Policy Evaluation

Lecture 1: What is Policy Evaluation?

Edmund Malesky, Ph.D.
June 18, 2018

Duke University

Review of 4 Policy Interventions

1. Scared Straight: Juvenile Delinquency

2. Drug Substitutions Programs

3. Second Responder Program

4. School-Based Social Information Processing

From MacAskill, William, Benjamin Todd, and Robert Wiblin, 2015. "Can you guess which government programs work? Most people can't," *Vox Policy and Politics, August 17, 2015* < http://www.vox.com/2015/8/13/9148123/quiz-which-programs-work>

Preventing Juvenile Delinquency: *Scared Straight*
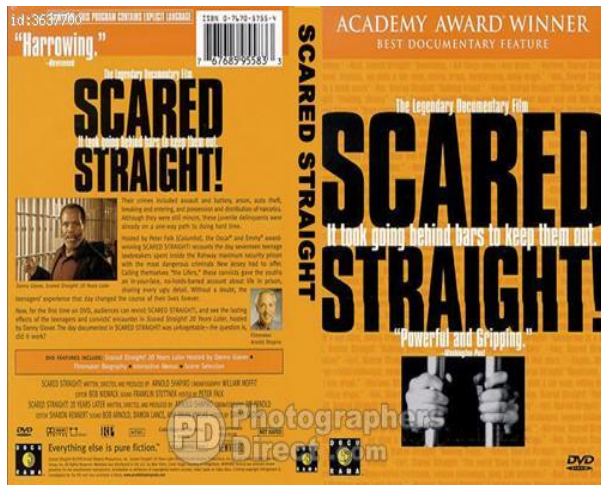
- *Scared Straight* is a program designed to discourage at-risk kids from committing crimes.

- The first program featured aggressive 'rap sessions' depicting life in prison, and received extensive media attention including a documentary.

- Following the documentary, over 30 prisons across the US adopted similar 'juvenile awareness' programs.

- All interventions in the Campbell review feature a prison visit, and most include a presentation by the prisoners which range from graphic to educational. Programs sometimes include tours and extended orientation sessions, such as living as a prisoner for 8 hours.

- ***Do Scared Straight and other 'juvenile awareness' programs reduce the rate that participants commit crimes in the future?***

Preventing Juvenile Delinquency:
Scared Straight

- ***Do Scared Straight and other 'juvenile awareness' programs reduce the rate that participants commit crimes in the future?***

- **Answer: No, they increase criminal activity.**

- [Nine randomized trials covering a 25-year period in eight different prisons](#) demonstrated that Scared Straight programs increase the odds that participants will commit crimes in the future. The results are the same whether the program is aggressive or passive.

- Each $1 spent on Scared Straight increased societal costs by $166.88.

- Despite the evidence, Scared Straight programs are still in use, and people [continue to believe](#) it's effective.

Drug Substitution Programs

- Drug substitution programs attempt to reduce criminal behavior and improve quality of life by giving addicts controlled amounts of drugs.

- This review focused on criminal behavior in opiate addicts, whether or not they had previously committed crime. Addicts were given heroin or substitutes such as methadone or buprenorphine, based on their needs.

- ***Does prescribing heroin or substitutes to addicts reduce crime?***

Drug Substitution Programs

- ***Does prescribing heroin or substitutes to addicts reduce crime?***

- Answer: Yes!

- Heroin prescription was found to "significantly reduce criminal involvement" by the [Cambell review](). All drugs covered in the review reduced crime, but heroin prescription had the most significant effect

Family Violence:
Second Responder Programs

- Second responder programs try to stop the repetition of family violence.

- A police officer and victim advocate come to your home a few hours or days after an incident has occurred.

- They talk to the victim about services and legal options, and may warn the offender of legal consequences if abuse continues.

- Second responder programs are used in many major US cities, with five states and over four thousand households covered in this review alone.

- *Do second responder programs reduce family violence?*

Second Responder Programs

- **Do second responder programs reduce family violence?**

- **Answer: No. Minimal and Conditional Impact**

- 10 high quality studies covering over 4,000 households were reviewed. The second responder intervention did not reduce the likelihood of repeat violence.

- However, it was found to slightly increase the odds that victims would report incidents to the police. The review concludes:

- "Policy makers will have to decide whether the small gain in willingness to call the police is worth the cost of these programs."

