

Week 1: Outline

- Motivation
- How does econometrics differ from statistics?
 - ⇒ Some early alerts!
- From an economic to an econometric model
- Data
- Bivariate regression model

What is Econometrics?

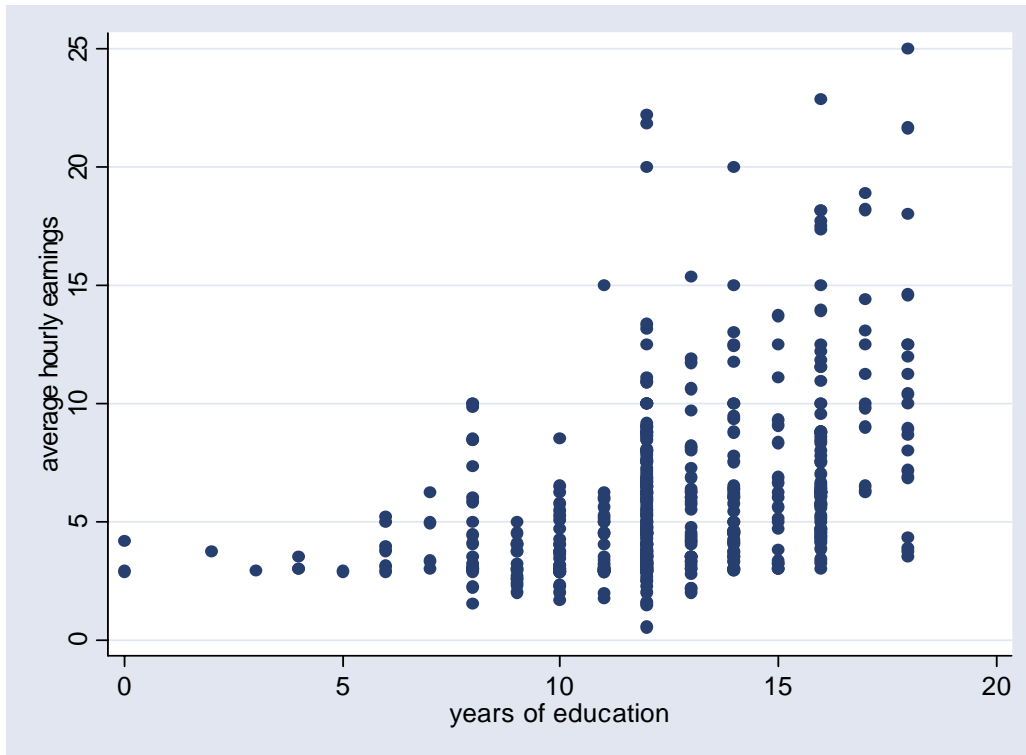
Use of statistical methods for:

- Estimating economic relationship
 - E.g. Each additional year of education increases average hourly wage by 54 cents in the 70s in the US.
 - E.g. 50% of agricultural output decline during early transition in Eastern Europe can be attributed to price declines.
- Testing economic theories
 - E.g. Is there a causal relationship between institutions/governance and growth?
 - E.g. Is there a causal relationship between the education of the mother and education of her children?
- Impact analysis
 - E.g. What is the impact of an agricultural extension program on yields in Indonesia?
 - E.g. What is the impact of conditional cash transfers on school enrollment of children in Mexico?

Why?

