

## Impact Evaluation using the Alkire-Foster Measures as Outcome Variables

Ana Vaz

Monday March 20th, 2017 - 14:00 GMT

Tabita, Kenya



Rabiya, India



Stephanie, Madagascar



Agatha, Madagascar



Dalma, Kenya



Ann-Sashia, Kenya



Valérie, Madagascar



# Agenda

- Research Project
- Motivation
- General Hypothesis
- Impact Evaluation
- Challenges
- Empirical Examples
- Steps

# Research Project

- Objective: explore the use of Alkire-Foster measures in impact evaluation
- Expected outcome: 8-12 articles to submit for academic publication, potentially as a special issue

# Research Project

- Research Proposal (2 pages):
  - Intervention/policy/program;
  - Dataset;
  - Why intervention and dataset support using an AF measure;
  - Rough draft of the structure of the AF measure;
  - Impact evaluation methodology;
  - The value added of the proposed paper; and
  - Potential co-authors.
- Sent proposal to Ana and Bilal, with the e-mail subject "Impact Evaluation - Proposal".

# Research Project

- Main criteria to assess research proposals:
  - Potential contribution to OPHI research agenda;
  - Rigor of the methodology proposed;
  - Suitability of the AF method for the evaluation of the program;
  - Feasibility to meet the timeline.
- Interaction between OPHI and researchers:
  - Technical advice; and
  - Potentially, co-authoring.

# Research Project

- Timeline:
  - April 7<sup>th</sup>: Submission of research proposal
  - April 24<sup>th</sup>: Identification of selected proposals
  - August 21<sup>th</sup>: Submission of preliminary draft
  - October 16<sup>th</sup>: Submission of draft
  - Late October: Internal workshop
  - December/January: Workshop
  - Early March: Submission of final paper for publication

# Motivation

- 2030 Agenda for Sustainable Development:
  - Target 1.2 of SDG's: “By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty **in all its dimensions** according to national definitions.”
  - Emphasis on intersectoral linkages and integrated approaches
- Increased importance of policy coordination and multisectoral programs. (Ex.: Conditional Cash Transfers, Millennium Villages Project, etc.)

# Motivation

- Impact/program evaluation became fundamental.
  - It generates evidence of what works;
  - It is a mechanism of accountability at the end of a project.
- OPHI hopes to make a contribution at the intersection of poverty measurement and impact evaluation, and also contribute to the body of knowledge on the effectiveness of poverty reduction programs.

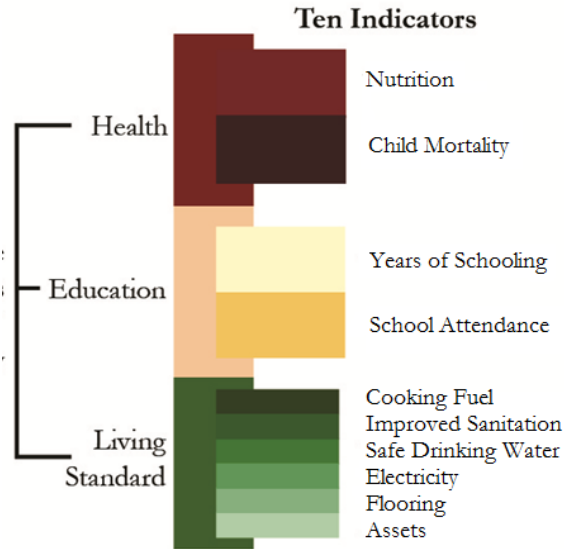


# General Hypothesis

- Hypothesis: In **certain circumstances**, the use of Alkire-Foster multidimensional measures can provide **valuable insights** to the evaluation of social protection programs/policies.

# What is an AF Measure?

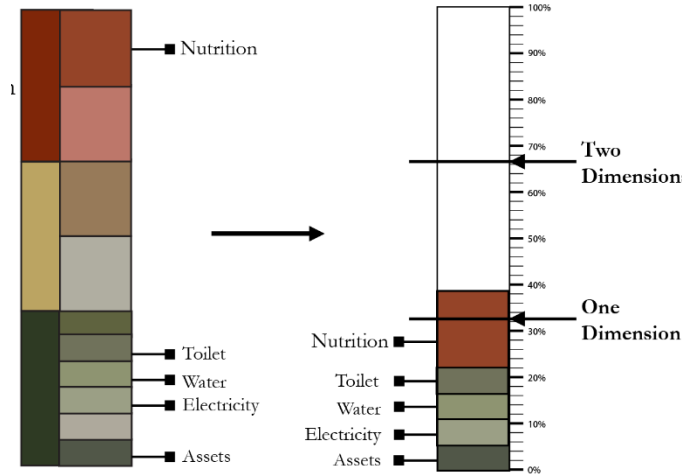
## 1. Select Indicators, Cutoffs, Weights



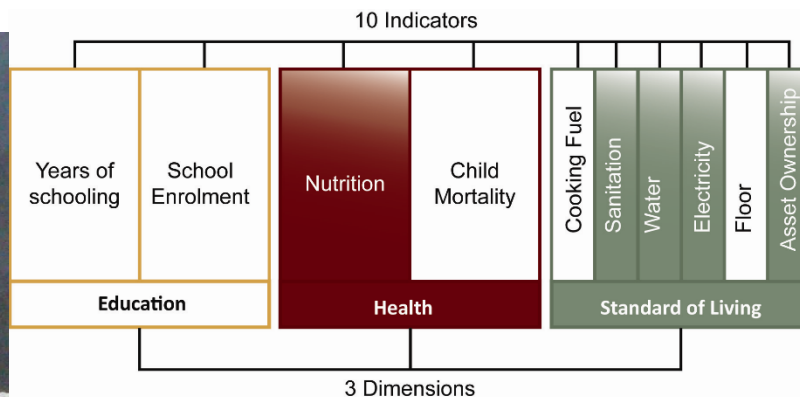
## 3. Identify who is poor

**4. Aggregation:**

- Incidence (H)
- Intensity (A)
- $MPI = H \times A$



## 2. Build a deprivation score for each person



# General Hypothesis

- Hypothesis: In **certain circumstances**, the use of Alkire-Foster multidimensional measures can provide **valuable insights** to the evaluation of social protection programs/policies.
- “**Certain circumstances**”. When program/policy has:
  - A multisectoral approach;
  - Multiple goals;
  - Potential synergies that may trigger effects at different levels.