School-based Social Information Processing

- Social information processing is an educational intervention designed to reduce aggressive and disruptive behavior in children.

- This structured intervention focuses on improving thinking skills in a social context. Children are trained to process social cues, reflect on what they want, and select the best response.

- This review focused on school-based programs that were applied to entire classrooms. Programs varied from short and intensive, to weekly sessions over a year.

- **Do social information processing programs reduce aggressive and disruptive behavior when given to entire classrooms?**

School-based Social Information Processing

- **Do social information processing programs reduce aggressive and disruptive behavior when given to entire classrooms?**

- **Answer: Yes!**

- [Seventy-three unique studies](#) were reviewed, finding a positive effect.

- The review concludes that "students who participated in social information processing programs with their classmates showed less aggressive and disruptive behavior after treatment than students who did not receive the program."

- The review also notes that short, intensive programs of a few weeks were more effective than year-long programs. They theorize that extended programs may "have a tendency to become routine and thus have less impact on the students."

Organization of Today's Lecture

- Some information about you, the students
- Goals
- Structure
- Assignments
- Grading
- What is policy evaluation?
- How is it different from monitoring?
- Why you should care?

Goals of the Course

- How do we know whether a particular policy, strategy, program, event, or institutional change achieved the desired goals? How do we select more effective strategies in the future?

Learning Objectives

1. Understand the main objectives of rigorous policy evaluation, including how to avoid common pitfalls that lead to incorrect conclusions.

2. Develop ability to select appropriate policy evaluation technique for specific government intervention.

3. Become proficient at reading, analyzing and critiquing data derived from policy evaluation.

4. Know how to design, implement, and interpret results from a simple Randomized Controlled Trial (RCT).

5. Develop ability to construct a Pre-Analysis Plan (PAP), which describes theory of change, outcome variables, analysis techniques, data visualizations for proposed evaluations.

Not your Typical Class

- This is a course about skill development. You are required to learn techniques and apply them on assignments.

- You will never be asked to regurgitate formulas or facts for a test.

- Attendance matters – The readings are difficult and you will need our help making sense of them.

Course Structure (4 Sections)

1. Fundamentals of Policy Evaluation
2. Randomized Controlled Trials
3. Natural Experiments
4. Survey Experiments

Grading

1. Participation (10%)
2. 2 Problem Sets (30%)
3. Research Proposal (20%)
4. Pre-Analysis Plan (40%)

Course Materials

- Angrist, Joshua and Jorn-Steffen Pischke. (2014). *Mastering 'Metrics: The Path from Cause to Effect.* Princeton University Press.

- Khandker, Shahidur R., Gayatri B. Koolwal, and Hussain A. Samad (KKS, 2010). *Handbook on Impact Evaluation Quantitative Methods and Practices*. The International Bank for Reconstruction and Development, The World Bank. eISBN: 978-0-8213-8029-1

- Evidence in Governance and Politics. (EGAP, 2018). Methods Guides. < https://egap.org/list-methods-guides>

- Additional short reading on specialized topics listed with hyperlinks below.

Evidence Based Inquiry

- **Inquiry involves investigating the answers to puzzles and questions.**

Examples:
  - How do we alleviate poverty?
  - How do we reduce pollution?
  - Do social inequalities affect civic participation?
  - Can better infrastructure improve economic performance?

From Theory to Testing

- **<u>Theory:</u>** From deductive theory, we begin with expectations of the world.

- **<u>Hypothesis:</u>** We construct testable implications of those expectations (hypotheses)

- **<u>Measurement:</u>** We find ways to operationalize our outcome and treatment variables.

- **<u>Test:</u>** We analyze evidence to see if it is consistent/inconsistent with our expectations.

- **<u>Build</u>**: We use our findings to help in the positive accumulation of knowledge.

Hypothesis Testing

- **We often want to know if one variable is dependent on another**
  - How does an outcome differ with exposure to a treatment?
    - e.g., Does agricultural extension increase farmers' productivity?
  - How does an outcome differ for different values of another variable?
    - e.g., Will provinces with higher PCI scores grow faster than others?
  - How does an outcome differ after a program is put in place?
    - e.g., Does a grant to transport students to better school improve literacy?

- We look to see if variation in our independent variables (i.e., treatment, causal factors, programs) leads to variation in our dependent variables (outcomes)

- **Correlation is not causation**
  - But a lack of correlation helps us rule out causation
  - Ideally, we would design our studies to rule out other potential causes

You Do This All The Time?