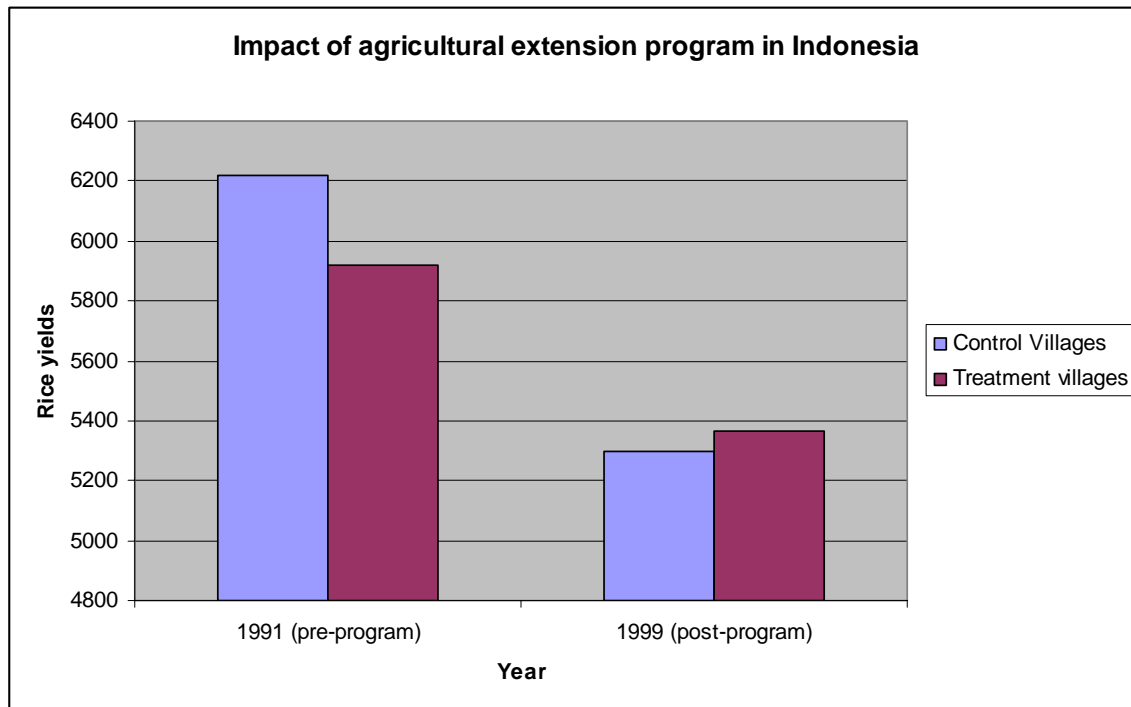
- Forecasting
- Predicting of behavior economic agents – advice on future policies
- Evaluating existing/past policies

Graph: Returns to education



Source: Wooldridge, WAGE1.dta (data from 1976 Current Population Survey)

Graph: Impact analysis of Farmer Field Schools in Indonesia



Source: Feder, Murgai and Quizon (2004).

How does Econometrics Differ from Statistics?

=> *Observational, non-experimental data*

Natural sciences: often experimental data

⇒ Causal inference automatic

Social sciences:

- Sometimes experiments
Random assignment of interventions (~policy evaluation)

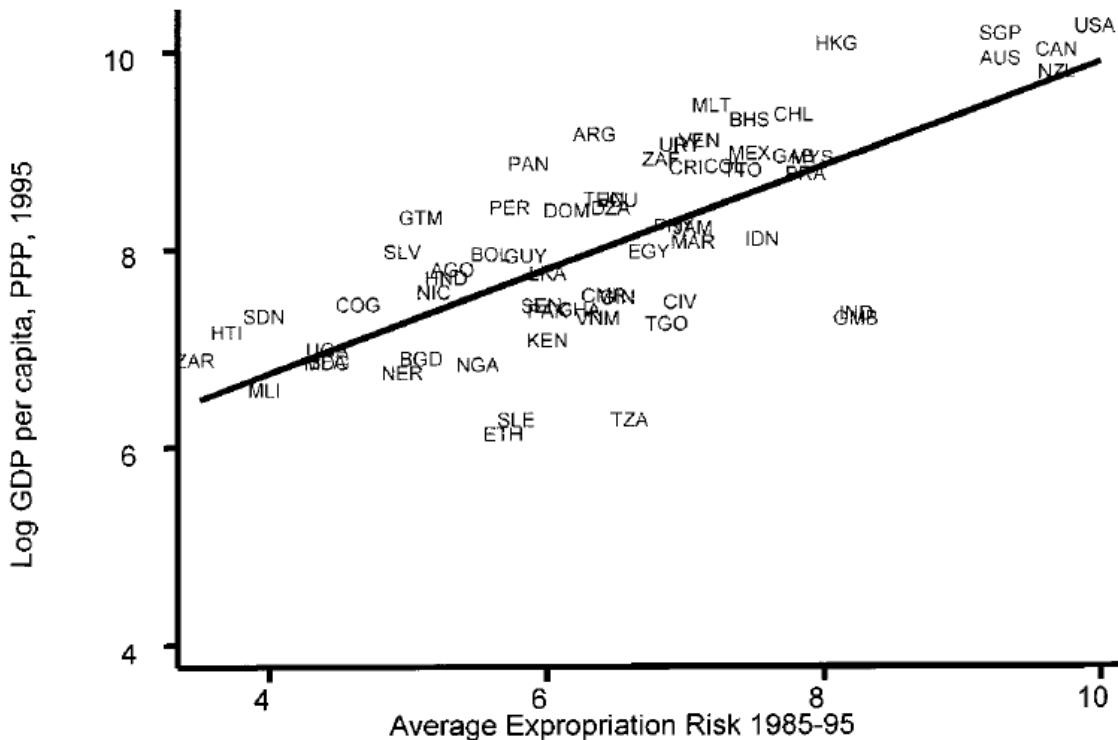
- Most of the time; can't do experiments