# General Hypothesis

- Example. Suppose:
  - A poverty reduction program with  $D$  objectives;
  - Each objective can be defined in terms of minimum achievement thresholds,  $\mathbf{z}$ , for each target unit (person, household, community, etc.);
  - $w_d$  is the weight/importance of objective  $d$ ;
  - The overall goal of the program can be defined as reducing the weighted sum of the targets' deprivations below a certain cutoff,  $k$ .

# General Hypothesis

- In these circumstances, we can ‘translate’ the program’s overall goal into a AF measure.
  - $D$  objectives  $\Rightarrow D$  or plus indicators
  - $z$  minimum achievement thresholds  $\Rightarrow z$  deprivation cut-offs
  - $k$  is the program cut-off
- And use AF measure as the outcome of interest in the evaluation of the program’s impact.

# General Hypothesis

- Hypothesis: In **certain circumstances**, the use of Alkire-Foster multidimensional measures can provide **valuable insights** to the evaluation of social protection programs/policies.
- What kind of “**valuable insights**” can provide an evaluation based on an AF measure?

# General Hypothesis

Matrices of achievements of treated individuals										
	Baseline					Post-treatment				
	Income	Edu	BMI	Sanitation	Water	Income	Edu	BMI	Sanitation	Water
1	85	4	16	0	0	85	5	18	1	1
2	90	6	16	0	0	90	8	18	1	1
3	75	6	17	1	1	75	6	17	1	1
4	50	4	17	1	1	50	4	17	1	1
5	100	4	17	1	1	110	4	17	1	1
6	100	6	17	1	1	110	6	17	1	1
7	100	7	21	1	1	110	7	21	1	1
8	100	8	20.5	1	1	120	8	20.5	1	1
z	100	8	18.5	1	1	100	8	18.5	1	1

## Program's average impact in each indicator

Average achievement at baseline	87.50	5.63	17.69	0.75	0.75
Average achievement at post-treatment	93.75	6.00	18.19	1.00	1.00
Program's average impact	6.25	0.38	0.50	0.25	0.25

# General Hypothesis

## Deprivation matrices

	Baseline						Post-treatment					
	Income	Edu	BMI	Sanit	Water	No. depriv	Income	Edu	BMI	Sanit	Water	No. depriv
1	1	1	1	1	1	5	1	1	1	0	0	3
2	1	1	1	1	1	5	1	0	1	0	0	2
3	1	1	1	0	0	3	1	1	1	0	0	3
4	1	1	1	0	0	3	1	1	1	0	0	3
5	0	1	1	0	0	2	0	1	1	0	0	2
6	0	1	1	0	0	2	0	1	1	0	0	2
7	0	1	0	0	0	1	0	1	0	0	0	1
8	0	0	0	0	0	0	0	0	0	0	0	0

### Program's average impact on raw headcounts

Raw headcount at baseline	0.50	0.88	0.75	0.25	0.25	2.63
Raw headcount at post-treatment	0.50	0.75	0.75	0.00	0.00	2.00
Program's average impact	0.00	-0.13	0.00	-0.25	-0.25	-0.63



# General Hypothesis

	Overall goal: no one experiences... deprivation(s) or more				
	k = 1	k = 2	k = 3	k = 4	k = 5
<b>Levels</b>					
Baseline					
Incidence	0.88	0.75	0.50	0.25	0.25
Intensity	0.60	0.67	0.80	1.00	1.00
Adjusted headcount	0.53	0.50	0.40	0.25	0.25
Post-treatment					
Incidence	0.88	0.75	0.38	0.00	0.00
Intensity	0.46	0.50	0.60	0.00	0.00
Adjusted headcount	0.40	0.38	0.23	0.00	0.00
<b>Program's impact</b>					
Incidence (change)	0.00	0.00	-0.13	-0.25	-0.25
Intensity (change)	-0.14	-0.17	-0.20	-1.00	-1.00
Adjusted headcount (change)	<b>-0.13</b>	<b>-0.13</b>	<b>-0.18</b>	<b>-0.25</b>	<b>-0.25</b>

# General Hypothesis

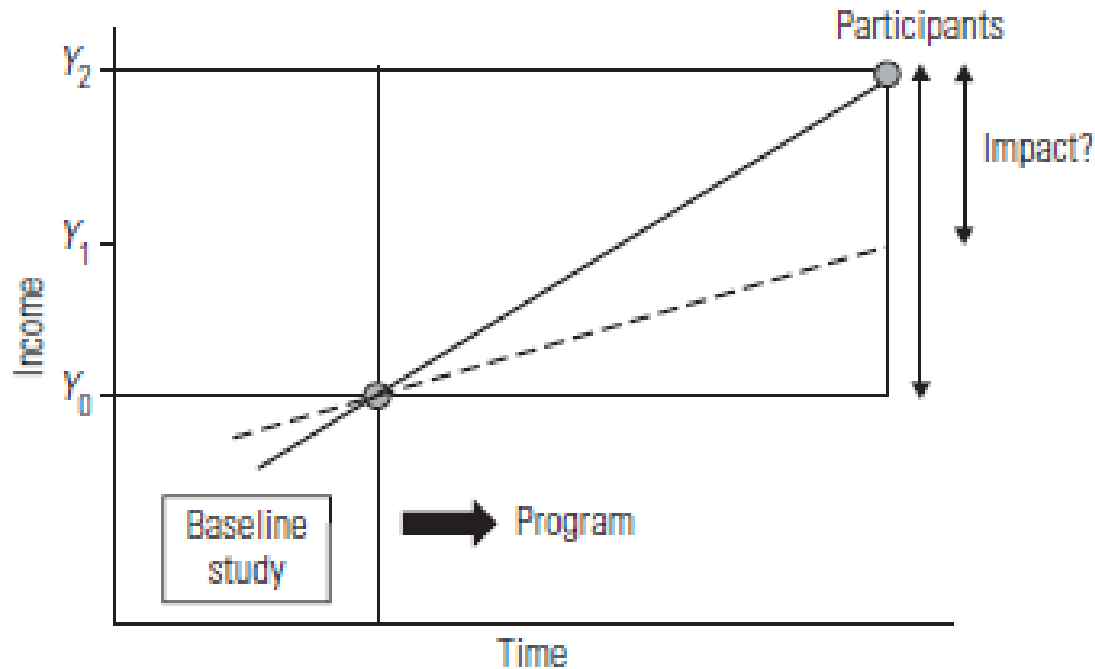
- Hypothesis: In **certain circumstances**, the use of Alkire-Foster multidimensional measures can provide **valuable insights** to the evaluation of social protection programs/policies.