- Thinking about the effects of certain foods on your waistline
- Thinking about the effects of the ordering of certain beverages on your stomach
- Thinking about the effects of injury on your favorite World Cup team?
- Thinking about the effects of certain gifts on the strengths of your relationships
- Thinking about the effects of your relationships on your academic performance

Ultimate Goal of Policy-Makers

- Conduct policies and programs that help citizens and businesses, and let them participate in economic growth and prosperity.

- But... How do we know that a particular policy or program "works"? Can different policy designs work even better? If yes/no, why and when?
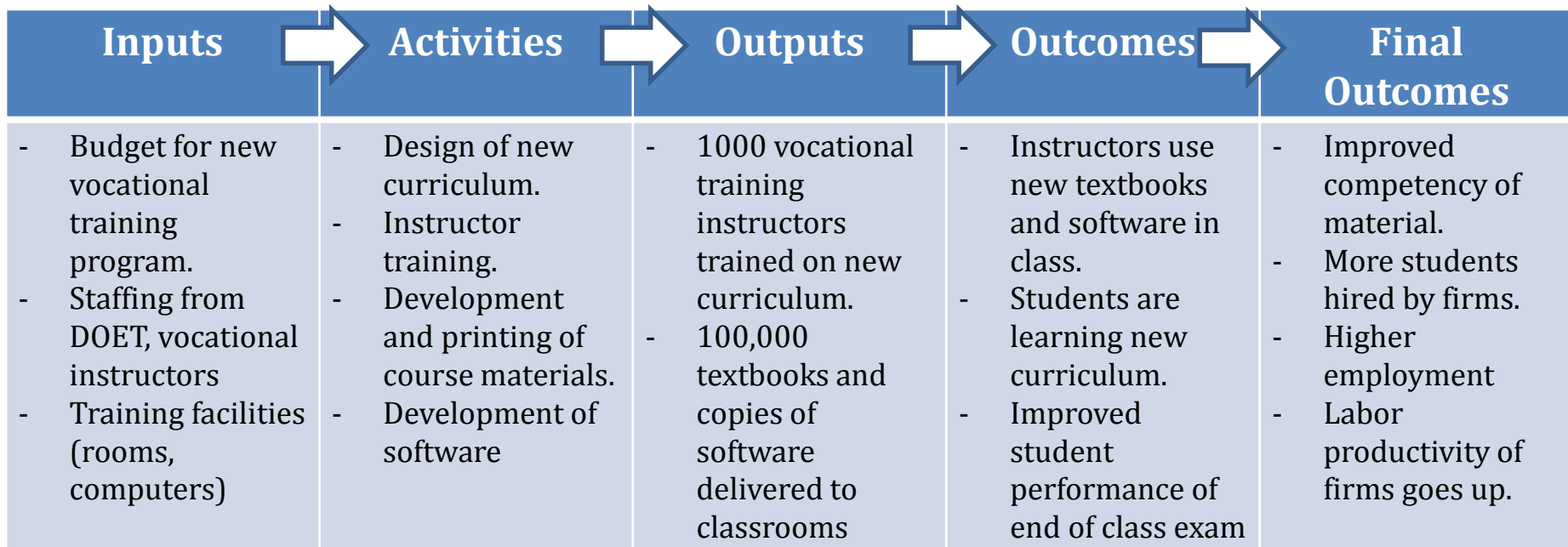
What is a Policy/Program?

| Resources/ Inputs | → | Activities | → | Outputs | → | Outcomes | → | Impact |

WE DO STUFF
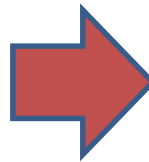
GOOD THINGS HAPPEN

Problem: Mismatch between worker skills and employer needs.
Solution: Better vocational training.
**Program: New vocational educational training curriculum**

Inputs	Activities	Outputs	Outcomes	Final Outcomes
- Budget for new vocational training program. - Staffing from DOET, vocational instructors - Training facilities (rooms, computers)	- Design of new curriculum. - Instructor training. - Development and printing of course materials. - Development of software	- 1000 vocational training instructors trained on new curriculum. - 100,000 textbooks and copies of software delivered to classrooms	- Instructors use new textbooks and software in class. - Students are learning new curriculum. - Improved student performance of end of class exam	- Improved competency of material. - More students hired by firms. - Higher employment - Labor productivity of firms goes up.

