Impact Evaluation and Evaluation of Financial Programs

Jacob Alex Klerman, Senior Fellow, Abt Associates

"Governance and Performance Management of Public Employment Services in the United States and the European Union" Brussels, September 2015

What is the Decision?

- In thinking about performance measurement and (impact) evaluation, it's useful to ask

What is the Decision?

1. How can we manage the program better (holding the program model fixed)?
   - Personnel actions (remediation, firing, promotion bonuses)
   - Contractor management (firing, rehiring, incentive payments)

2. Should we change the program model?
   - Continue vs. close the program
   - Modify the program

These are different decisions

An American Perspective

- If you're going to ask "impact questions"—e.g., Does this program work at all? Would some change make it work better?—use impact analytic methods

- The United States has a strong and growing tradition of random assignment evaluations
  - But, I'm a biased observer: I am part of that tradition, and one of my major funders is in the room!

- The United States also has a strong tradition in rigorous non-experimental impact evaluation methods
  - PSM/Propensity Score Matching (Heinrich)
  - RD/Regression discontinuity
  - DiD/Difference in differences

A/B Testing

- If the decision concerns program improvement, consider A/B Testing (sometimes called "Planned Variation" or "Rapid Cycle Evaluation"); i.e.,
    - Randomly assign clients (or sites) to different program models
    - Compare outcomes

- As long as you are willing to focus on proximate outcomes, A/B Testing can be
    - (relatively) Fast,
    - (relatively) Cheap,
    - Raises few ethical issues

Continuous Evaluation

- For mandatory programs, another option is "continuous evaluation"; i.e.,
  - Exempt a randomly selected subsample from the program
  - Compare outcomes

- This is the approach in the American REA/Reemployment and Eligibility Assessment Program

- Implementation needs to be done with care
  - Need a truly random sample
  - With constant control fractions
  - <or more subtle analytic strategies>

On Benchlearning I

- Benchlearning asks impact evaluation questions; i.e., (how) Should we change the program model?

- But, the American literature suggests that such cross-national comparisons are unlikely to identify best practices
  - Differences in measurement (Heckman, Ichimura, and Todd, 1997)
  - Differences in local labor market conditions (Heckman, Ichimura, and Todd, 1997)
  - Other (unmeasured, not accounted for) policy variation (standard critique of observational studies)
  - Too many possible explanations (Greenberg, Meyer, Michalopoulos, and Wiseman, 2003)

On Benchlearning II

- Some versions seem more promising

- With longer time-series and changes in policy within countries, use difference-in-differences
  - See Klerman and Danielson (2009), Heinrich and Fournier (2004) for examples

- With proper within country impact estimates (e.g., random assignment, PSM) this is a meta-analysis problem
  - See Bloom, Hill, Riccio (2003) for a widely cited example

- It is often striking how little things matter
  - Both narrow policy choices and programs overall

Apply Evaluation Ideas to Performance Measurement/Management

- Consider using evaluation ideas to identify good performers (employees or contractors); i.e.,
  - Randomly assign cases to employees/contractors
  - Then differences in outcomes are either due to true differences in performance or chance
    (and statistics allows us to bound the impact of chance)
  - See Klerman (2005); Baehler and Klerman (2015)

- New York City does this for their welfare programs (Feldman 2011)

- Idea has been used in lots of research papers
  - Disability Policy: Dahl, Kostøl, and Mogstad (2014); Recidivism: Di Tella and Schargrodsky (2014); Child Protection: Doyle (2008); Incarceration: Kling (2006)

Impact Evaluation and Evaluation of Financial Programs

Jacob Alex Klerman, Senior Fellow, Abt Associates

"Governance and Performance Management of Public Employment Services in the United States and the European Union" Brussels, September 2015