=> Need for a more "careful" use of the data

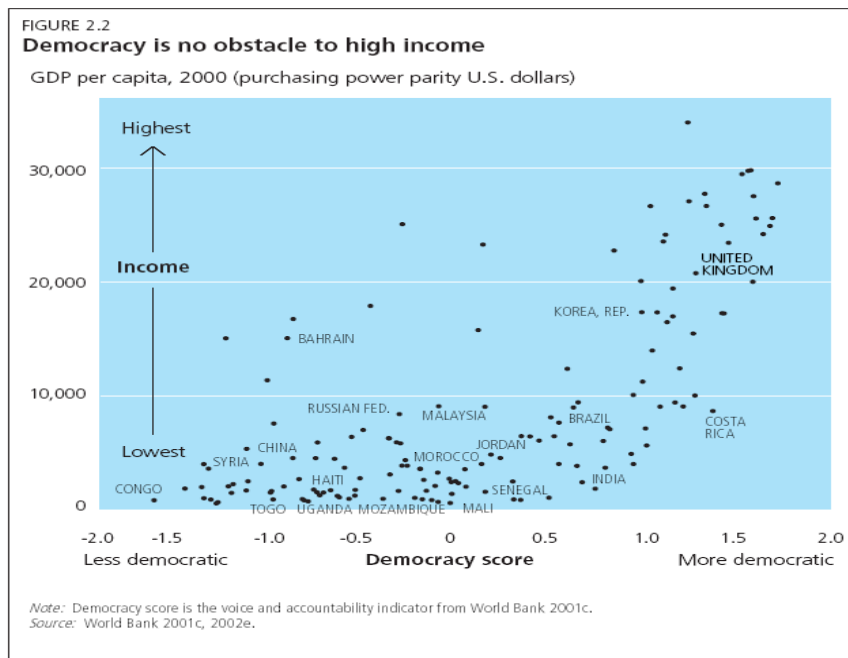
- Causal inference becomes difficult
E.g. be aware of
 - Spurious correlation
 - Reverse causality:
 - Always watch out for *identification* problems
- Sample selection matters

Graph : OLS Relationship between expropriation risk and income – spurious correlation?

Source: (Acemoglu, Johnson, and Robinson, *American Economic Review*, 2001)



Graph: Example of possible reverse causality



Source: UNDP, Human Development Report 2002.

Econometric Model (1) :

Example 1: relationship between GPA and income

Economic model: $Y = f(GPA)$

Econometric model: $Y_i = \beta_0 + \beta_1 * GPA_i + u_i$

With:

- GPA: Independent (or right-hand side) variable
- Y: Dependent (or left-hand side) variable
- β : Parameter: to be estimated!
- u: Error term
 - Captures the unobserved part of the relationship
 - We will need to make assumptions about these

Econometric Model (2) :

Example 2:

Economic model: $Y = f(GPA, age)$

Econometric model: $Y_i = \beta_0 + \beta_1 * GPA_i + \beta_2 * age_i + u_i$

Hypothesis: $\beta_1 > 0, \beta_2 > 0$

Interpretation of parameters: “Ceteris paribus”
i.e. “keeping everything else constant”

The Data:

- Cross-section, time series or panel
 - Cross-section: sample of many entities (individuals, firms, ..., countries) at one point in time
 - Time series: Observations on one entity at many points in time
 - Panel: Observations on several entities at several points in time.
- Existing records or targeted surveys
- Different levels of aggregation
- Very often: proxies!