Implementation (Supply Side)	**Results (Demand + Supply)**

Monitoring v. Evaluation

	Monitoring	**Evaluation**
Frequency	Regular, Continuous	Periodic
Coverage	All programs	Selected program, aspects
Data	Universal	Sample based
Depth of Information	Tracks implementation, looks at What?	Tailored, of to importance, and impact. Asks Why?
Cost	Cost spread out	Can be high
Utility	Continuous program, improvement, management	Major program decisions

Distinguishing Evaluation from Monitoring

M = Monitoring

Inputs   Activities   Outputs   Outcomes   Impacts

E = Evaluation

Internal   **External**
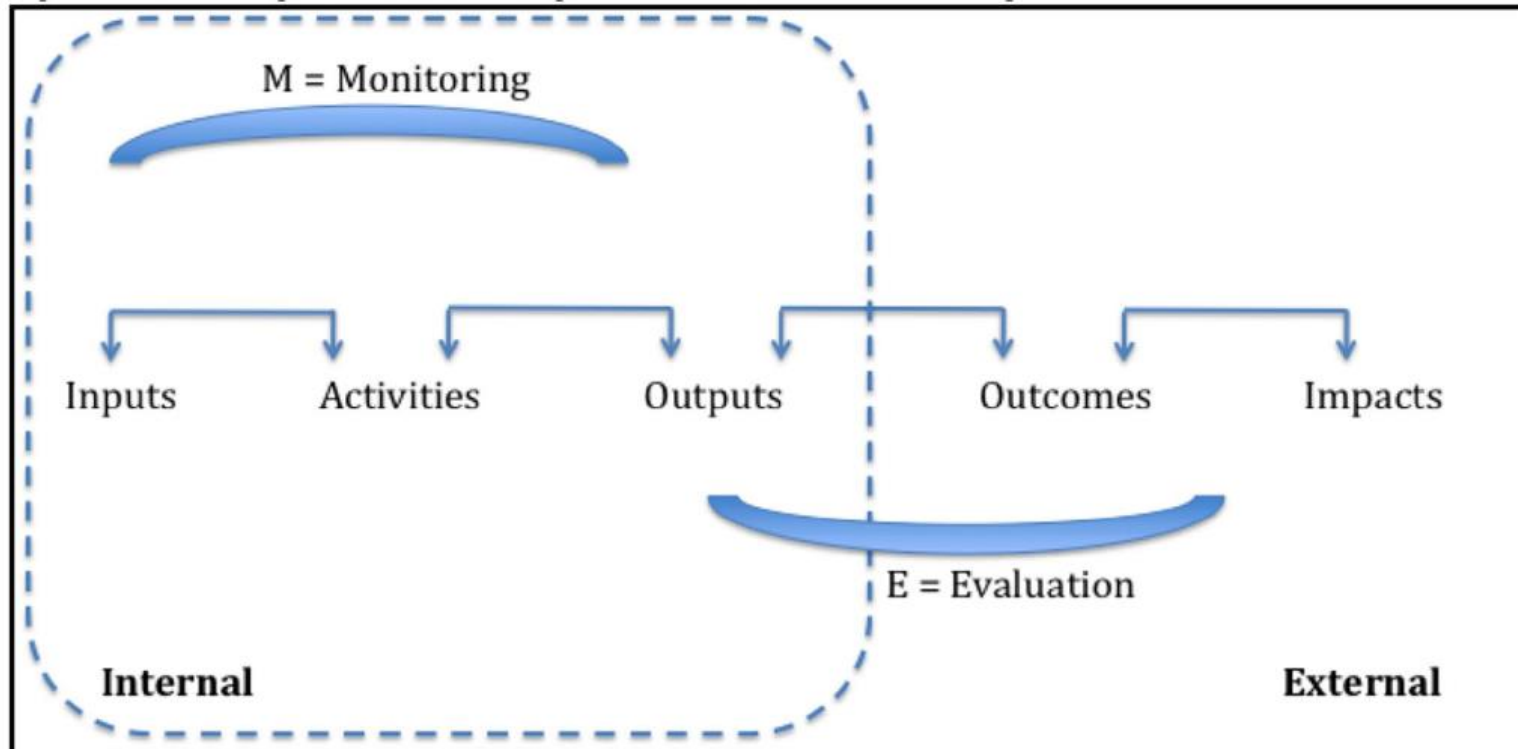
Source: Pritchett, Samji and Hammer et al, 2012.