# Impact Evaluation

- Impact evaluation assesses the changes in the well-being of individuals that can be attributed to a particular project, program, or policy.
- The problem of the counterfactual
  - What would have been the beneficiaries outcome in the absence of the intervention?
  - Need to create a convincing and reasonable comparison group
    - the control group.

# Impact Evaluation

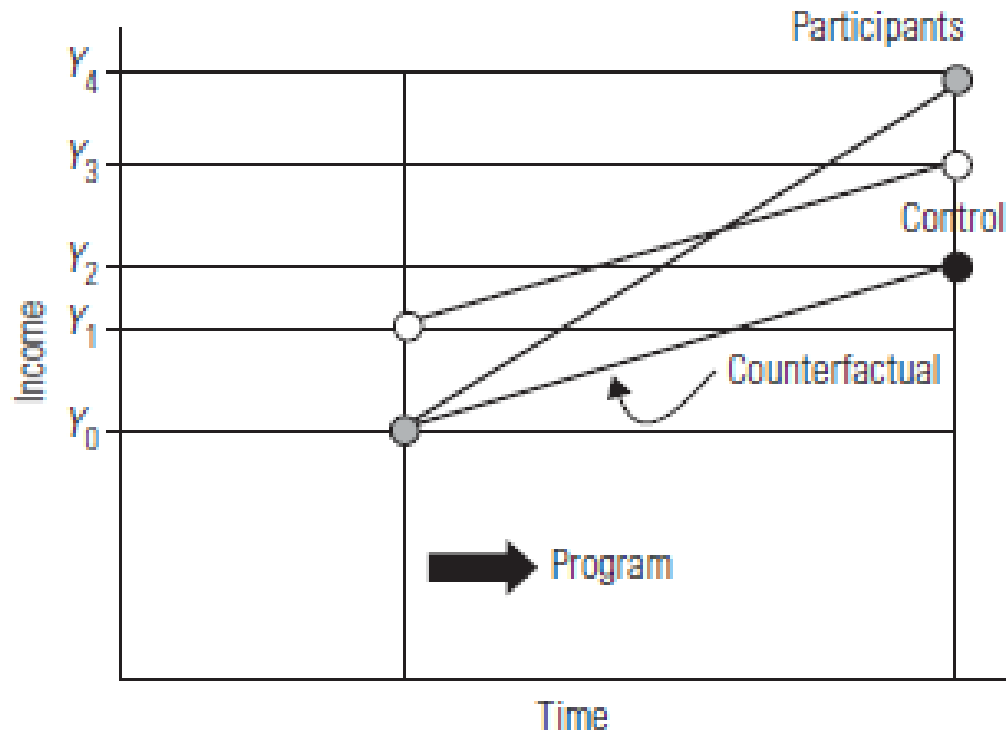
- Before-and-after comparisons



Source: Khandker, Koolwal and Samad (2010)

# Impact Evaluation

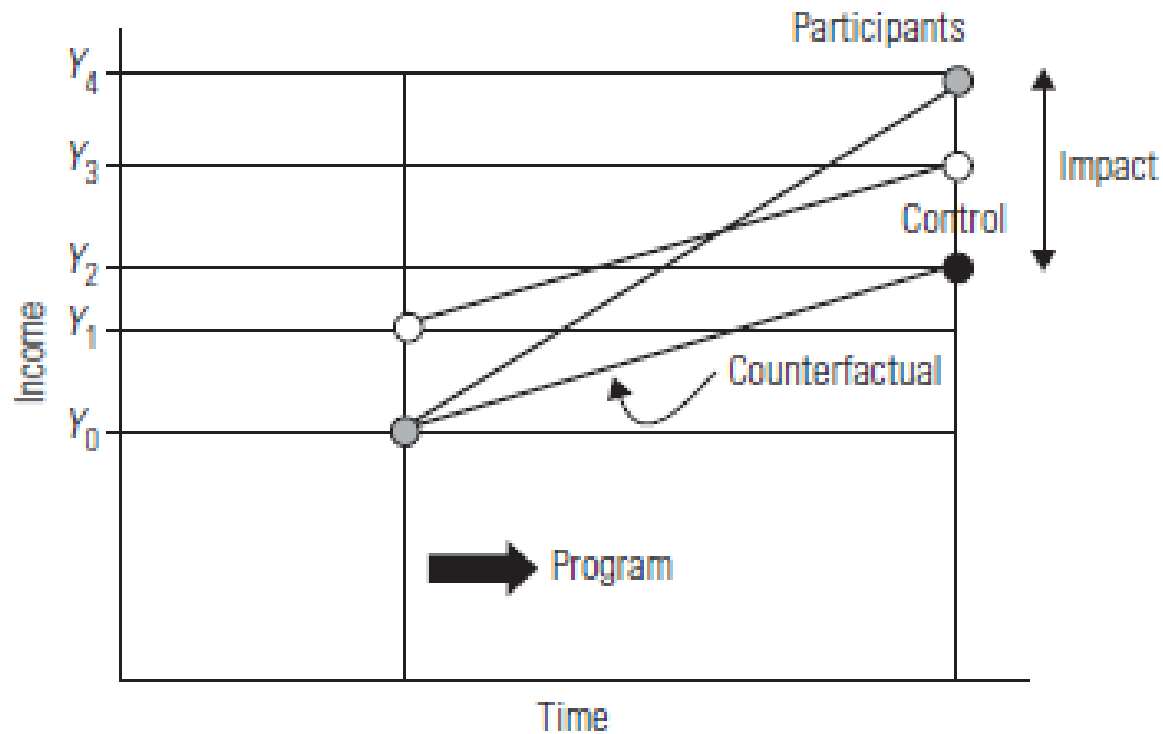
- With-and-without comparisons



Source: Khandker, Koolwal and Samad (2010)

# Impact Evaluation

- With-and-without comparisons



Source: Khandker, Koolwal and Samad (2010)

# Impact Evaluation

- The problem of selection bias

$$Y_i = \alpha X_i + \beta T_i + \epsilon_i$$

- Problem: if the treatment assignment is not random (program placement or self-selection into the program) and it is correlated with unobserved factors, the error term and  $T$  will be correlated and this will bias the estimate of the program effect  $\beta$ .
- The success of an impact evaluation depends on identifying a good comparison group, finding ways to get rid of the selection bias or finding ways to account for it.

