Policy Impact Evaluation: A Quick Overview

Gelson Tembo
Department of Agricultural Economics
The University of Zambia

Paper presented at the Training Workshop and Policy Symposium,
Smallholder-Led Agricultural Commercialization and Poverty Reduction: How to Achieve It

Held at Lemigo Hotel in
Kigali, Rwanda, April 18-22 2011

Outline

- Monitoring & Evaluation
- Impact evaluation, and why?
  - Need for a good counterfactual
- Social experiment vs observational studies
- Difference-in-Difference Estimator
- An Application
- Exercises

Monitoring & Evaluation

- Monitoring
  - Regular collection and reporting of information
    - Compare actual results with targets
    - Is the project doing things right?
  - Understanding different levels of indicators
    - Inputs -> Outputs -> Outcomes -> Impacts
  - Selecting indicators
- Evaluation
  - Analytical efforts to answer specific questions
    - About performance of program activities
    - Useful for understanding processes, not causality. e.g. NPV
    - Is the project doing the right things?

Why Impact Evaluation?

- Determine if policies work?
  - Existence and size of benefits
  - Distributional effects
- Logical justification for the programs
- Can generate lessons for scale up, etc
  - Learning from others!

Impact Evaluation

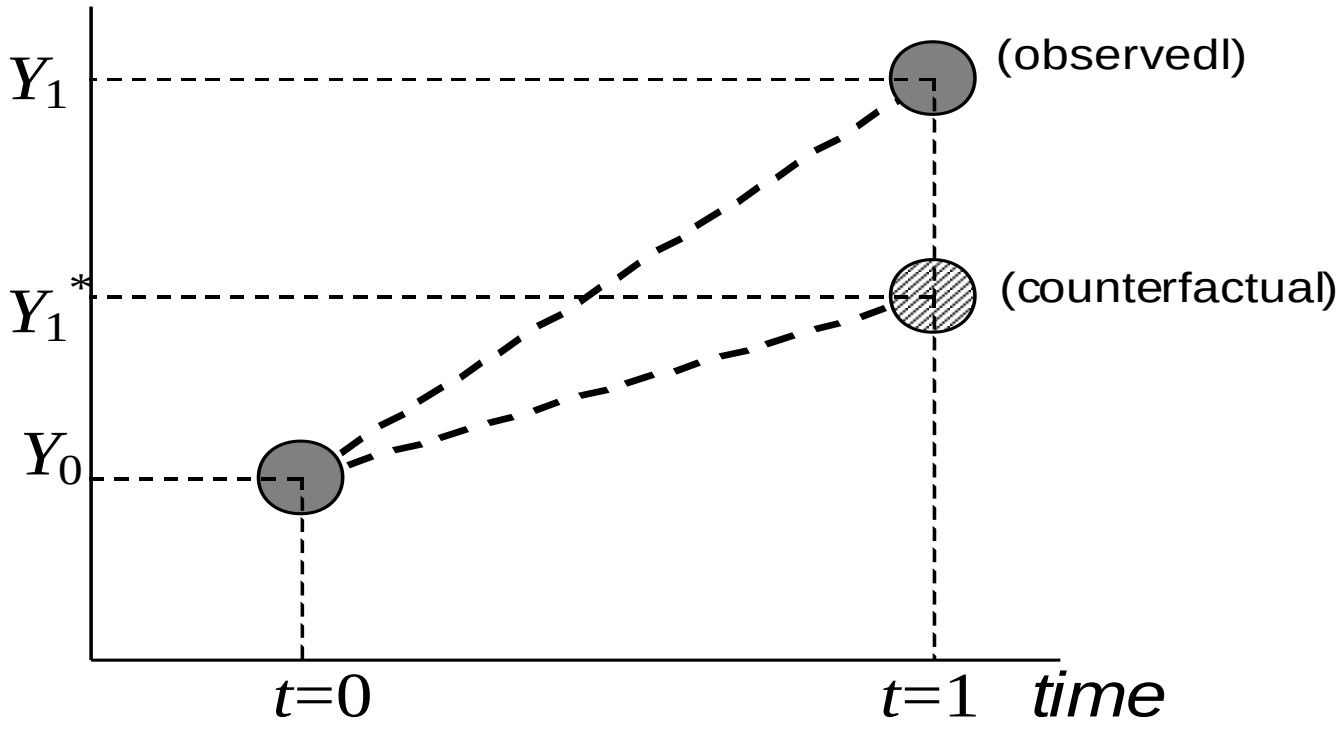
- Impact is the difference in outcome with and without the project
- IE is concerned with measuring this difference in a manner that it can be attributed to the program
  - Goal: To establish causal impact

