L15: Policy Evaluation in Development Economics: A Primer on Statistical Methodology

Dilip Mookherjee

Ec320 Lecture 15, Boston University

Oct 28 2014
Introduction

Start second half of the course: focus on development policy in specific areas

In the first half, we took a macro view to identify determinants of development success:

- savings/investment rates (financial development, investment climate, public infrastructure)
- human capital (education and health) investments
- lowering population growth (fertility) rates
- raising productivity (e.g., reducing misallocations, raising efficiency/technical progress, improving governance)
- enhancing access of poor to assets, skills and finance
Development Policy: Key Questions

- How exactly can these objectives be secured?
- What kinds of policies will be effective?
- How should governments/aid agencies/nonprofits allocate scarce budgetary resources across different development programs: weigh benefits against costs
Policy Questions

- Arise at various layers, e.g.:
  - Macro: how much should be spent on education
  - Micro: spending on different items in state schools: school construction, teacher salaries, curriculum support, student loans/scholarships
  - Regulatory/institutional changes: allow private schools? how to regulate private schools? monitor teacher performance in state schools? provide citizens choice of different schools?
Evaluating Policies: What Works?

- Rational evaluation procedures require assessment of effectiveness of different policy initiatives: to what extent do they help improve enrollment rates (quantity), or student learning (quality)?
- Yet most governments/aid organizations evaluate policy programs on the basis of how much they spend, not how effective they are in terms of ultimate goals.

- Evaluation criteria: some broad macro numbers (trends..), anecdotes, gut feeling...
- B/Millions get spent without any evaluation of effectiveness
- Problem of ‘self-serving bias of development practitioners’ (Mullainathan)
The Movement for Evidence-Based Evaluation of Policies

- In the last 10-15 years, there has been a big movement among development economists to use credible quantitative evidence on policy effectiveness.
- Use of refined statistical procedures for analysis of large scale micro-data data concerning past policy experience and outcomes.
- Natural and controlled field experiments to overcome statistical biases in estimates.
- Many governments, aid organizations and aid donors are now incorporating evidence-based evaluation procedures of policies.
Example of Policy Questions

- **Example 1:** How much would one additional year of schooling \((x)\) enable young adults to increase their earnings \((y)\) by?
- **Example 2:** What would the impact of reducing class-size (by \(x\)) be on student learning outcomes \(y\) (e.g., measured by test scores)?
Key Conceptual Issues

- Usually try to answer these questions by studying past experience, with past data on $x$ and $y$, and estimating how correlated they are.
- Cross-sectional evidence: compare across countries, across regions, across individuals.
- Time-series/longitudinal evidence: take a given unit of observation (country/region/person) observed at different points of time.
- **Key problem:** how relevant is the observed correlation in past data in making a prediction of what would happen to $y$ if $x$ were to be increased in future.
Regression Equation: Estimating Rate of Return to Education

- Cross-section across countries/regions/groups/individuals denoted $i$: estimate $\beta$ in
  \[ \log W_i = \alpha + \beta S_i + e_i \]

- Time-series for given unit across time $t$:
  \[ \log W_t = \alpha + \beta S_t + e_t \]

- Here $e_i$, $e_t$ are error (unexplained residual) terms
Spurious Correlations

- Yule’s example: strong positive correlation between stork population in Norway and number of babies born, 1870-1900 data
- Hendry’s example: strong positive correlation between annual cost of living index in UK and cumulative rainfall in outer Mongolia, 1940-1970 data
Why Correlation Need Not Indicate Causation

- **Omitted Variables/Selection Bias:** There may be other third factors \((z)\) that caused \(x\) and \(y\)

- **Endogeneity/Reverse Causality:** Causation may run from \(y\) to \(x\)
Example: Returns to Schooling

- High positive correlation between schooling and earnings could be due to unobserved ability/effort/background (omitted variable)
- High negative correlation between learning and class-size could be due to effect of student performance on school budgets/enrollment (reverse causality)
Micro-level Data

- Don’t just use macro or country-level data
- Go to micro-level data: study the relationship at the level of communities, individuals
- Many Reasons:
  - seek corroborating evidence
  - expand sample size
  - avoid aggregation/measurement errors
Step 1: Dealing with Omitted Variables, Time Series Data

- Don’t omit them!!
- Include time trend as a control in the time-series regression: e.g.

\[ \log W_t = \alpha + \beta S_t + \gamma t + \epsilon_t \]

- Amounts to checking if deviations from time trend of \( S_t \) and \( W_t \) are correlated
- Or changes from one year to the next are correlated:

\[ \log W_t - \log W_{t-1} = \beta (S_t - S_{t-1}) + \gamma + (\epsilon_t - \epsilon_{t-1}) \]
Step 1: Dealing with Omitted Variables, Cross-Section data

- Include controls for possible third variables $V_i$ that may be driving both wages and schooling, such as IQ, family background (parental education, wealth), neighborhood characteristics (distance from school, library etc):

  $$\log W_i = \alpha + \beta S_i + \gamma V_i + \epsilon_t$$

- Check if the estimate of $\beta$ is robust to inclusion of such controls
Table: Estimated Returns to Education for Men in NLSY (1979 cohort, 2002 survey)

<table>
<thead>
<tr>
<th>Controls:</th>
<th>RoR (s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>.132 (.007)</td>
</tr>
<tr>
<td>Age</td>
<td>.131 (.007)</td>
</tr>
<tr>
<td>&amp; Background</td>
<td>.114 (.007)</td>
</tr>
<tr>
<td>&amp; AFQT Score</td>
<td>.087 (.009)</td>
</tr>
<tr>
<td>&amp; Occupation Dummies</td>
<td>.066 (.010)</td>
</tr>
</tbody>
</table>
Step 1: Dealing with Omitted Variables, Cross-Section data, contd.