# Impact Evaluation

- Methods commonly used to address the problem of the missing counterfactual:
  - Randomization
  - Matching methods, especially propensity score matching (PSM)
  - Differences-in-differences (DID)
  - Instrumental variable (IV)
  - Regression discontinuity design (RDD)



# Impact Evaluation

- Randomization:
  - Allocating a program or intervention randomly across a sample of observations.
  - Has been growing in popularity
  - Once the survey has been designed and the data collected, the empirical exercises to infer impacts are quite straightforward.
  - Concerns: ethical issues, external validity, partial or lack of compliance and spillovers.

# Impact Evaluation

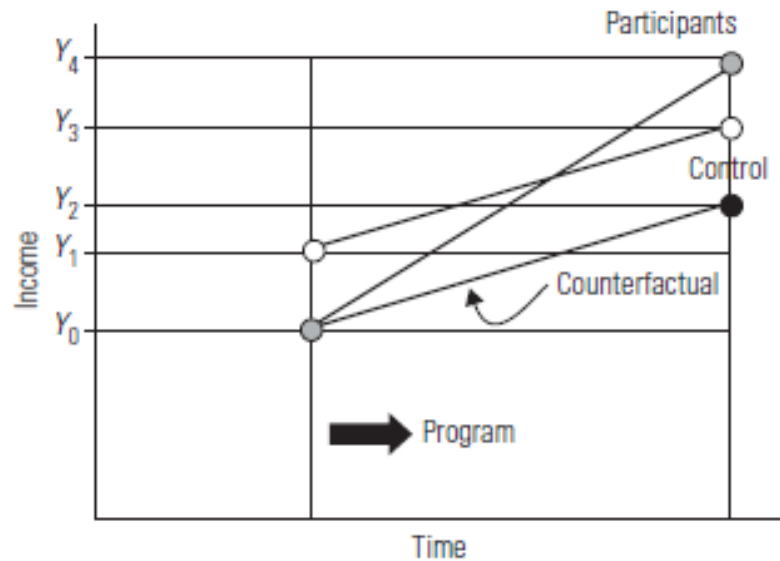
- Propensity Score Matching
  - For each participant and nonparticipant unit, it computes the probability that this unit will enroll in the program based on the observed values of its characteristics (the explanatory variables). Participants are then matched on the basis of this probability, or propensity score, to nonparticipants.
  - Validity depends on:
    - assumption that there are no unobserved characteristics that simultaneously affect program participation and outcomes;
    - sizable common support or overlap in propensity scores across the participant and nonparticipant samples.
  - Use bootstrapping to estimate standard errors

# Impact Evaluation

- Differences-in-Differences
  - Compares the changes in outcomes over time between participants and non-participants.

$$DD = E(Y_1^T - Y_0^T | T_1 = 1) - E(Y_1^C - Y_0^C | T_1 = 0)$$

- DD assumes unobserved heterogeneity is time invariant, so the bias cancels out through differencing.



# Impact Evaluation

- Instrumental Variable Estimation

$$Y_i = \alpha X_i + \beta T_i + \varepsilon_i$$

- Problem: if  $cov(T, \varepsilon) \neq 0$ , OLS estimates are biased
- The IV approach aims to clean up the correlation between  $T$  and  $\varepsilon$ . To do so, one needs to find an instrumental variable, denoted  $Z$ , that satisfies the following conditions:
  - Correlated with  $T$ :  $cov(Z, T) \neq 0$
  - Uncorrelated with  $\varepsilon$ :  $cov(Z, \varepsilon) = 0$
- 2 Stage Least Squares
- Common sources of IV: randomization, geographic variation, and exogenous factors affecting program placement.

# Impact Evaluation

- Regression Discontinuity Design
  - Can be used for programs that have a continuous eligibility index with a clearly defined eligibility threshold to determine who is eligible and who is not.
  - Estimates impact around the eligibility cutoff as the difference between the average outcome for units on the treated side of the eligibility cutoff and the average outcome of units on the untreated (control) side of the cutoff .
  - Because it estimates the impact of the program around the cutoff score the estimate cannot necessarily be generalized to units whose scores are further away from the cutoff score.

# Challenges

- Construct an appropriate AF measure
  - Data availability
  - Criteria to selection of indicators
- Potential for sample drop increases with the number of indicators included in the AF measure
- Potential for endogeneity issues may increase with the number of indicators included in the AF measure
- More limited scope for control and matching variables

# Examples

- Example 1: *Oportunidades*
  - <https://www.youtube.com/watch?v=g4S8GdMM-QA>
- Example 2: Shelter Assistance Program (Loschmann, Parsons and Siegel, 2015)

# Example: *Oportunidades*

- Why *Oportunidades*?
  - Pioneer in Conditional Cash Transfer Programs.
  - Multi-sector program:
    - Education,
    - Health,
    - Nutrition.
  - Experimental design:
    - Randomization of localities into control and treatment groups;
    - Data collected before and after the start of the treatment.



# Example: *Oportunidades*

- Impact of *Oportunidades* in single indicators documented:
  - Positive impact on enrolment (Schultz, 2000)
  - No impact on school attendance (Schultz, 2000)
  - Significant reduction in school grade gaps (Behram, Sengupta & Todd, 2000, 2005)
  - Positive impact on the number of grades completed (Behram, Parker & Todd, 2005)
  - Increase in number of visits to public health centres (Gertler, 2000)
  - Negative impact on probability of illness of children under 5 (Gertler, 2000)
  - Negative impact on children's labor (Parker & Todd, 2000)
  - Increase in food expenditure (Hoddinott & Skoufias, 2004)

# Example: *Oportunidades*

- Select indicators that:
  - Reflect the program's minimum goals;
  - Based on previous evaluation literature;
  - For which we have data for all time periods.
- Select weights