  → ATET = Average Treatment Effect on the Treated

But there is a problem!

- Individuals cannot be both in the treatment and the control
  - The missing data problem
- Programs are seldom randomly placed
  - Participation is endogenous (selection bias)
    - Programs are targeted
    - Participants choose to participate
  - Naïve methods will not work!
    - Before and after differences
    - With and without differences

Role of a good counterfactual



- A comparison group as similar as possible to the treatment group
  - Observables --- > Data
  - Unobservables
- Except for participation status

Two alternative strategies

- Social experiment
  - Experimental designs
  - Randomization
- Observational studies
  - Quasi- or non-experimental designs

Social experiment

- Social experiment
  - Pure randomization can reveal the counterfactual but is rare/impractical
    - Ethical challenges!
  - In practice, conditional on observables
  - Examples
    - The famous PROGRESA in Mexico
    - E.g. social cash transfer schemes in Zambia
    - The WB's Results-Based Financing (RBF)

Observational studies

- Two conditional independence assumptions
  - A correlate of placement is independent of outcomes given *D* and *X*
    - Instrumental variables estimator
  - Placement is independent of outcomes given *X*
    - Single difference methods assuming conditionally exogenous program placement
    - Double difference methods assuming placement is independent of outcome changes

Difference-in-Difference
(Double-Difference)

- Observed changes over time for non-participants provide the counterfactual
  - Allows for time-invariant and additive selection bias
- Steps:
  - Collect baseline data on non-participants and (probable) participants before the program
  - Compare with data after the program
  - Subtract the two differences, or use a regression with a dummy variable for participant

