many different entities in the data set and from the evolution over time of the variables for each entity.

Data: Sources and Types

Regional Overview of Biogas Generation