# Example: *Oportunidades*

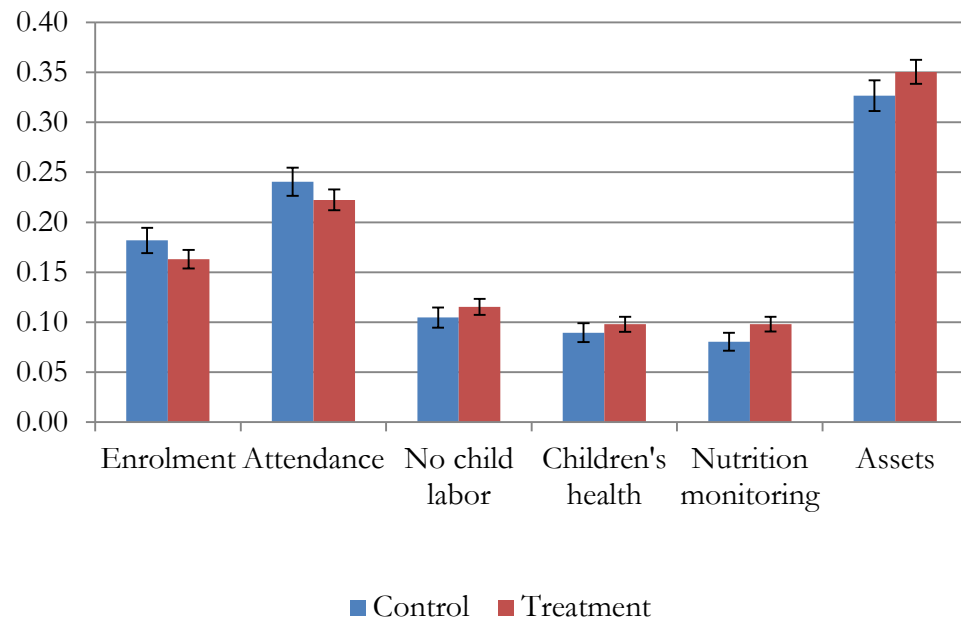
Indicator	Deprived if:	Weight
Enrolment	at least one member aged 6-14 not attending school (secondary school)	0.1
School attendance	at least one member aged 6-14 attended less than 90% of the school days (past month) OR is not enrolled	0.1
No child labor	at least one member aged 8-14 had a job or worked during last week (even if unpaid)	0.2
Children's health	at least one member aged 0-2 was ill in the past 4 weeks for more than 5 days	0.2
Health visits for nutrition monitoring	at least one member aged 0-2 has not made any visit in the past 6 months <sup>(1)</sup>	0.2
Assets	at least one of the following assets: refrigerator, television or radio	0.2

- **Problem:** Most indicators are defined with reference to children, the poverty status of the household is highly dependent on its the demographic structure.

# Differences at baseline?

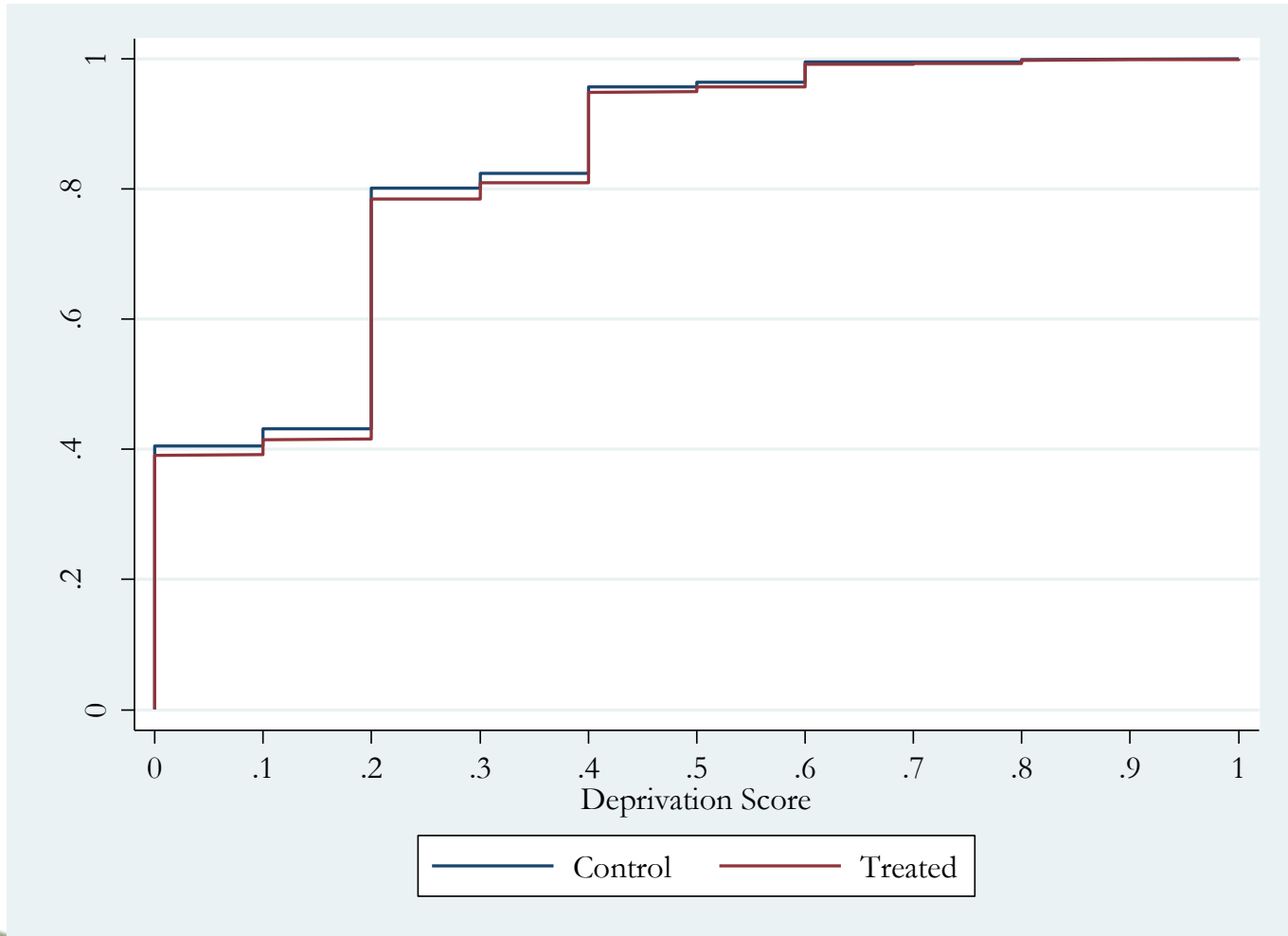
## Raw Headcount Ratios

	<b>Control</b>	<b>Treatment</b>
Enrolment	0.182	-0.019 (0.014)
Attendance	0.241	-0.018 (0.016)
No child labor	0.105	0.011 (0.012)
Children's health	0.090	0.008 (0.009)
Nutrition monitoring	0.080	0.018 (0.012)
Assets	0.327	0.024 (0.024)

