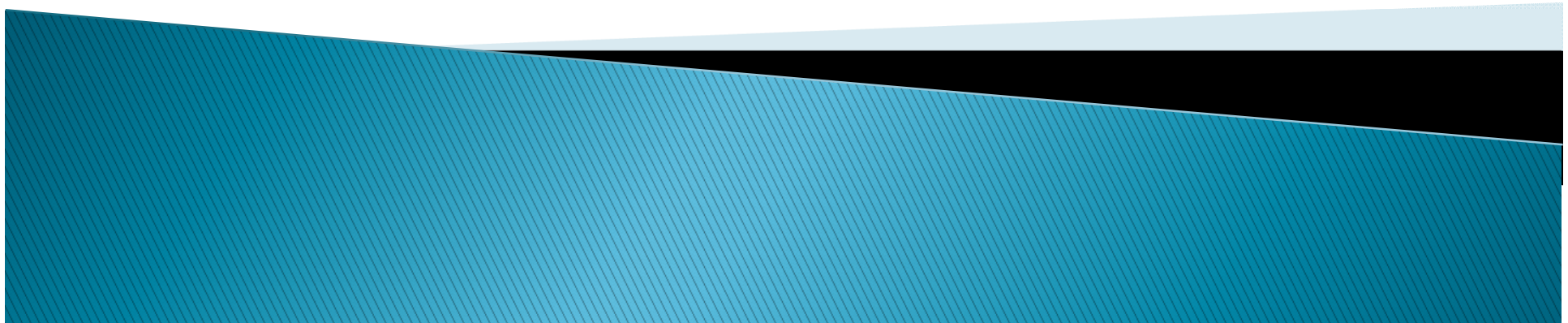


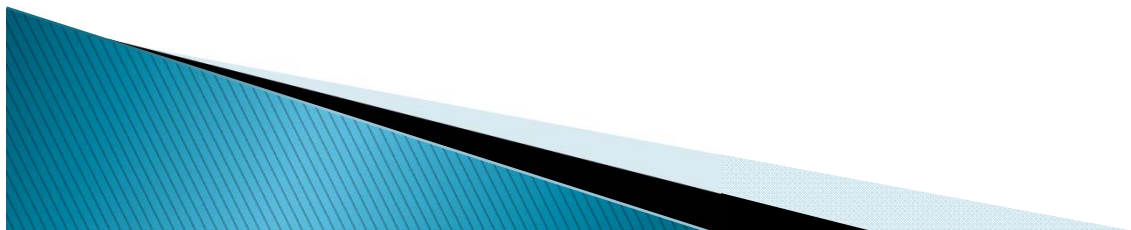
EE 325

The Nature of Regression Analysis

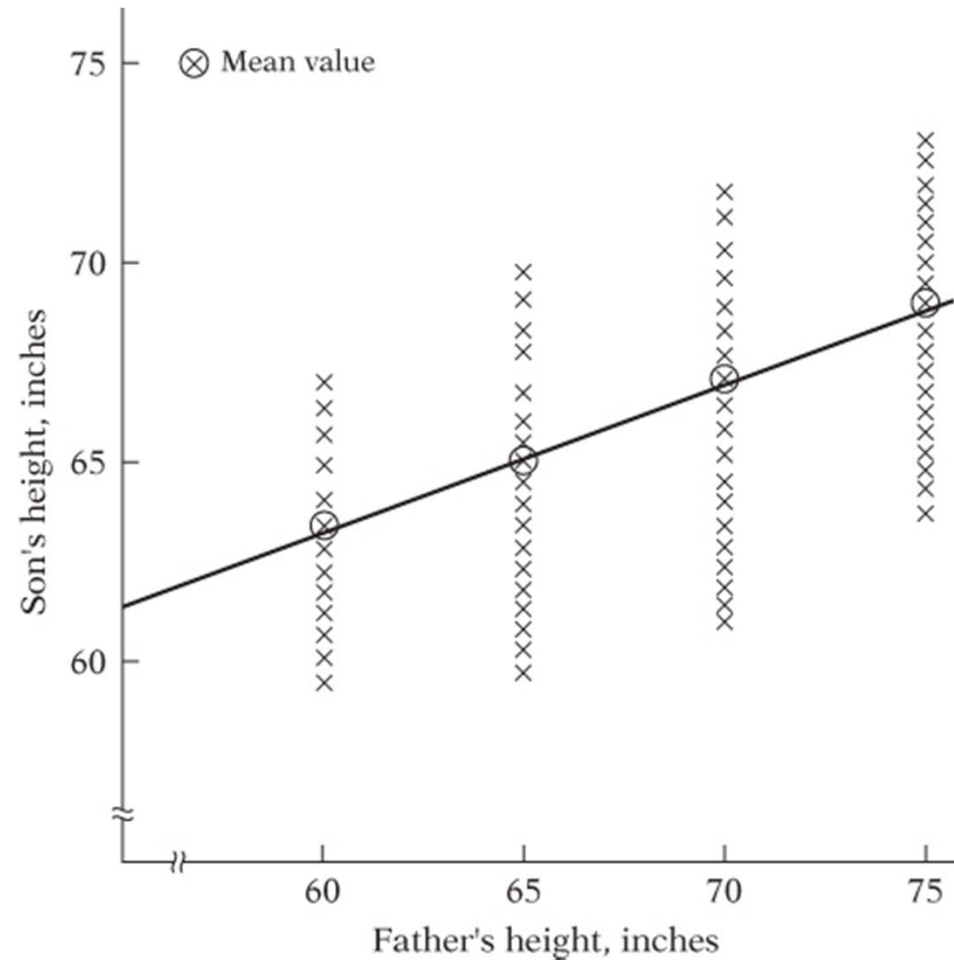


Regression

Regression analysis is concerned with the study of the dependence of one variable, the dependent variable, on one or more other variables, the explanatory variables, with a view to estimating the population mean of the former in terms of the known repeating sampling values of the latter.



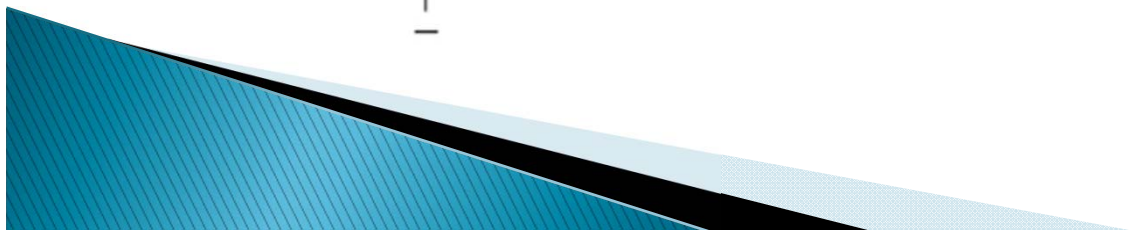
Hypothetical distribution of sons' heights corresponding to given heights of fathers



