AGRODEP

Household survey data course

Dakar, 8-10 October 2012

Introduction to microeconometrics

IFPRI®

HarvestChoice

BETTER CHOICES, BETTER LIVES

THE WORLD BANK

LSMS

Living Standards Measurement Study

Overview

- Purpose and scope of econometrics
- Type of data
- Causality
- Type of models
- Simple linear regression model: theory and application

Purpose and scope of econometrics

- Estimating economic relationships
- Testing economic theories
- Evaluating programs/policy
- Forecasting

....in general

finding association among phenomena

Type of data

- Cross-section → 
- Time series → 
- Cross-section (different unit) over time: pooled cross-section → 
- Cross-section (same unit) over time: panel (longitudinal) → 

Non-experimental vs/ experimental

Causality

All econometric applications must be extremely cautious on:

- Problem of *other factors being equal*
- Problem of spurious correlation

Type of models

- Linear regression (simple, multiple): OLS, IV and 2SLS
- Simultaneous equation
- Binary regression: again OLS (LPM)
- Limited dependent variable: binary response probit and logit, multinomial probit and logit, tobit and censored regression, count (poisson), sample selection (heckit)
- Advanced time series: IDL, geometric (Koyck) distributed lag, RDL, ECM

Simple linear regression model /1

The basic model (simplest case) is:

$y=\beta_0+\beta_1 x + u \qquad (1)$

where y is the *dependent* variable

x is the *independent* variable

u is the *error term* (or unobserved *disturbance*)

What do β_0 and β_1 represent?

Simple linear regression model /2

Let's assume that $\Delta u=0$ (holding other factors in u unchanged):

$\Delta y=\beta_1 \cdot \Delta x$

Then, β_1 is the **slope** parameter (linear relationship between x and y) and β_0 is the **intercept** parameter.

β_1 is the most interesting parameter in econometrics

Simple linear regression model /3

Let's go back on *(1).* Now let's assume

$E(u)=0$

how far is it from reality?

Is this condition enough for *(1)* to be valid?

What about u was *not linearly* correlated with x (x^2, log x,...)?

$E(u|x)=E(u)=0$ *(2)*

This is the **zero conditional mean assumption**

Simple linear regression model /4

Taking the E(.) of *(1)* (cond. on x) and using *(2):*

$E(y|x) = \beta_0 + \beta_1 x$

population regression function (**PRF**) linear in x.

→Graph

How do we calculate β_0 *and* β_1*?* ?