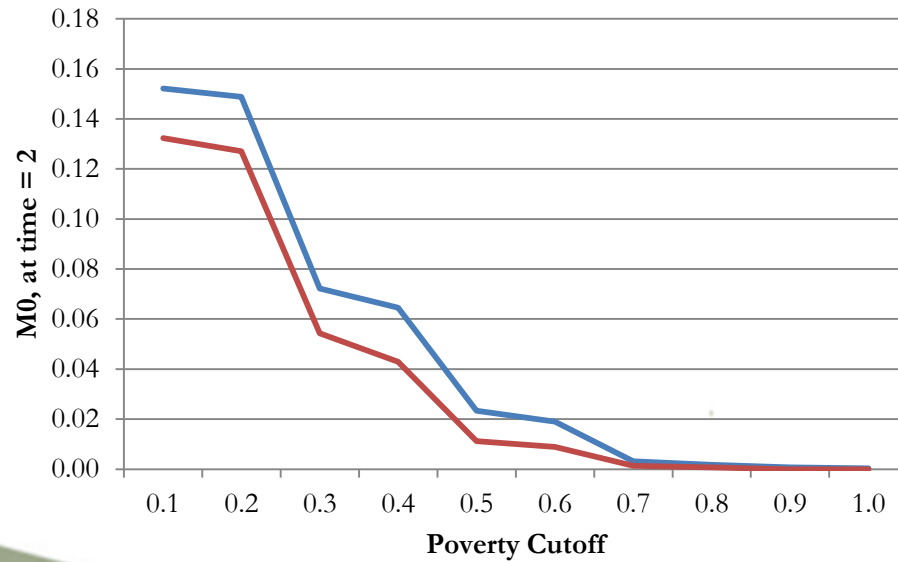
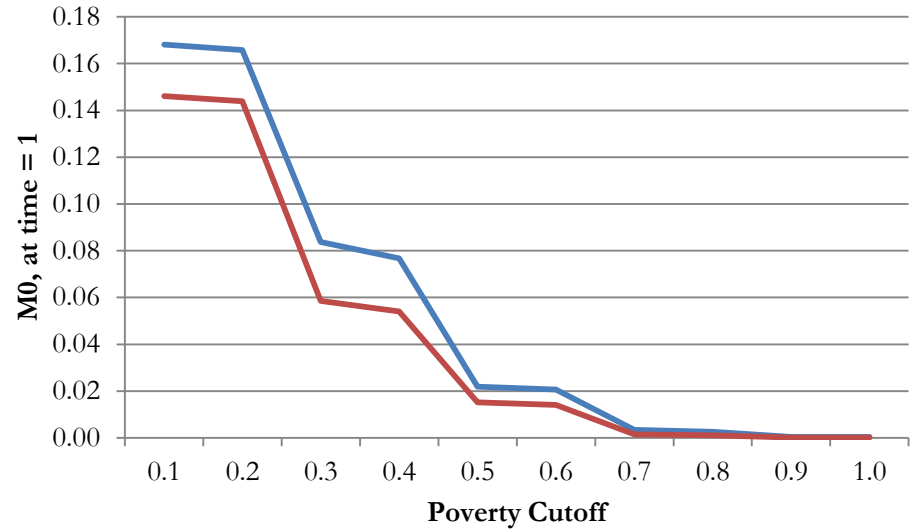
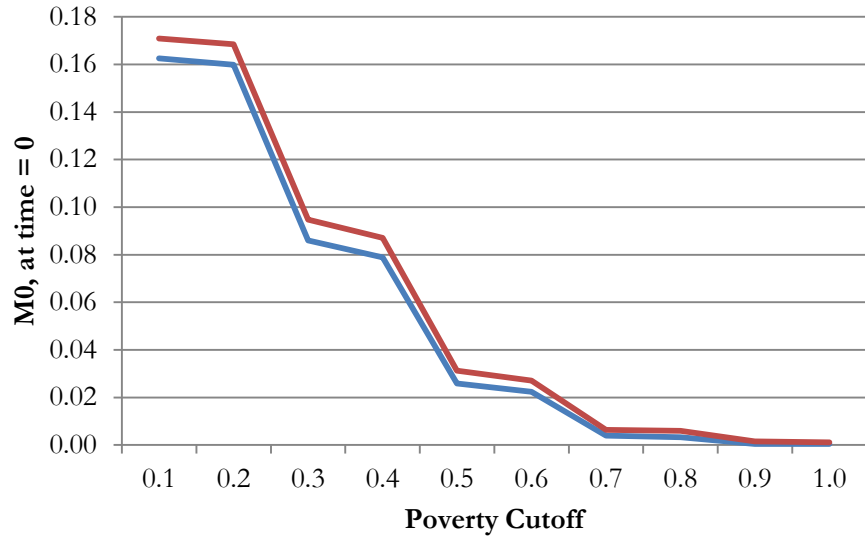


