

## Why study Econometrics?

- research tools used in accounting, finance, marketing and management
- used by other social scientists  
history, sociology, polj sci.
- As econ, as business,

Differentiates you from others. +

skill marketable.

Makes econ much more practical.

Necessary for grad work -

Economic Theory tends to be a rather abstract science. For example, economists would speculate that, based on the Theory of Demand

$$Q^d = f(P, P_s, P_c, \text{Income})$$

$Q^d$  = Qty demanded

$P$  = Price of the good

$P_s, c$  Price of subs and complements

Income.

What economic Theory lacks, in many instances, is any ability to quantify the sizes these effects or the specification of  $f(\cdot)$  or the precise way that  $P, P_s, P_c, \text{Inc}$  affect  $Q^d$ .

Econometrics is ~~the~~ a sub discipline of economics concerned with quantifying ~~re~~ economic relationships. It also falls on econometrics to distinguish one theory from another based on evidence - hypothesis testing.

So, the process of econometric model building starts with economic theory, which helps to identify relevant variables and makes predictions about the direction of the effect of one on another or of the direction of causality.

Then, ~~the economist~~ the econometrician builder

then samples are collected data on the variables of interest. At that point, a statistical model is formed that is consistent with the S.O.P. of the sample - or stat - population of interest.

- 1) specifies a functional form
- 2) expresses the relationship as a function of parameters
- 3) recognises and specifies the nature of the unobservable error associated with your model.

$$\ln(Y^d) = \beta_1 + \beta_2 \ln(P) + \beta_3 \ln(P_s) + \dots + \epsilon_c$$

For the ex-

The errors include unobservable random events, omitted variables, measurement errors associated with  $\ln G^d$ .

So, after specifying how  $e$ 's behave  
(some variance, ~~and~~ random sample?)  
estimate  $B$ .

Evaluate the adequacy of model  
If adequate, then make inference  
If not, figure out what went wrong  
and start over.

## Where Do we get Data?

In general, in our discipline we are or have to use data that are not outcomes of experiments. This is what makes econometrics a little more specialized than regular statistics, (although every thing we do can be found there.)

### Experiment

$$g^d = f(P, P_s, P_c, \text{Income})$$

If on experimental world we'd be able to set Price and Income and observe what <sup>is</sup> sold how many houses are sold.

By varying

$$P = \$25,000$$

$$i = \$42,000$$

$$P_s = \$25,000$$

$$P_c = \$1.35 \quad \text{gasoline}$$

Before experiment,  $g^d$  is unknown

$\Rightarrow$  Random variable.

(a) The outcome

$$g^d = \text{systematic part} + \text{Random unobsr. part}$$

our goal is to measure the  
Systematic part

$$\underline{g^d = \beta_1 + \beta_2 P + \beta_3 P_s + \beta_4 P_c + \beta_5 I_{nc}} + c$$

In experiment, we could  
hold  $P_c$ ,  $i$ , and  $P_s$  constant,  
raise Price & rent <sup>by \$1,000</sup> and  
see how much money we sell.

$$\Delta i, \Delta P_c, \Delta P_s$$

The advantage here is that we can change the control R.H.S. Vais to anything we want. Helps economize on H of experiments we do, since there are ways to choose these that improve how precisely we can measure their effects.

### Nonexperimental

Here, the researcher is passive observer. — takes data as they are generated by the system under study.

→  $P, P_c, P_s, F_{nc}$ . ~~may~~  
not all changing at the same time.

If If  $P$  and  $P_s$  both go up by \$1,000 and we measure  $f$  how can be tell if how much of  $f$  depends on  $P$  and how much of  $P_s$ ?

So, data may not move  
independently of one another  
and precise measurement  
could be hard.

also, no control over  $\epsilon$ .

That can complicate things as well  
— as we shall see later in  
course.

time series - data collected at  
discrete intervals — Q GDP  
 $1960 - 1999 \quad 40 \times 4 = 160$

cross section - data collected over  
sample units at 1 point in  
time

Avg tax rate in each state  
for 1999. (50 obs)

panel data -

time series are collected  
on several individuals

Avg tax rate in each state  
1960-1999. (40 x 50 obs)

# Inference

9

To learn

- EST Params
- make prediction
- test hypothesis.

Quantitative - outcomes are expressed as #s

Qualitative - outcomes ~~are~~  
are not naturally expressed  
as # - Possessing (or not)  
some quality.

Gender

Race

Political Party

etc.

Research Factual.

(1) Problem

(2) Economic Theory

(3) Statistical Model

(4) Sample

(5) Estimate

(6) Model diagnostics

(7) Implications & conclusions

## Basic Probability Concepts

Economic variables are by nature random. We don't know what their values are until we observe them. Probability is one way of expressing uncertainty about economic events and outcomes so we need to spend a little time talking about probability.

### Controlled Experiments

The conditions surrounding the experiment are under the control of the researcher

seed corn =  $\{$  identical plots, equal fertility,  
some amt of fertilizer, pesticide, etc)

measure yield, # ears/stalk, moisture, sugar content, etc.

- ① Experiment is can be repeated under the same circumstances by others
- ② It can be repeated using different settings to measure their effects on the random variable of interest.

The experimental outcome is measured using one or more characteristics of yield

$$X = \# \text{ bushels/acre} \quad Y = \# \text{ ears/stalk}$$

These are Random Variables - a variable whose value is unknown until it is observed. R.V.  $X$  its value  $x$ .

### Uncontrolled Experiments

GDP, interest rates, of beer demand, etc are also unknown until observed. They are not the product of controlled experimentation.

- (1) observed rather than collected
- (2) usually collected for purposes other than economic analysis.

One goal is to adopt stat models that are consistent with economically relevant variables the way the data are generated. Another is to obtain economically relevant data whenever possible.

## Randomized Controlled Experiment

- Basically a way to measure CAUSAL relationships.

take two groups. Individuals are randomly assigned to each group.

- Control Group
- Treatment Group.

The individuals in the control group receive no treatment and those in the other group do.

Any differences in outcomes are consequences of treatment

- Why? Because by randomly assigning or determining who receives treatment - all other factors that determine outcome

should average out between  
the ind. in each group.