- However, it may be impossible to include all possible omitted variables (e.g., how caring or attentive parents are)
- So then there may still be unobserved factors that could be driving the observed correlation: problem of *unobserved heterogeneity*
Pooled Cross-Section-Time-Series (Longitudinal/Panel) Data

- One solution to problem of unobserved heterogeneity: use pooled (longitudinal) data
- For example, follow a set of individuals \((i)\) over time \((t)\) and see how their earnings \(W_{it}\) change as their education level \(S_{it}\) changes
- Can include an *individual fixed effect* as a control in the regression, which captures all individual-specific characteristics such as ability, family background etc.:

\[
\log W_{it} = \alpha_i + \beta S_{it} + \gamma V_{it} + \epsilon_{it}
\]
Another way to rewrite the panel regression:

\[ \log W_{it} - \log W_{i,t-1} = \beta [S_{it} - S_{i,t-1}] + \gamma [V_{it} - V_{i,t-1}] + \text{error} \]

This way we washed out the fixed effect of background, connections etc \((\alpha_i)\) entirely.

We are ignoring *level* differences between people, and focusing on *changes* in earnings and their relation to *changes* in schooling over time of each person.
Problem No. 2: Reverse Causality/Endogeneity

- Even if the estimated $\beta$ remains significant and robust with the preceding formulation, there may still be concerns about the direction of causality.
- Is it possible that the direction of causation is reverse (from earnings to education? from test scores to class size?)
- Is it possible there are time-varying omitted variables that are driving both earnings and education (such as wealth increase; new room-mate/partner...?)
Solution to Endogeneity Problem: Instrumental Variables (IV)

- One way of getting around this problem: find an instrumental variable $l_{it}$: which
  - causes schooling to vary,
  - apart from effect on schooling, has no direct effect on earnings

- Example: quarter of birth, combined with mandatory school attendance rules in US
- Those born early in the year tend to attain less education
Other Examples of Instrumental Variables

- Passage of mandatory schooling attendance law at some point of time (provided data includes cohorts both affected and not affected)
- In study of immigrant flows on wages, use external events that affect immigrant flows (lifting of Cuban embargo on emigration, on wages in Florida)
IV/2SLS estimates

- Obtain IV/2SLS estimate following a two step procedure:
  - predict variation in $S_{it}$ explained by IV $I_{it}$:
    \[ \hat{S}_{it} = \theta_i + \mu I_{it} + \text{residual} \]
  - then obtain variation in earnings associated with these predicted variations in schooling:
    \[ \log W_{it} = \alpha_i + \beta \hat{S}_{it} + \gamma X_{it} + \text{residual} \]

- **Main Idea:** Confining attention only to variations in earnings associated with ‘externally caused’ variations in schooling
Conditions of Validity of IV estimate

- First requirement: the instrument should be a strong predictor of schooling (can be checked)
- Second requirement: *exclusion* restriction: conditional on schooling the instrument should have no relation to earnings whatsoever (cannot be tested; matter of judgment)
What If There Is No Valid Instrument?

- Then there is lack of solid evidence of causal impact of schooling on earnings, or an estimate of the rate of return.
- Have to be lucky: exploit a *natural experiment*.
- Only remaining option: conduct a *controlled experiment*.
RCEs: The ‘Gold Standard’ of Policy Evaluation

- Take a bunch of individuals, conduct a lottery and allow the winners to receive a large scholarship conditional on an extra year of schooling.
- Then the lottery outcome forms an instrumental variable.
- By construction it is unrelated to anything that might affect earnings, except schooling.
- Once the experiment is conducted, compare earnings of winners with losers.
RCEs: the ‘Gold Standard’ continued

- RCEs have become very popular among development economists over the past decade
- Surefire way to deal with problems of omitted variables and concerns about causality
- Allows researchers to design and test new policies that have not been tried before
- Standard method used in pharmaceutical drug trials, agronomy etc.
- Governments/aid organizations tend not to like them, but are slowly yielding...
Gold or Just Superior Metal?

- Many debates in the literature over whether RCEs are ‘gold’ or ‘silver’ or ‘bronze’
- Number of shortcomings:
  - External validity
  - Short-term, narrow impacts
  - Infeasibility, ethical concerns in some contexts
- Certainly a valuable addition to research tool-kit
Summary of Steps in Statistical Rigor and Precision

- Macro to micro data
- Raw correlations
- Regressions with controls
- Regressions with panel/longitudinal data
- Endogeneity checks: Instrumental Variable Regression
- RCEs