# Differences at baseline?

## Weighted Deprivation Score



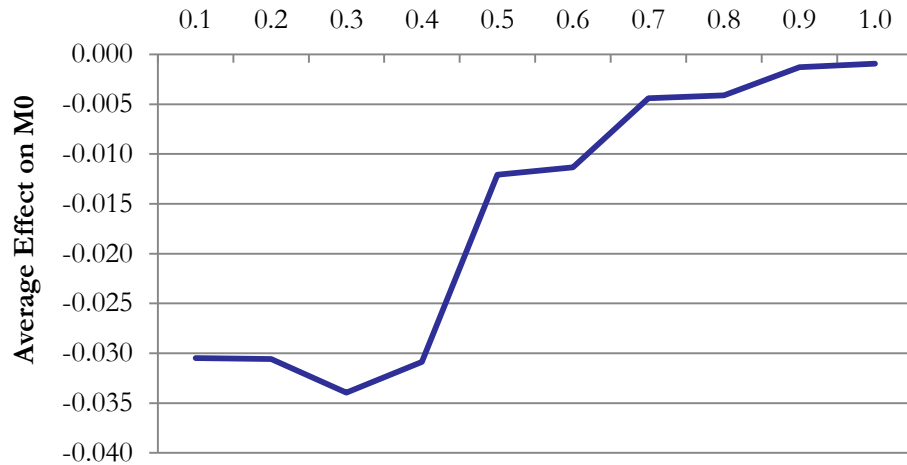
# Impact on Aggregated Measures



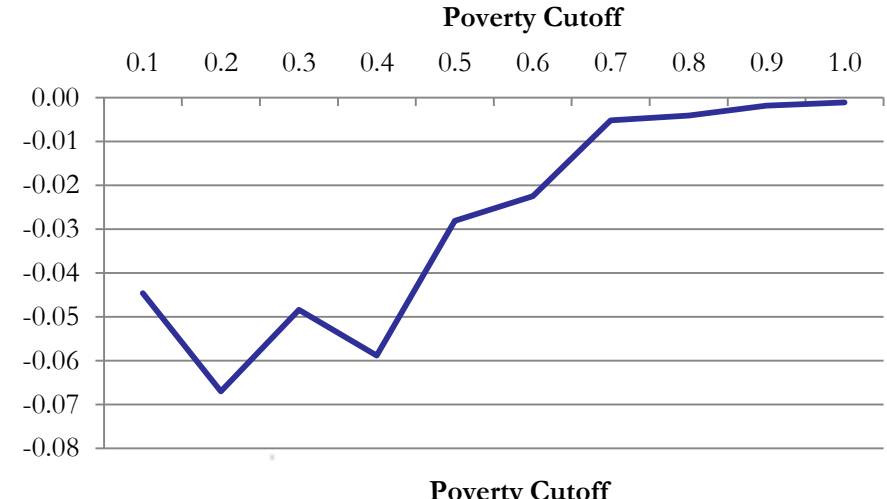
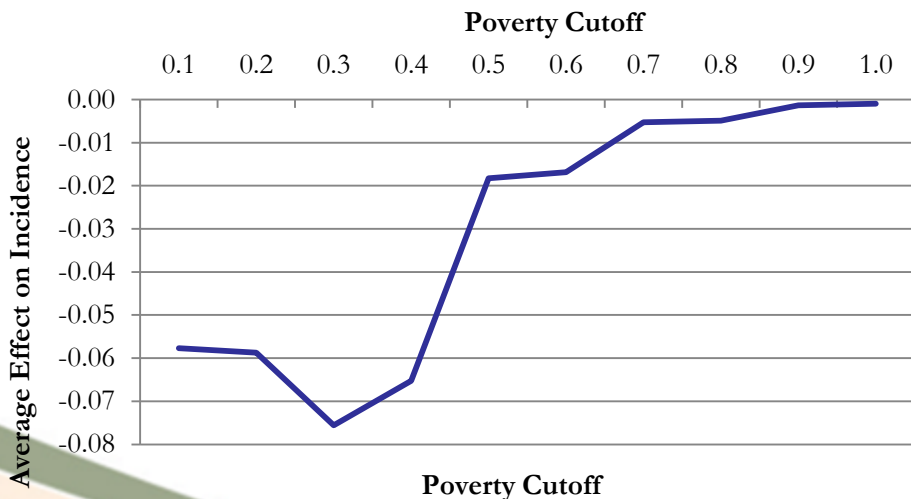
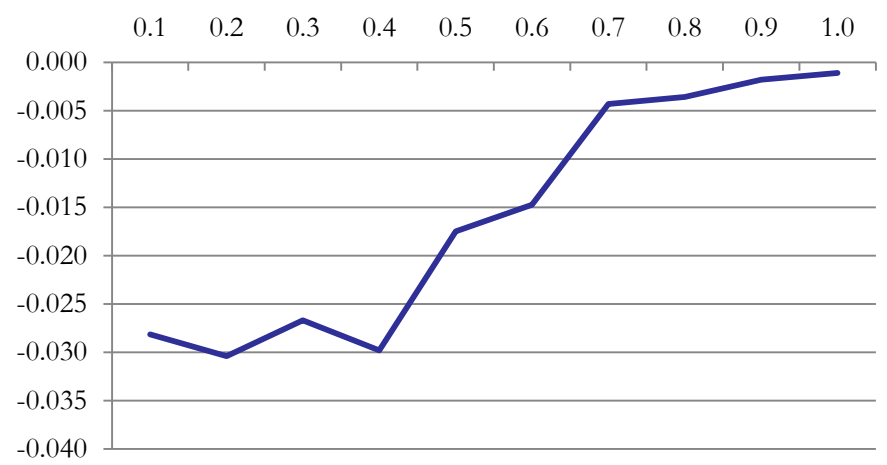
— Control — Treated

# Impact on Aggregated Measures

After 1 period:



After 2 period:



— Dif-in-Dif

— Dif-in-Dif

# Impact on Raw and Censored Headcounts

Time	Indicator	% Pop. Deprived in...			% Pop Poor & Deprived in...		
		Control	Treatment	Dif-in-Dif	Control	Treatment	Dif-in-Dif
0	Enrolment	18.2%	16.3%		10.2%	9.9%	
0	Attendance	24.1%	22.2%		13.4%	13.4%	
0	No child labor	10.5%	11.5%		8.0%	8.3%	
0	Children's health	9.0%	9.8%		5.5%	6.4%	
0	Nutrition monitoring	8.0%	9.8%		5.1%	6.8%	
0	Assets	32.7%	35.0%		12.6%	14.3%	
1	Enrolment	17.9%	11.9%	-0.041***	10.9%	7.2%	-0.034***
1	Attendance	23.0%	15.9%	-0.052***	13.6%	9.0%	-0.046***
1	No child labor	6.5%	5.2%	-0.024**	5.7%	4.2%	-0.018*
1	Children's health	5.9%	5.9%	-0.009	3.9%	3.2%	-0.016*
1	Nutrition monitoring	8.8%	5.1%	-0.054***	5.6%	3.2%	-0.039***
1	Assets	42.5%	43.0%	-0.019	14.5%	10.6%	-0.057***
2	Enrolment	17.7%	12.2%	-0.037***	9.6%	6.7%	-0.026***
2	Attendance	24.7%	21.8%	-0.011	13.3%	11.0%	-0.023
2	No child labor	6.3%	5.3%	-0.021*	5.6%	3.9%	-0.020*
2	Children's health	5.6%	4.4%	-0.020*	3.4%	2.5%	-0.017**
2	Nutrition monitoring	5.5%	2.3%	-0.049***	3.7%	1.5%	-0.038***
2	Assets	37.5%	37.2%	-0.026	12.0%	10.3%	-0.034***