- ▶ This figure shows the distribution of heights of sons in a hypothetical population corresponding to the given or fixed values of father's height
- ▶ Despite the variability of the heights of sons for given value of fathers' height, the average height of sons generally increases as the height of the father increases
- ▶ Regression line shows how the average height of sons increases with the father's height



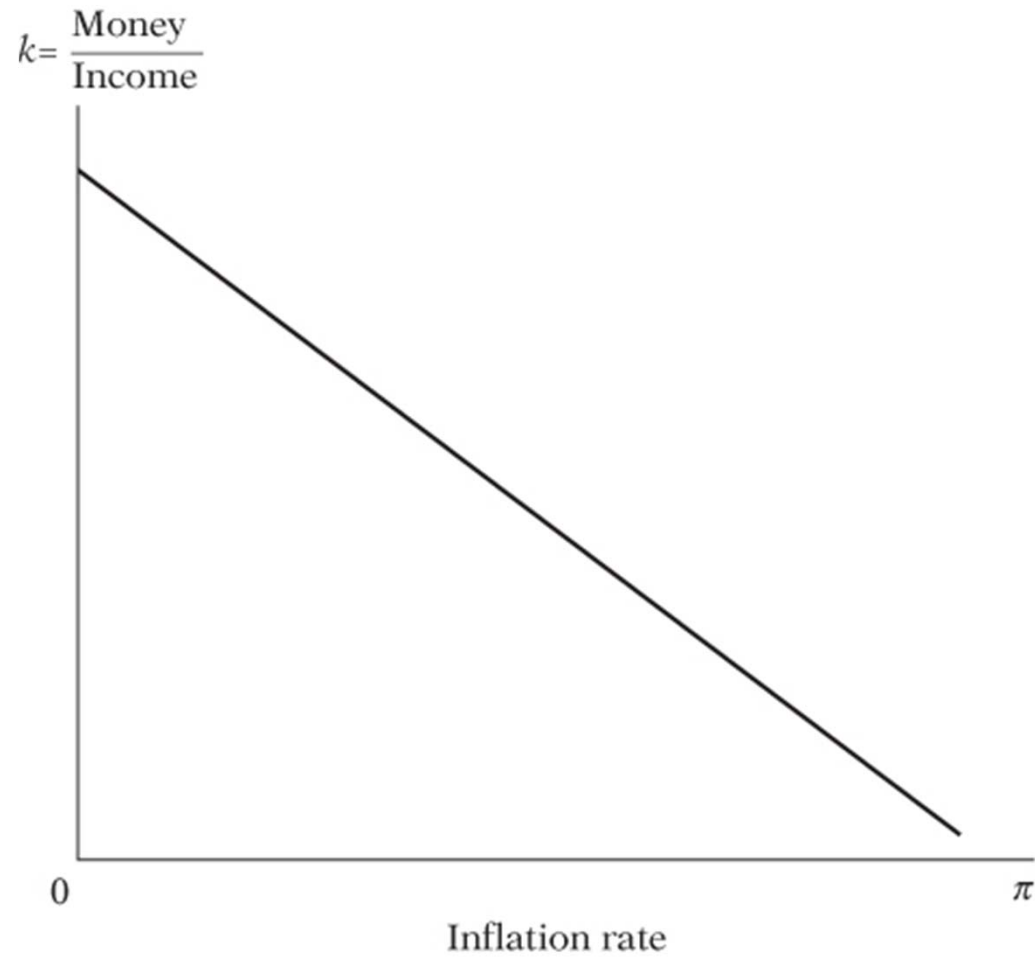
Hypothetical phillips curve



- ▶ Phillips curve relating changes in the money wages to the unemployment rate
- ▶ To predict the average change in money wages given a certain unemployment rate



