Introduction to Applied Research in Economics

Kamiljon T. Akramov, Ph.D.
IFPRI, Washington, DC, USA

Training Course on Applied Econometric Analysis

June 1, 2015, WIUT, Tashkent, Uzbekistan

Why do we need applied economic research?

Introduction

- Excitement about applied economic research comes from the opportunity to learn about cause and effect in real world
- Economic theory suggests important relationships, often with policy implications, but virtually never provides quantitative magnitudes of causal effects
- Applied economics is more vulnerable than physical sciences to models whose validity never will be clear, because necessity for approximation is much stronger
- Nevertheless, economics has an important quantitative side, which cannot be escaped
- The important questions of the day are our questions

Introduction (cont.)

- Will loose monetary policy spark economic growth or just fan the fires of inflation?
- What is the effect of a 1 percentage point increase in broad money on inflation?
- What is the effect of a 1 percentage point increase in interest rates on output growth?
- Will mandatory health insurance really make people healthier?
- How does an additional year of education change earnings?
- What is the effect of farm size on agricultural productivity?
- How does agricultural diversity impact nutritional outcomes?

Introduction (cont.)

- Economists' use of data and tools to answer cause-and-effect questions constitutes the field of applied econometrics
- Tools of applied econometrics are disciplined data analysis combined with statistical inference
- We are after truth, but truth is not revealed in full, and messages the data transmit requires interpretation
  - Examples
- Comparisons made under ceteris paribus conditions may have a causal interpretation
- Ceteris paribus comparisons are difficult to engineer

Objectives of this course

- The main objective of the course is to strengthen the capacity of young researchers to conduct empirical research
- The emphasis of the course will be on empirical applications with a focus on detection of causality in data
- Ideally, we would like to have an experiment
- But most of the time we only have observational (non-experimental) data

In this course you will learn

- “Furious Five” of modern applied econometric research
  - Regression, RCT, Regression discontinuity, Instrumental variables, Difference-in-differences
- Other methods for estimating causal effects using observational data
  - Fixed effects, Propensity score matching, etc.
- Productive use of these techniques requires a solid conceptual foundation and a good understanding of the machinery of statistical inference
- Focus on applications – theory is used only as needed to understand the “why” of methods
- Learn to evaluate the empirical (regression) analysis of others – this means you will be able to read/understand empirical economics papers
- Get some hands-on experience with econometric analysis using real world data

Hypotheses in empirical research

- Construction of research hypotheses is an important step in applied economics research
- Hypotheses argue that one phenomenon or behavior causes or is associated with another phenomenon or behavior
  - These phenomena are called constructs
- Various sources of support routinely used to develop hypotheses
  - Theory and logical analysis (intuition)
  - Past studies: authority and consensus
  - Real life experiences and observations

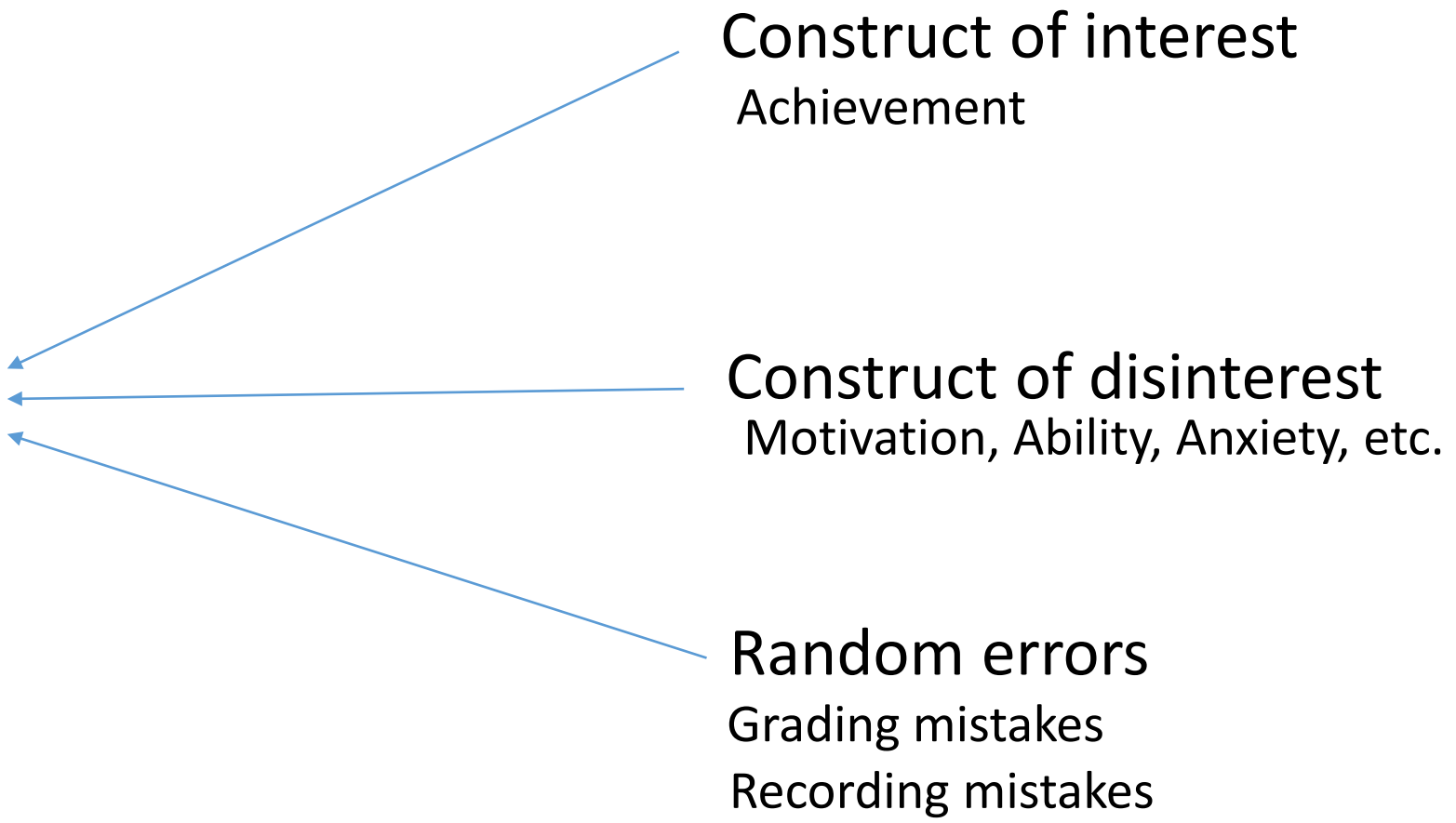
Framework for assessing empirical research