Why is Policy Evaluation Important?

- Thought experiment
  - Child drowning in lake – Do you ruin your clothing (worth $100) to save her?
  - Do you send $100 now to an NGO in Malawi to save child?
- Role of evaluation is to overcome inaction
- Would we prescribe medicine without studies of impact of side effects?
- We need to learn from mistakes
  - Every year we spend millions on policy programs, intervention, and overseas aid.
  - On many of these programs, evaluation has suggested little impact (i.e. village councils and corruption).
- Need to learn from our mistakes to change tact
- Impact evaluation can help us do this.

Course Focus: Impact Evaluation
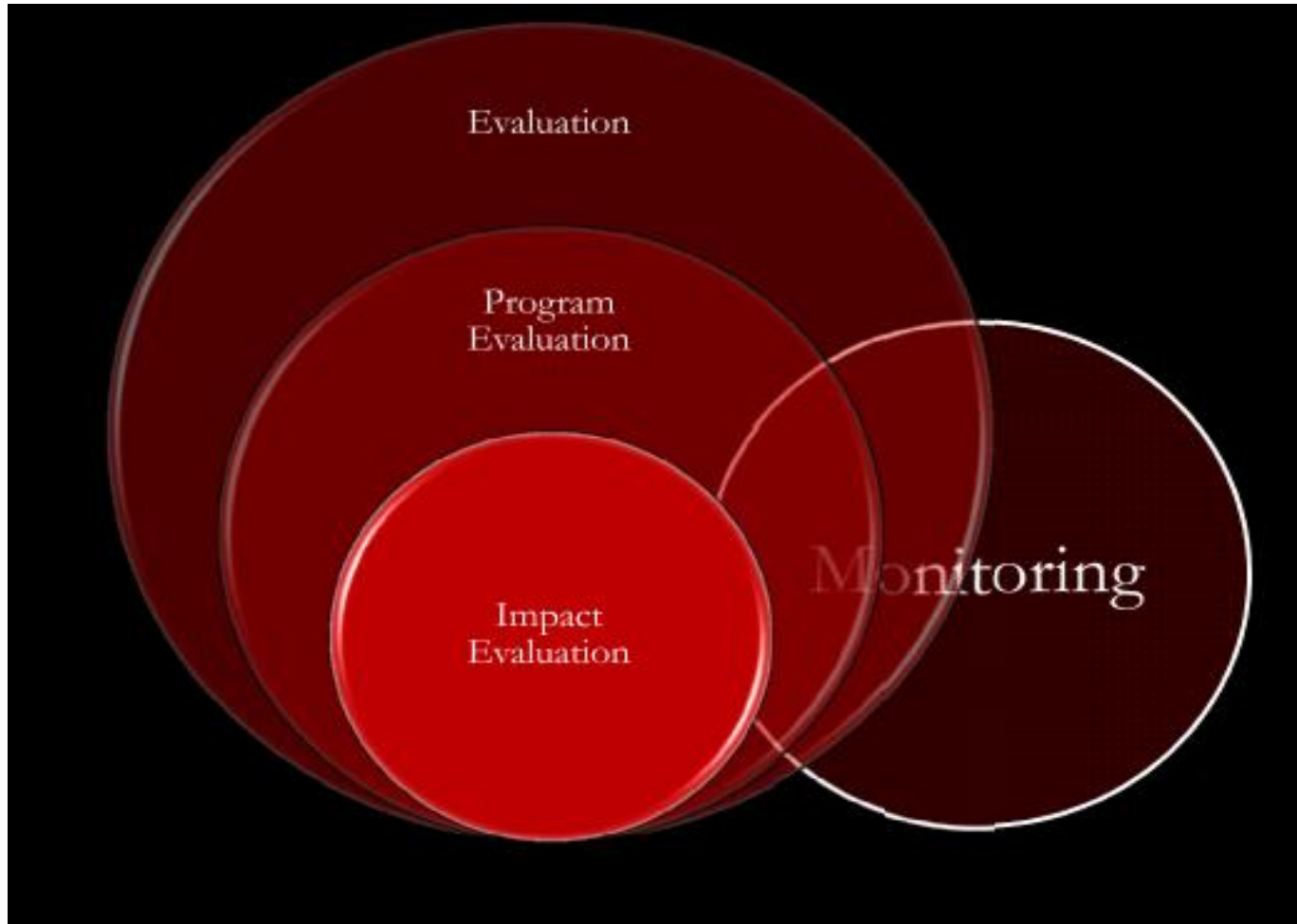
Components of Policy Evaluation