Money holding in relation to the inflation rate



- ▶ The higher rate of inflation, the lower the proportion k of their income that people would want to hold in the form of money
- ▶ The slope of this relationship will enable the monetary economist to predict the amount of money, as a proportion of their income, that people would want to hold at various rates of inflation



Regression versus Causation

A statistical relationship in itself cannot logically imply causation

Regression analysis deals with the dependence of one variable on other variables, it does not necessarily imply causation



Ceteris paribus

- ▶ **Other (relevant) factors being equal**
- ▶ E.g. Analyzing consumer demand
 - We are interested in knowing the effect of changing the price of good on its quantity demanded, while **holding all other factors –such as income, price of other goods, and individual tastes-fixed**
- ▶ E.g. The effect of another week of job training on wages- **with all other components (education and experience) being equal**

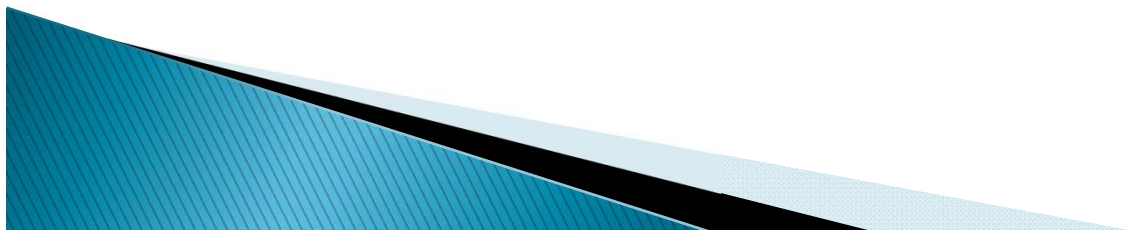


Correlation

Correlation analysis measures the strength or degree of linear association between two variables

Example

- ▶ The correlation between smoking and lung cancer
- ▶ The correlation between high school grades and college grades



Regression versus Correlation

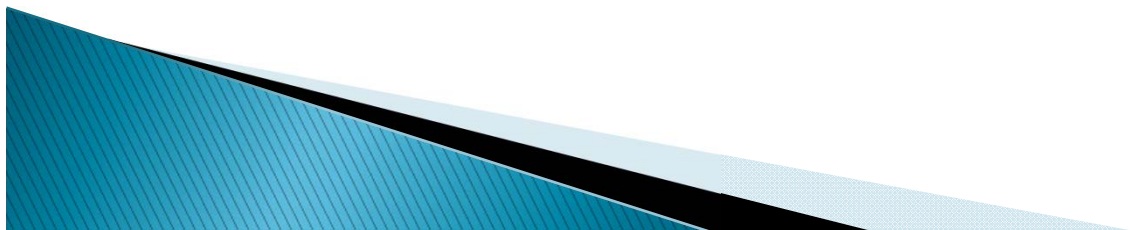
Regression	Correlation
<p>An asymmetry in the way the dependent and explanatory variables are treated</p> <p>Dependent variables are assumed to be statistical, random, or stochastic, to have a probability distribution</p> <p>Explanatory variables are assumed to have fixed values (in repeated sampling)</p>	<p>No distinction between the dependent and explanatory variable</p> <p>E.g. Scores on mathematics and statistics examination</p>

Terminology and Notation

Y	X
Dependent variable	Independent variable
Explained variable	Explanatory variable
Regressand	Regressor
Responses variable	Control variable
Outcome	Covariate

The nature of the Y variable

- ▶ Ratio scale
- ▶ Interval scale
- ▶ Ordinal scale
- ▶ Nominal scale



Ratio scale

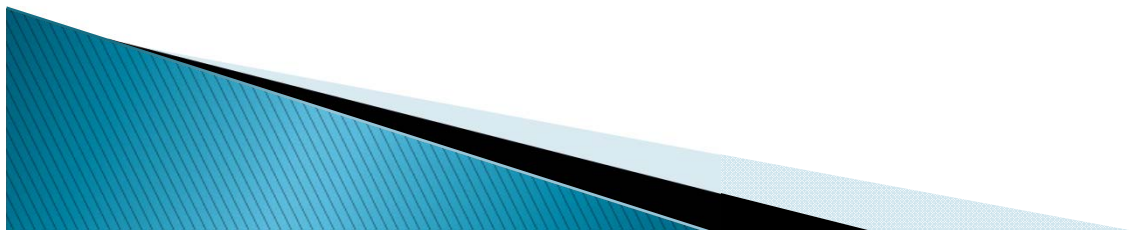
Three properties:

- ▶ Ratio of two variables
- ▶ Distance between two variables
- ▶ Ordering of variables

Most economic variables belong to this category

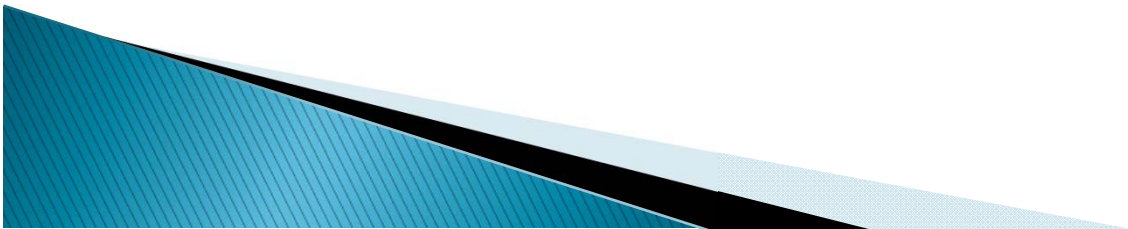
E.g. GDP is greater this year than last year

The ratio of GDP this year to the GDP last year is greater than or less than one



Interval scale

- ▶ Distance between two time periods



Ordinal scale

- ▶ Variables satisfy the ordering property of the ratio scale, but not the other two properties
- ▶ E.g. Grading systems
 - Income classification (low income, middle income, and high income)



Nominal scale

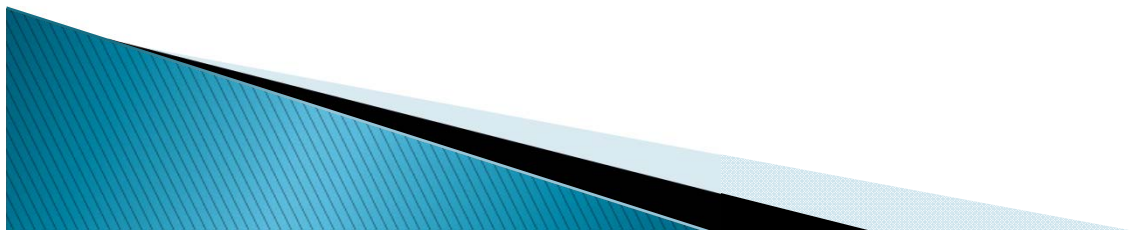
- ▶ Do not have any of the features of the ratio scale variables
- ▶ E.g. Gender
Marital status
Religious
- ▶ Such variables are often called dummy or categorical variables



Types of Data

Types of Data for empirical analysis:

1. Cross-Section
2. Time Series
3. Pooled
4. Panel, Longitudinal, or Micropanel Data

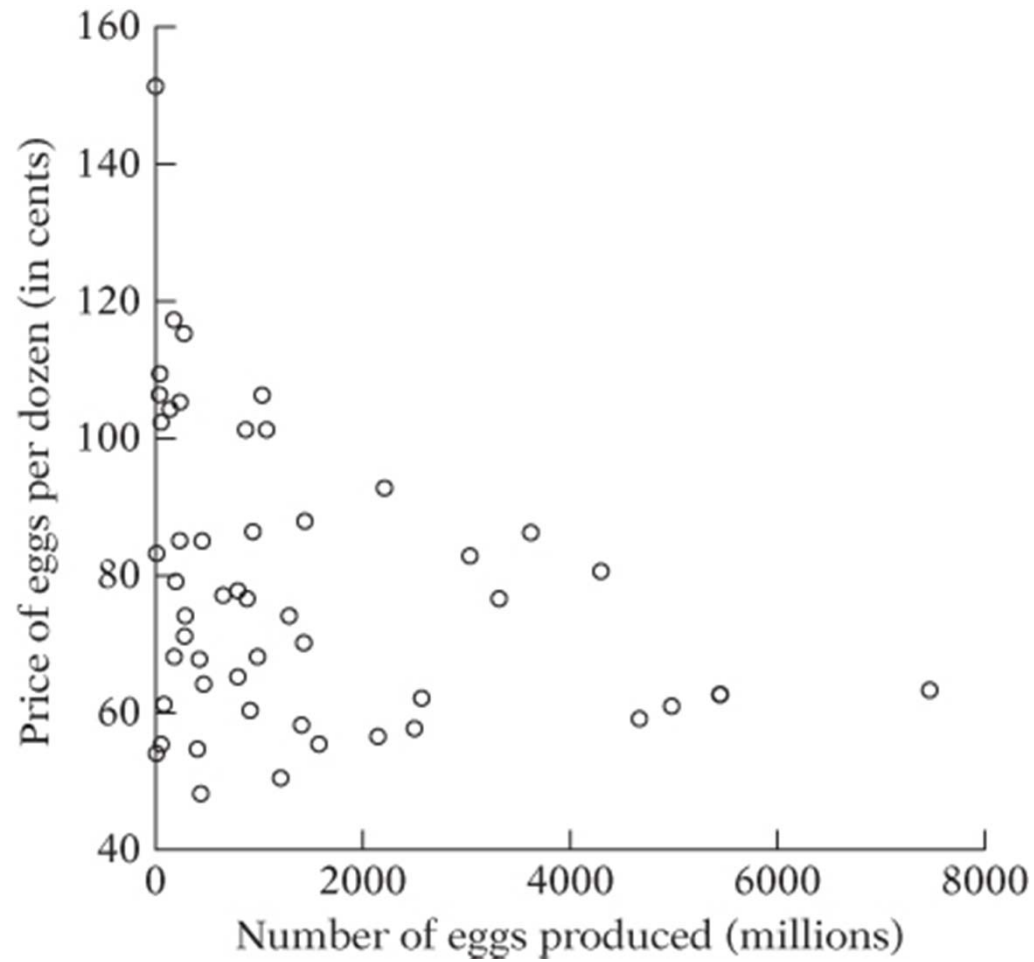


Cross Section Data

- ▶ Cross-sectional data: Data collected at the same point in time
- ▶ Consists of a sample of individuals, households, firms, cities, states, countries, or a variety of other units, taken at a given point in time.
- ▶ Often obtained by random sampling from the underlying population
- ▶ The analysis of cross-sectional data is closely aligned with the applied economics fields, such as labor economics, health economics, industrial organization, urban economics, and local public finance