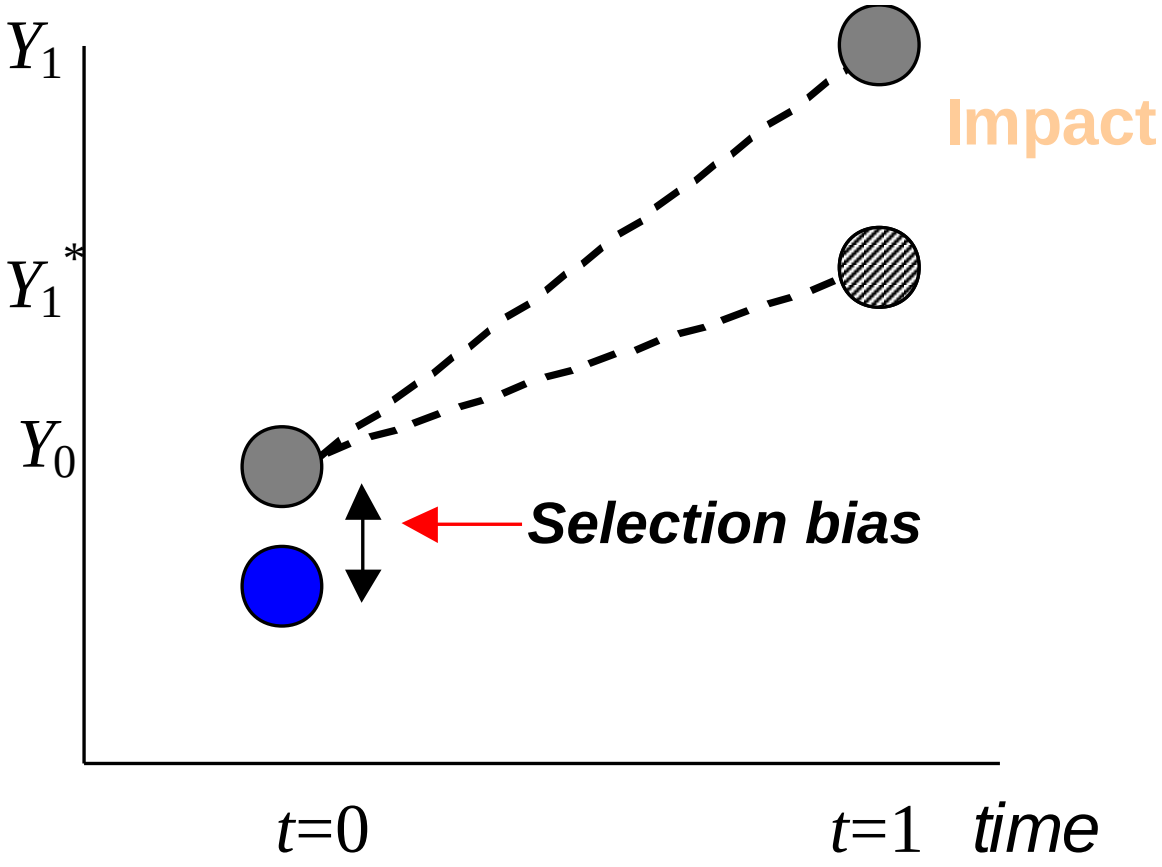
DD Estimator

$$DD \equiv E[(Y_{i1}^{T} - \hat{Y}_{i1}^{C}) - (Y_{i0}^{T} - \hat{Y}_{i0}^{C})]$$

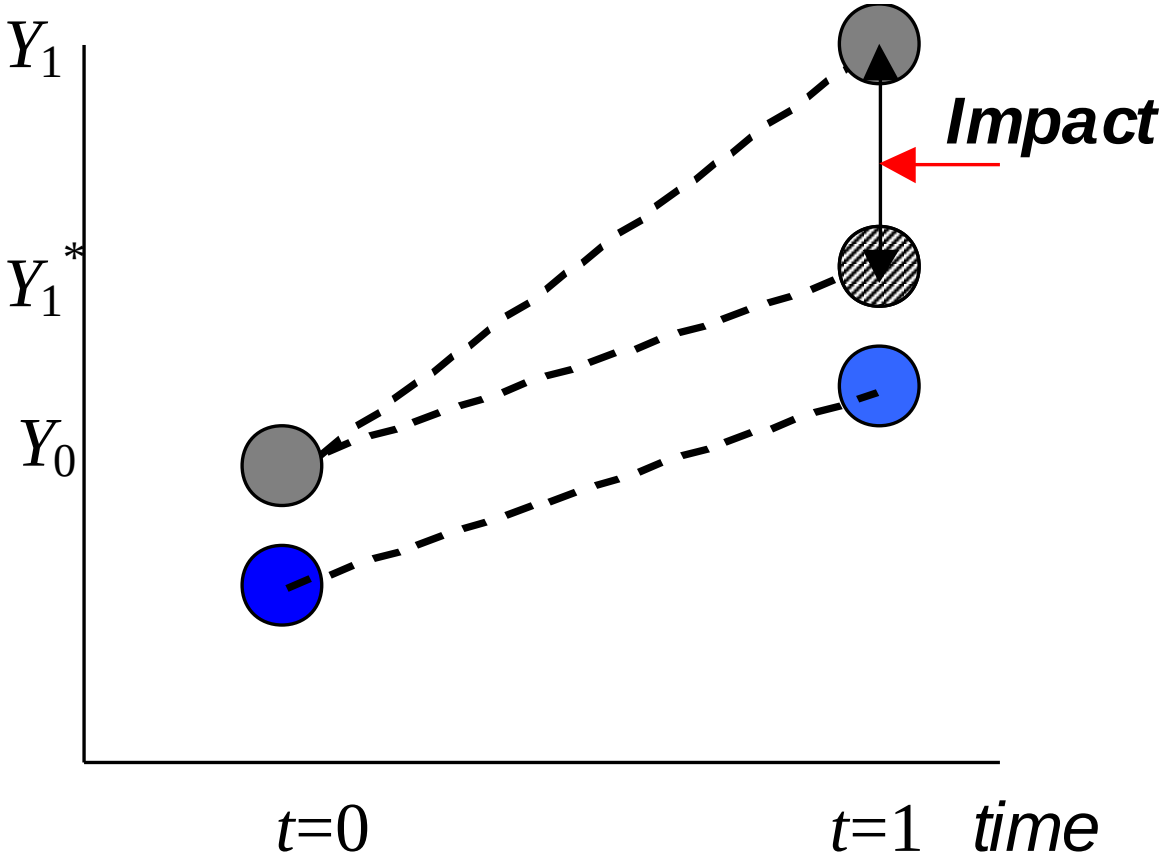
Post-intervention difference in outcomes