Needs Assessment	What is the problem to be solved?
Program Theory Assessment	How does the program fix the problem?
Process Evaluation	Does the program work as planned
Impact Evaluation	Were goals achieved? Magnitude of effect?
Cost Effectiveness	Given magnitude and costs, how does it compare to alternatives?

Program Theory Assessment

- How will the program address the needs put forth in your needs assessment?
    - What are the prerequisites to meet the needs?
    - How and why are those requirements currently lacking or failing?
    - How does the program intend to target or circumvent shortcomings?
    - What services will be offered?

- Tools
    - Theory of Change
   - Logical Framework (LogFrame)

What is Impact Evaluation?

- **Evaluations** are periodic, objective assessments of a planned, ongoing, or completed program
  - Can ask descriptive, normative, or cause-and-effect questions

- **Impact evaluations** – the primary focus of this course – assess changes in outcomes *attributable* to program, policy or event.
  - Try to identify *causal relationship* between program and outcomes
  - Ask about difference between what happened with program & what would have happened without it (i.e., counterfactual)
  - As much as possible, try to re-capture the experimental ideal

- Also, when a program can be implemented in several ways, which one is the most effective?
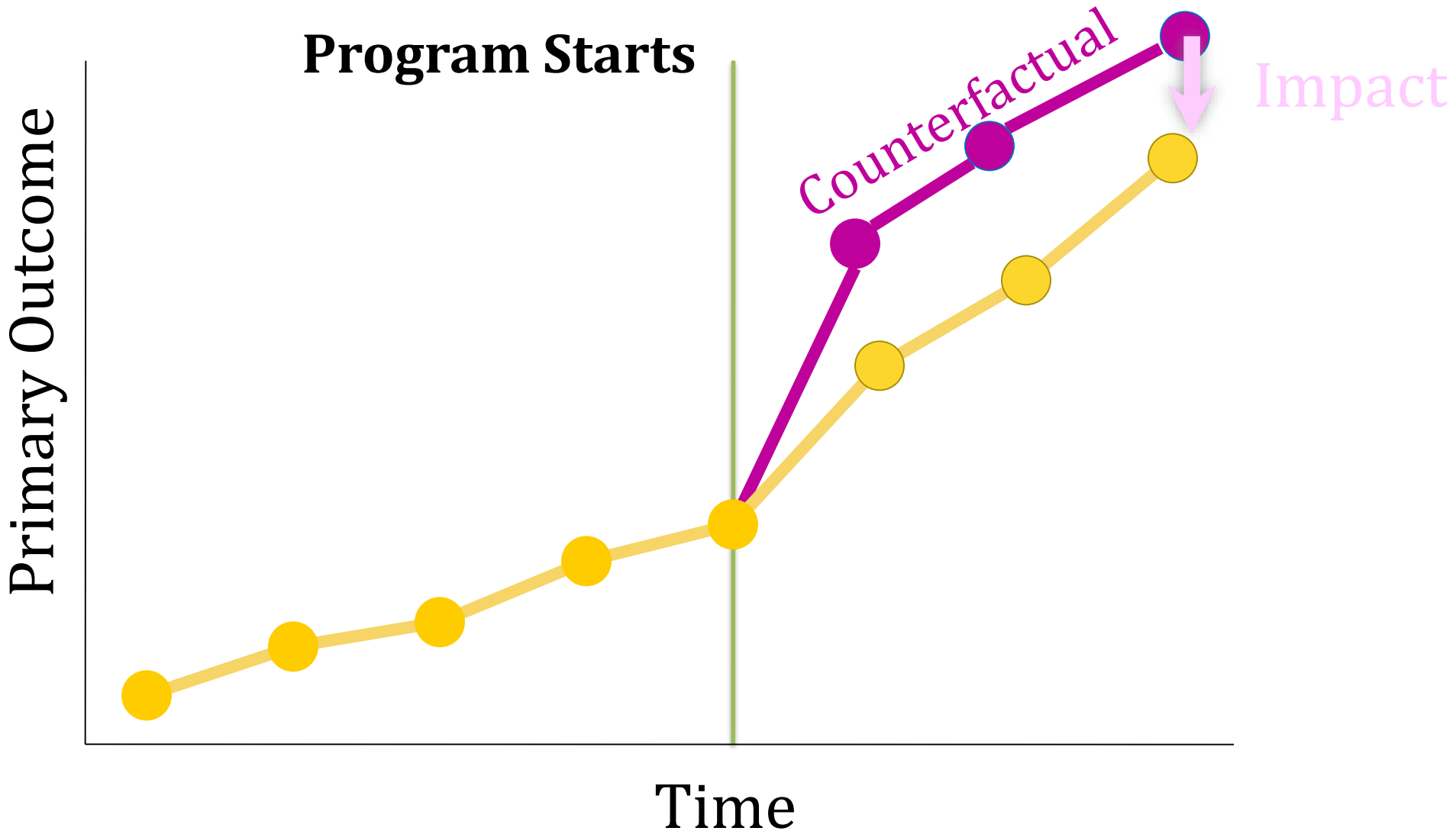
Learning about Impact is Hard