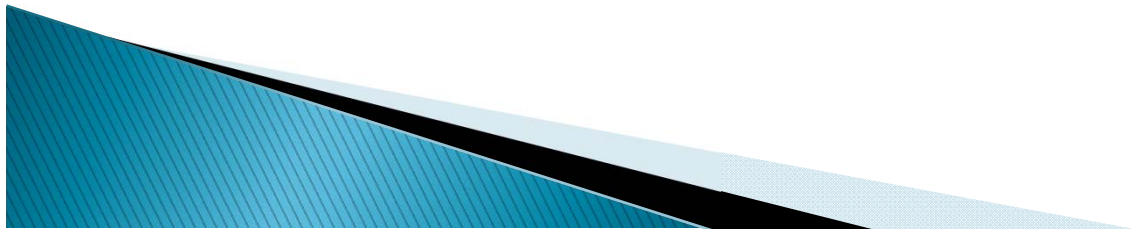


Relationship between eggs produced and prices, 1990

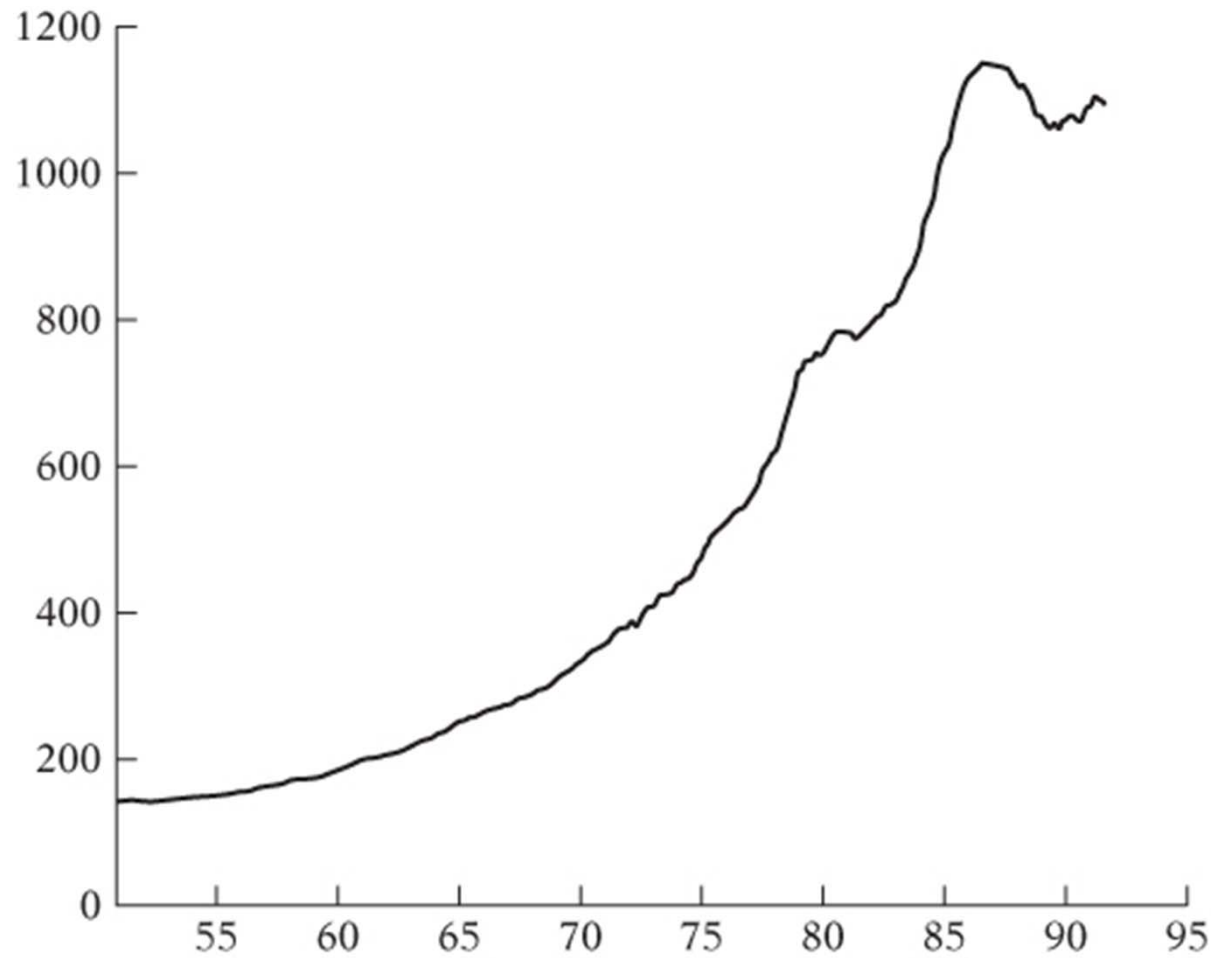


Time Series Data

- ▶ Time series data consists of observations on a variable or several variables over time
- ▶ Data may be collected at regular time intervals such as
 - Daily – Stock prices or Weather reports
 - Weekly - Money supply in the U.S. economy
 - Monthly – Unemployment rate or CPI
 - Quarterly - GDP
 - Annually – Government budgets

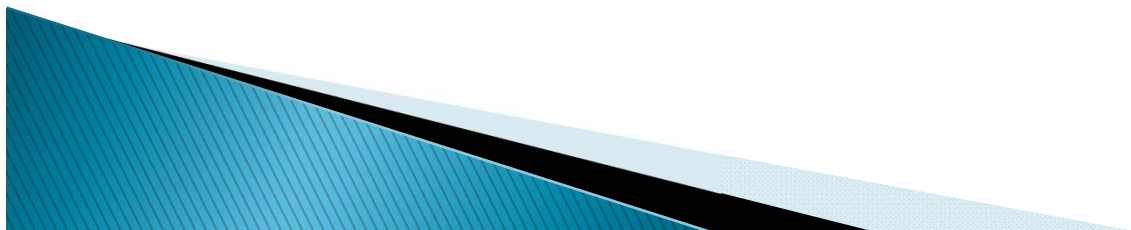


Money supply: U.S., Jan 1959-Sept 1999



Pooled Data

- ▶ Data are elements of both time series and cross-section data
 - Two cross-sectional household surveys are taken in the U.S., one in 1985 and one in 1990



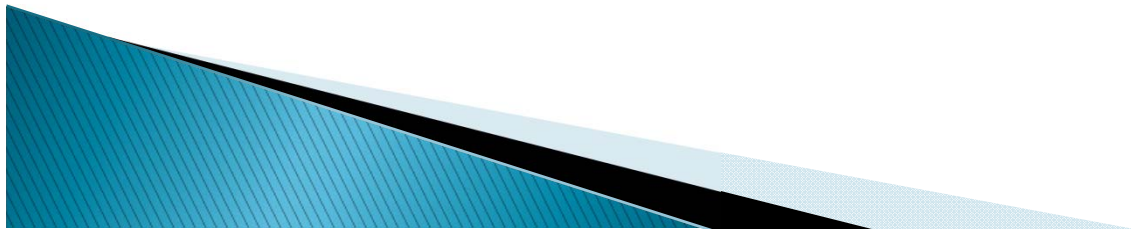
Panel or Longitudinal Data

- ▶ Consists of a time series for each cross-sectional member in the data set
- ▶ E.g. we have wage, education, and employment history for a set of individuals followed over a ten-year period
- ▶ Key feature of panel data that distinguishes them from a pooled cross section is that the same cross-sectional units are followed over a given time period



The Sources of Data

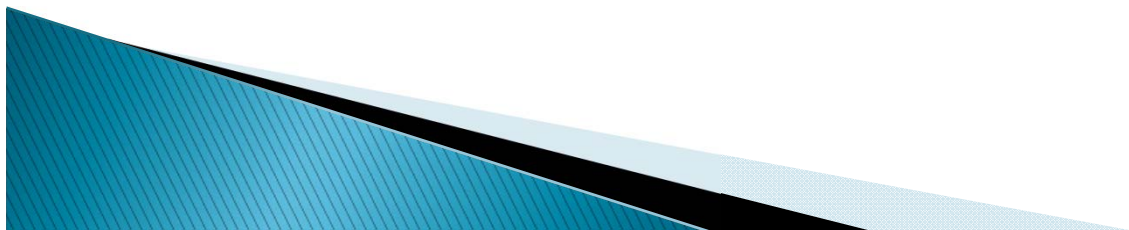
- ▶ The data used in empirical analysis may be collected by government agency, an international agency (e.g., IMF or the World Bank), a private organization (e.g., the Standard & Poor's Corporation)
- ▶ Experimental or Nonexperimental data