- Construct validity: to what extent are the constructs of interest successfully operationalized in the research?
- Internal validity
  - To what extent does the research design permit us to reach causal conclusions about the effect of the independent variable on the dependent variable?
- External validity
  - To what extent can we generalize from our sample and setting to the populations and settings specified in research hypothesis?

Maximizing construct validity

- Constructs are the abstractions and any one construct can be measured in different ways because there are a variety of concrete representations of any abstract idea
- Variables are partial representations of constructs, and we work with them because they are measurable
- Operational definitions specify how to measure a variable
- Reliability of a measure is defined as the extent to which it’s free from random error components
- Validity is defined as the extent to which the measure is free from systematic errors

Three components of variable



Stock, J. and M. Watson (2010)

Maximizing internal validity

- Threats to the internal validity of regression studies
  - Omitted variable basis
  - Misspecification or incorrect functional form
  - Measurement error
  - Sample selection bias
  - Simultaneous causality bias
  - Unobserved heterogeneity
- All of these imply that conditional mean independence assumption fails – in which case OLS is biased and inconsistent

Maximizing external validity

- A population is the aggregate of all of the cases that conform to some designated set of specifications
  - All the people residing in a given country, all households residing in a given state, etc.
- A stratum may be defined by one or more specifications that divide the population into mutually exclusive segments
- Nonprobability versus probability sampling
- Probability sampling
  - Simple random sampling gives each element in the population an equal chance of being selected
  - Stratified random sampling

Better research design helps to mitigate threats to validity

- Edward Leamer's (1983) reflection on the state of empirical work in economics: "take the con out of econometrics," "hardly anyone takes data analysis seriously"
- Empirical economics has experienced credibility revolution, with a consequent increase in policy relevance and scientific impact
- Randomized controlled trials (RCTs)
  - Treatment and comparison groups
- RCT designs
  - Randomized two-group design
  - Before-after two group design
  - Solomon four-group design
- Strengths and weaknesses of RCTs
  - Experimental artifacts
  - External validity issues

Quasi-experimental designs

- Difference-in-difference methods
  - Card and Krueger (2000) study the effects of minimum wages on employment
- Instrumental variables
  - Relevance
  - Exogeneity
- Regression discontinuity
  - Exploits precise knowledge of the rules determining treatment
  - Angrist and Lavy (1999) study of the effects of class size on achievement
- Matching methods

Data: sources and types

- Experimental versus observational data
- Cross-sectional data: data on different entities for a single period of time
- Time-series data: data on a single entity collected at multiple time periods
- Panel or longitudinal data: data for multiple entities in which each entity is observed at two or more time periods

Questions?

- What is the policy question?
- What is the causal relationship of interest?
- What is the dependent variable and how is it measured?
- What is (are) the key independent variable(s)?
- What is the data source?
- What is the identification strategy?
- What is the mode of statistical inference?
- What are the main findings?

Future readings

- Chetty, R. Yes, Economics Is a Science. New York Times, Oct. 20, 2013. http://www.nytimes.com/2013/10/21/opinion/yes-economics-is-a-science.html?_r=0
- Angrist, J. D. and J. S. Pischke. 2010. The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. Journal of Economic Perspectives, vol. 24, No. 2, pp. 3-30.
- Angrist, J. D. and J. S. Pischke. 2009. Mostly Harmless Econometrics: An Empiricists's Companion. Princeton University Press. Chapter 1.
- Angrist, J. D. and J. S. Pischke. 2015. Mastering Metrics: The Path from Cause to Effect. Introduction

Thank you and good luck