- What effect (if any) did the policy have?

  - How would individuals who experienced the program have fared in the absence of the program?

  - How would those who did not experience the program have fared if they had been exposed to the program?

What is Impact?

What is Impact?

Program Starts

Counterfactual

Impact

Primary Outcome

Time
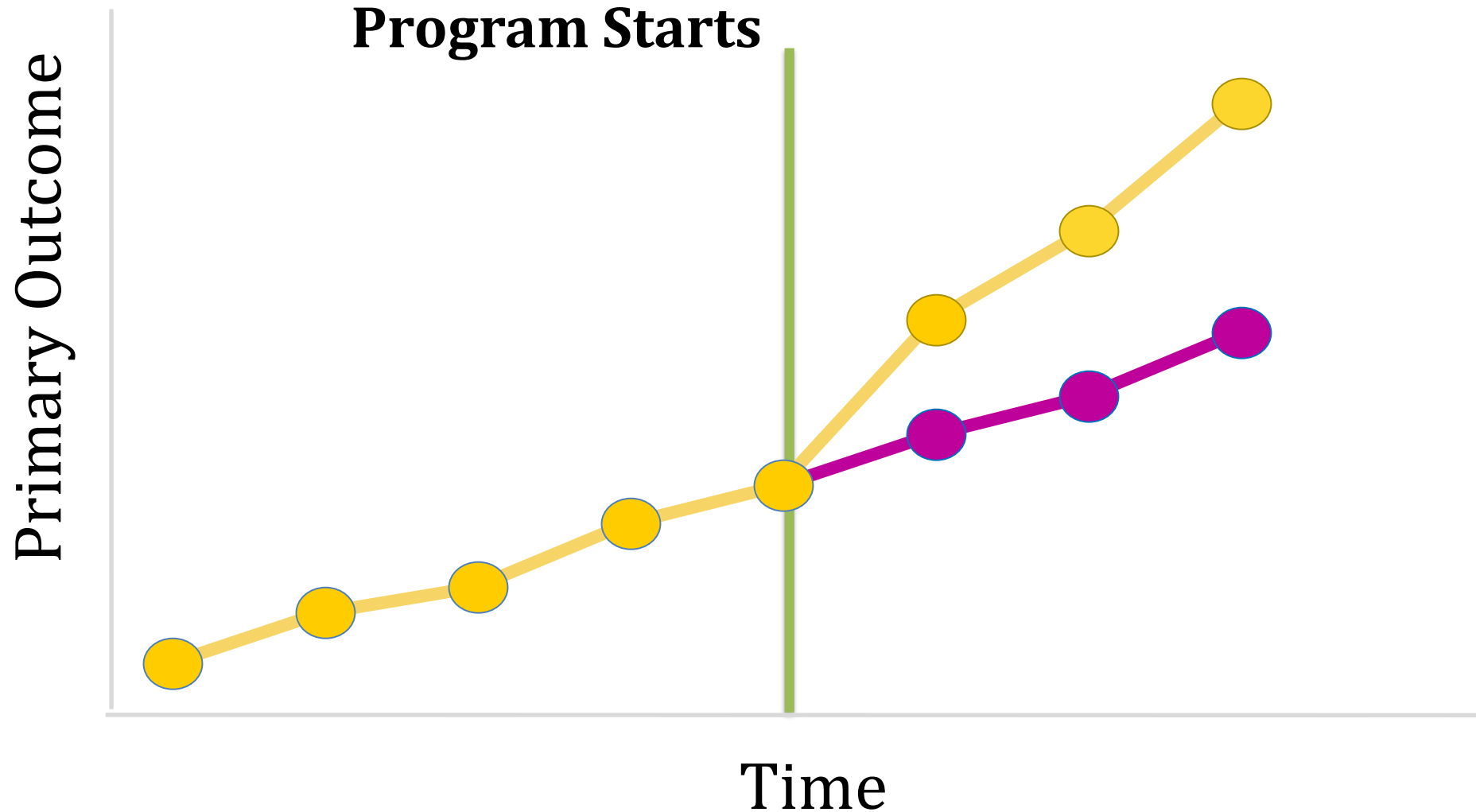
The Counterfactual

- Idea: Select a group that is <u>**exactly like**</u> the group of participants in all ways except one: their exposure to the program being evaluated
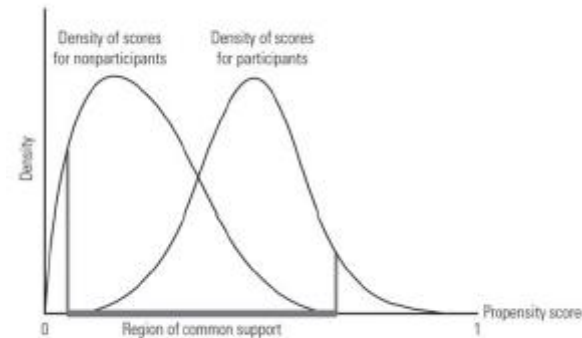

- Goal: To be able to <u>**attribute**</u> differences in outcomes between the group of participants and the comparison group to the program (and not to other factors)
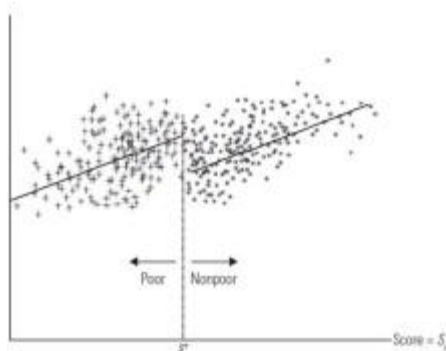