Web-Links

INTERNATIONAL ORGANIZATIONS

- ▶ [The Organization for Economic Cooperation and Development \(OECD\)](http://www.oecd.org)
(<http://www.oecd.org>).
The Organization for Economic Cooperation and Development (OECD) reports statistics on labor market conditions in many advanced economies.
- ▶ [IZA](http://www.iza.org) (<http://www.iza.org>).
IZA is a Bonn-based research institute that conducts labor research. Their discussion paper series provides up-to-date research on labor issues in many countries.
- ▶ IMF
(<http://www.imf.org/external/data.htm>)
- ▶ WORLD BANK
(<http://data.worldbank.org/>)



▶ **THAILAND MINISTRY OF LABOUR** (<http://www.mol.go.th/en>)

▶ **INTERNATIONAL LABOUR ORGANIZATION (ILO)**
(<http://www.ilo.org>)

▶ **BUREAU OF LABOR STATISTICS (BLS)**

The Bureau of Labor Statistics (BLS) is the government agency responsible for calculating the monthly unemployment rate as well as the Consumer Price Index. Their website contains a lot of information on many aspects of the U.S. labor market, as well as comparable international statistics. The BLS also publishes a detailed description of how it defines and measures the concepts of labor force and unemployment. The BLS is officially a part of the U.S. Department of Labor, which also collects detailed information on various aspects of the state-run unemployment insurance program.

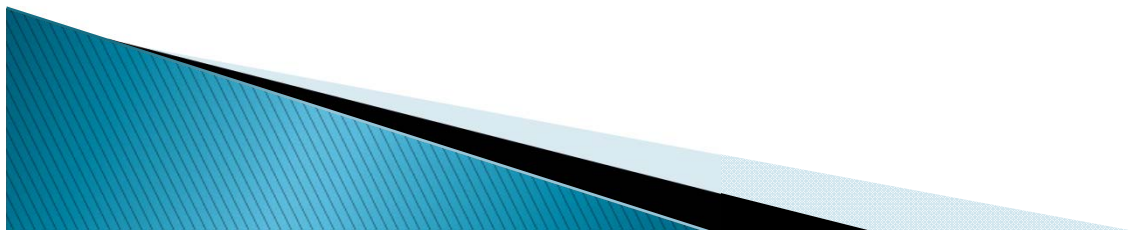


▶ **NATIONAL STATISTICAL OFFICE THAILAND**

(<http://www.nso.go.th>)

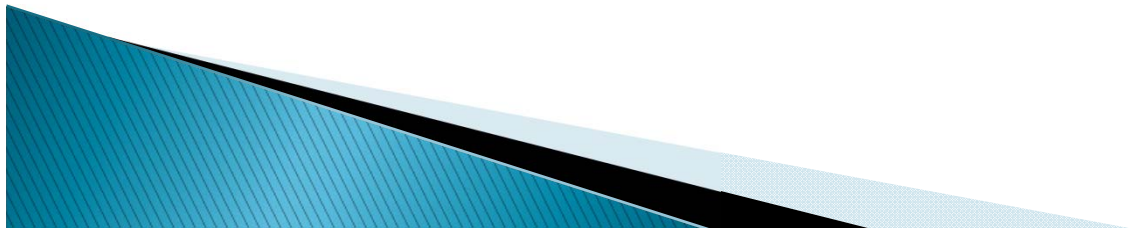
▶ **BUREAU OF THE CENSUS**

The Bureau of the Census reports detailed demographic and labor market information, including historical data by race and gender, mobility/migration patterns, and health insurance coverage. The Bureau of the Census also publishes annually the Statistical Abstract of the United States , which contains data on a wide variety of topics including many aspects of the U.S. economy.