Baseline difference in outcomes

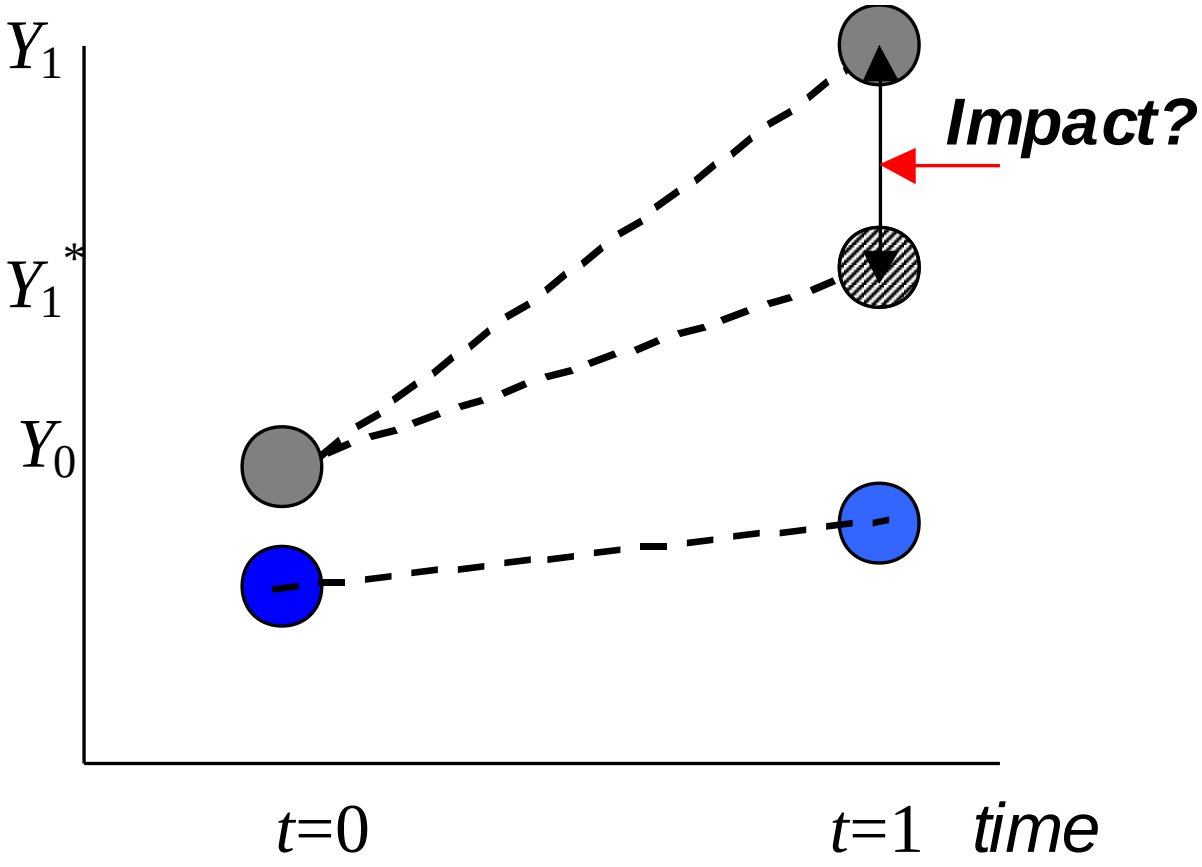
Selection bias



DD requires additive and time-invariant bias

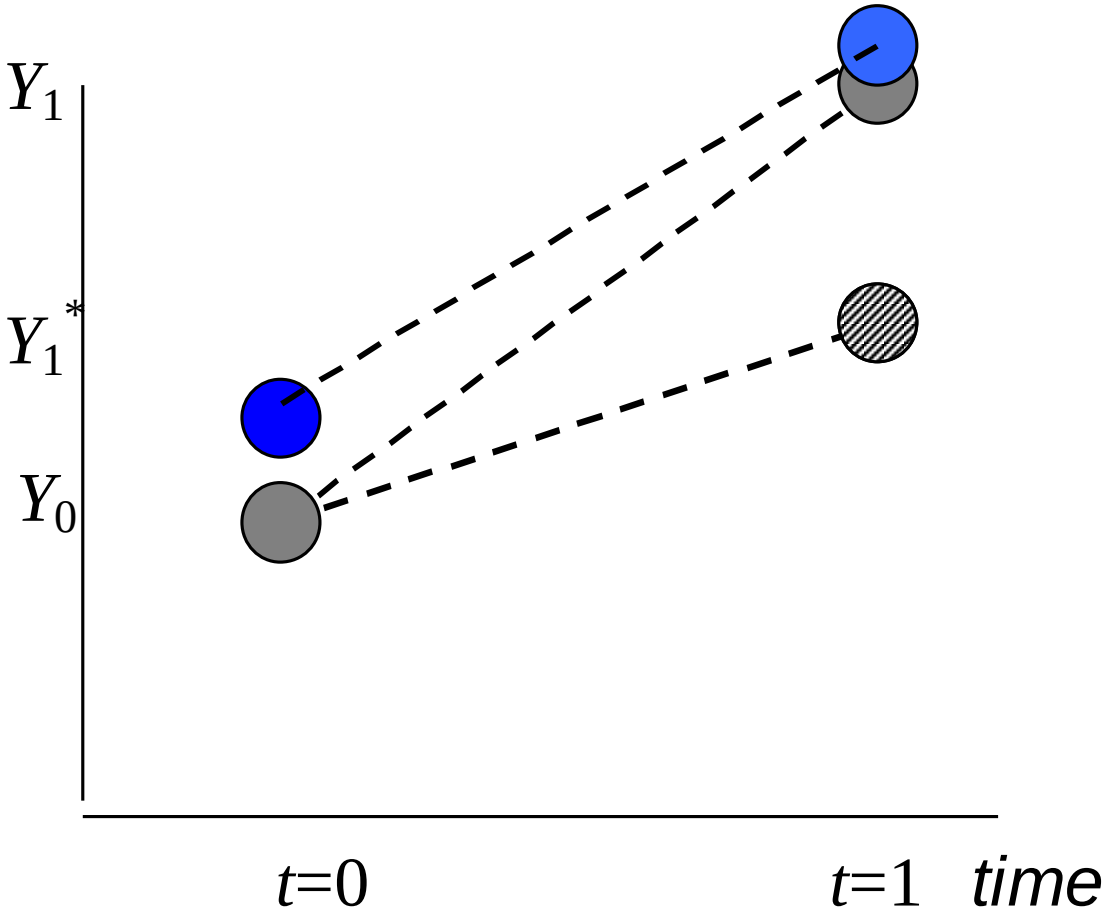


Impact over-estimated by DD



Comparison group on a different trajectory

Impact under-estimated by DD



Common problem in assessing impacts of development projects?