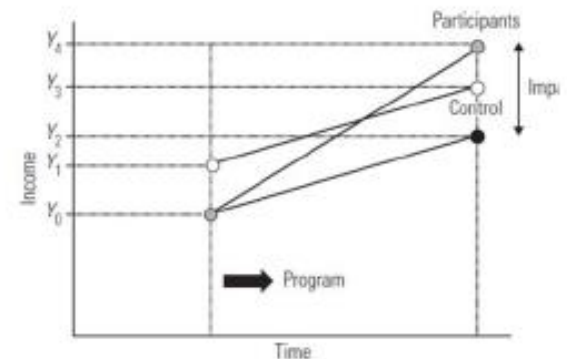# Techniques We Will Learn

Randomized Evaluation
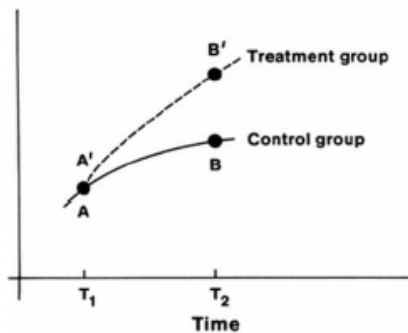
Matching

Regression Discontinuity

Differences in Differences

37

Review

- Impact evaluation assesses impact of program on outcome(s)
- Thus, causal inference is central focus of impact evaluation
    - Did program, and program alone, lead to Δ (change) in outcome?
- Correlation ≠ causation warning doesn't satisfy policy makers
    - Seek rationale for decisions: If we do X, will we get Y?
- Causality can be viewed as problem of counterfactual
- Evaluators' main challenge: determine what counterfactual state of world looks like → comparison groups
- Must avoid common pitfalls of invalid comparison groups
- Numerous statistical techniques, but shoe leather also key