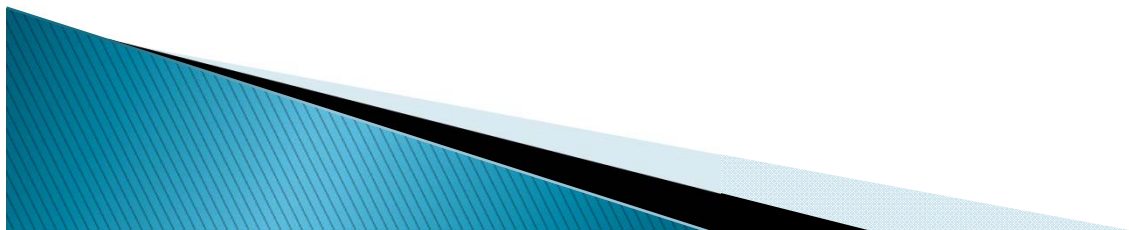
IMMIGRATION

- ▶ [Bureau of Citizenship and Immigration Services \(BCIS\)](http://www.uscis.gov/graphics/index.htm) (<http://www.uscis.gov/graphics/index.htm>). The website of the Bureau of Citizenship and Immigration Services (BCIS) contains information on U.S. immigration policy.
- ▶ [Department of Homeland Security](http://www.dhs.gov/ximgtn/statistics/) (<http://www.dhs.gov/ximgtn/statistics/>). The Department of Homeland Security reports detailed statistics on immigration to the United States.
- ▶ [International Trade Administration](http://www.ita.doc.gov) (<http://www.ita.doc.gov>). The International Trade Administration publishes detailed information on trade patterns and regulations.
- ▶ [Citizen and Immigration Canada](http://www.cic.gc.ca/english/index.asp) (<http://www.cic.gc.ca/english/index.asp>) The website of Citizen and Immigration Canada has the "test" that allows a potential applicant to determine if he or she qualifies for a visa.
- ▶ [Australia's Department of Immigration and Multicultural and Indigenous Affairs](http://www.immi.gov.au/) (<http://www.immi.gov.au/>) The website of Australia 's Department of Immigration and Multicultural and Indigenous Affairs has the similar test required by Australian authorities.



EDUCATION

- ▶ [The U.S. Department of Education](http://www.ed.gov/index.jsp) (<http://www.ed.gov/index.jsp>). The U.S. Department of Education's website provides an essential introduction to many of the education programs in the United States.
- ▶ [General Education Development \(GED\)](http://www.acenet.edu//AM/Template.cfm?Section=Home) (<http://www.acenet.edu//AM/Template.cfm?Section=Home>). The American Council on Education gives useful information to prospective takers of the General Education Development (GED) exam.



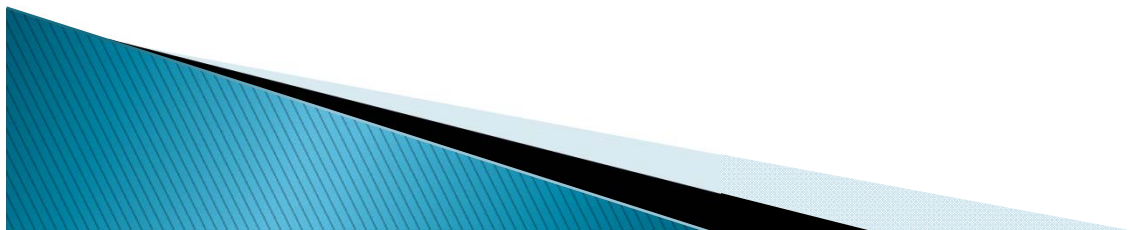
POVERTY / MINIMUM WAGES / DEVELOPMENT

- ▶ [The Green Book](http://waysandmeans.house.gov/Documents.asp?section-5813) (<http://waysandmeans.house.gov/Documents.asp?section-5813>). The *Green Book*, also known as the *Overview of Entitlement Programs*, is published by the Committee on Ways and Means, U.S. House of Representatives. Each chapter describes a particular type of public assistance program.
- ▶ [The Employment Policies Institute](http://www.epionline.org) (<http://www.epionline.org>)
- ▶ [The United Nations Development Programme](http://www.undp.org/) (<http://www.undp.org/>). The United Nations Development Programme maintains an extensive database describing income inequality in many countries.
- ▶ [OECD](http://www.oecd.org/) (<http://www.oecd.org/>)
The OECD reports unemployment statistics for many advanced economies and frequently publishes reports comparing the unemployment situation in different countries.
- ▶ [Population Resource Center](http://www.prcdc.org/) (<http://www.prcdc.org/>)
Many organizations are concerned with the politically charged issues of population growth and control. The Population Resource Center is an organization that limits itself to promoting the use of accurate data for objective analysis.



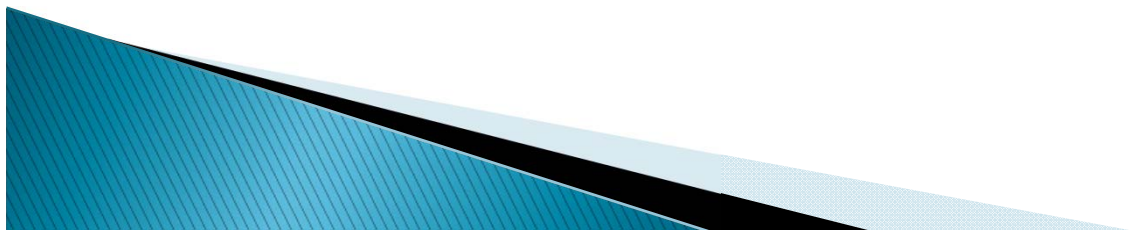
NON-PROFIT ORGANIZATIONS

- ▶ [The National Bureau of Economic Research \(NBER\)](http://www.nber.org)
(<http://www.nber.org>). The National Bureau of Economic Research (NBER) publishes a working paper series that represents the frontier of empirical research in economics. Their website also contains a number of widely used data sets. The working papers and data can be accessed and downloaded by students and faculty at many universities.



The Accuracy of Data

The researcher should always keep in mind that the results of research are only as good as the quality of the data



Source

Gujarati, D.N. (2009) Basic Econometrics. 5th ed.
Singapore, McGraw-Hill.