Example: Impact of the Agricultural Support Program

- Objective
  - Determine the impact of the Agricultural Support Programme (ASP) on productivity and household incomes
  - Determine the relevance and efficacy of ASP land and enterprise productivity-enhancing strategies and approaches

Conceptual issues

- ASP households self-selected themselves
  - They were not representative of the population
    - Some household attributes may influence the decision to participate
    - Some may also affect some of the welfare indicators of interest

    → Endogeneity!!
- ASP was not the only program offering similar services!
- Spillover effects

Stratified sampling strategy

- To control for
  - Spillover effects
  - Effects of other, non-ASP, programmes
- Allow for over-sampling of some strata

Six strata

Level of participation in the ASP	Experience with similar interventions implemented by non-ASP programmes	
	Without experience	With experience
	(1)	(2)
	---------- Stratum, $w \supset$ [illegible] k ----------	
Non-ASP communities	1	2
Non-participating household in ASP communities	3	4
ASP participants	5	6

Four DDs of interest

DD	Description	Remarks
DD51	• Impact compared to control communities • No exposure to other programs	• Counterfactual communities may be different
DD53	• Impact compared to non-ASP households within ASP communities	• Danger of spill over effects
DD62	• Impact compared to control communities; • With exposure to other programs	• Counterfactual communities may be different • Interaction effects with other programs
DD64	• Impact compared to non-ASP households within ASP communities	• Danger of spill over effects • Interaction effects with other programs

Household income

Crop/Agricultural season	Households in non-ASP communities		Households in ASP communities				Single Differences				Differences-in-Differences			
			Non-ASP households		ASP households									
	No other programs	Other programs	No other programs	Other programs	No other programs	Other programs	D51	D53	D62	D64	DD51	DD53	DD62	DD64
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(12)	(13)	(14)	(15)
	Real USD (2006)													
2005/06 agricultural season (*t = 1*)														
Crop income	425	382	204	240	322	363	-103	118	-18	123	-384	-142	-341	-129
Off-farm income	394	264	479	256	173	300	-220	-306	36	44	152	-254	-51	-164
Value of assets	2,522	1,369	360	2,183	1,589	3,030	-932	1,230	1,661	847	41	407	965	-61
Livestock income	64	56	0	9	11	66	-53	11	10	57	215	-200	-99	-72
Income from sale of assets	64	56	0	10	13	67	-51	13	11	57	318	-221	-77	-115
2002/03 agricultural season (*t = 0*)														
Crop income	47	182	68	252	328	504	281	260	322	252				
Off-farm income	562	268	242	148	189	355	-372	-52	87	208				
Value of assets	1,807	464	11	251	834	1,160	-974	823	696	908				
Livestock income	478	155	0	135	211	264	-268	211	110	130				
Income from sale of assets	610	228	7	144	242	315	-369	234	88	172				

Selected household characteristics across strata

Sample and household characteristic	Households in non-ASP communities		Households in ASP communities			
			Non-ASP households		ASP households	
	No other programs	Other programs	No other programs	Other programs	No other programs	Other programs
	(1)	(2)	(3)	(4)	(5)	(6)
Number of households	19	80	22	49	90	130
	---------- Mean ----------					
Effective dependency ratio	0.55	0.58	1.30	1.01	0.79	0.77
Household size	9.26	9.29	7.14	7.29	8.16	8.27
Age of the head (years)	53.93	45.35	48.24	45.63	49.34	46.12
Education (years)	6.00	6.12	4.05	6.73	6.41	6.92
Single female-headed (%)	26.3	16.5	27.3	8.2	18.9	13.1

Concluding remarks

- All the foregoing point to the need
  - For a good counterfactual
  - To control for observables
- Program historical and institutional knowledge is key
  - Be sure to understand the program well!
- There may be need for alternative estimators
  - PSM, Instrumental variable estimators, OLS
- Good idea to present results of more than one estimator - robustness

Exercises (Excel Spreadsheet)

$$DD \equiv E[(Y_{i1}^{T} - \hat{Y}_{i1}^{C}) - (Y_{i0}^{T} - \hat{Y}_{i0}^{C})]$$

$$\Delta y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$$

$$y_{it} = \beta_0 + \beta_1 D_i + \beta_2 T_t + \beta_3 (D_i * T_t) + \varepsilon_{it}$$