# Impact –Deprivation Score

$$C_{i,t} = \alpha + \beta_1 T_i + \beta_2 P1_i + \beta_3 P2_i + \beta_4 T_i P1_i + \beta_5 T_i P2_i + \varepsilon_i$$

	Deprivations count
Treatment	0.008 (0.008)
Period 1	0.006 (0.005)
Period 2	-0.010** (0.005)
Treatment x Period 1	<b>-0.030***</b> (0.006)
Treatment x Period 2	<b>-0.028***</b> (0.007)
Constant	0.163*** (0.006)
Observations	28,932
F-statistic	24.39***
R-Squared	0.0081

# Example 2: Shelter Assistance

# Steps

1. Introduction
2. Program / Policy
3. Evaluation Design
  - 3.1 Sample
  - 3.2 Outcome variables
  - 3.3 Empirical Strategy
4. Results
  - 3.1 Effects
  - 3.2 Analysis of the effects
5. Conclusion

# Steps

## 1. Introduction

- Research question or purpose of the study.
- Why is this question or issue important?
- How the study contributes to the body of knowledge?  
Including methodological and empirical contributions
- Identify avenues for further research

## 2. Program / Policy

- Background / context
- Objectives
- Implementation

# Steps

## 3. Evaluation Design

### 3.1 Sample

- (Examples of datasets that can potentially be used after the References)

### 3.2 Outcome variables

- Brief explanation of AF measures

### 3.3 Empirical Strategy

- Method used
- Justification
- Limitations

# Steps

## 4. Results

### 4.1 Effects

### 4.2 Analysis of the effects

- Additional analysis that can give insights on how the program works

### 4.3 Robustness checks

## 5. Conclusion

# References

- Gertler, P. J., S. Martinez, P. Premand, L. B. Rawlings and C. M. J. Vermeersch (2016). *Impact Evaluation in Practice*, Second Edition. The World Bank, Washington, DC.
- Khandker, S., G. B. Koolwal and H. A. Samad (2010). *Handbook on Impact Evaluation – Quantitative Methods and Practices*. The World Bank, Washington. DC.
- Loschmann C., C. R. Parsons and M. Siegel (2015). “Does Shelter Assistance Reduce Poverty in Afghanistan?” *World Development*, 74, 305-322.
- Robano, V., S. C. Smith (2014). “Multidimensional targeting and evaluation: A General Framework with an Application to a Poverty Program in Bangladesh”. *OPHI Working Paper* 65.

# List of Datasets

- Conditional Cash Transfer Program in Kazakhstan  
(OPM): [http://microdata.worldbank.org/index.php/catalog/2218/related\\_materials](http://microdata.worldbank.org/index.php/catalog/2218/related_materials)
- Social Assistance Grants for Empowerment in Uganda  
(OPM): <http://microdata.worldbank.org/index.php/catalog/2654/study-description>
- Hunger Safety Net Programme in Kenya  
(OPM): <http://microdata.worldbank.org/index.php/catalog/1917/study-description>
- Community Monitoring for Better Health and Education Services in Burkina Faso  
(ISSP): <http://microdata.worldbank.org/index.php/catalog/1550/study-description>



# List of Datasets

- National Community Empowerment Programme in Indonesia  
(WB): <http://microdata.worldbank.org/index.php/catalog/1049/study-description>
- Zomba Cash Transfer Program in Malawi  
(WB): <http://microdata.worldbank.org/index.php/catalog/1049/study-description>
- Social Action Fund for Vulnerable Groups  
(WB): <http://microdata.worldbank.org/index.php/catalog/1004/study-description>
- Upgrading Informal Settlements Programme in South Africa  
(WB): <http://microdata.worldbank.org/index.php/catalog/1041/study-description>

# List of Datasets

- Community Development Grants in the Philippines  
(MCC): <http://microdata.worldbank.org/index.php/catalog/2298/study-description>
- Cash and Food Transfers in Uganda  
(IFPRI): <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/3REX7R>
- Cash, Food Vouchers and Food Transfers to Colombian Refugees and Poor Ecuadoreans  
(IFPRI): <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/AXGCHT>
- Commitment Savings Products in Philippines  
(IPA): <https://www.povertyactionlab.org/evaluation/commitment-savings-products-philippines>

# List of Datasets

- Primary School Deworming in Kenya  
(JPAL): <https://www.povertyactionlab.org/evaluation/primary-school-deworming-kenya>.
- Measuring the Impact of Microfinance in Hyderabad, India  
(JPAL): <https://www.povertyactionlab.org/evaluation/measuring-impact-microfinance-hyderabad-india>
- Project Generasi: Conditional Community Block Grants in Indonesia  
(WB): <https://www.povertyactionlab.org/evaluation/project-generasi-conditional-community-block-grants-indonesia>
- Graduating the Ultra-Poor in Peru  
(IPA): <https://www.povertyactionlab.org/evaluation/graduating-ultra-poor-peru>

# List of Datasets

- Graduating the Ultra-Poor in Pakistan  
(IPA): <https://www.povertyactionlab.org/evaluation/graduating-ultra-poor-pakistan>
- Returns to Secondary Schooling in Ghana  
(JPAL): <https://www.povertyactionlab.org/evaluation/returns-secondary-schooling-ghana>
- Impact of Female Leadership on Aspirations and Educational Attainment for Teenage Girls in India  
(JPAL): <https://www.povertyactionlab.org/evaluation/impact-female-leadership-aspirations-and-educational-attainment-teenage-girls-india>
- Assessing the Impact of Microcredit in Ethiopia  
(JPAL): <https://www.povertyactionlab.org/evaluation/assessing-impact-microcredit-ethiopia>

# List of Datasets

- The Impacts of Microcredit: Evidence from Bosnia and Herzegovina (JPAL): <https://www.povertyactionlab.org/evaluation/impacts-microcredit-evidence-bosnia-and-herzegovina>
- The Oregon Health Insurance Experiment in the United States (JPAL): <https://www.povertyactionlab.org/evaluation/oregon-health-insurance-experiment-united-states>
- Graduating the Ultra-Poor in Ethiopia (IPA): <https://www.povertyactionlab.org/evaluation/graduating-ultra-poor-ethiopia>
- The Impact of Employment on High-Risk Men in Liberia (IPA): <https://www.povertyactionlab.org/evaluation/impact-employment-high-risk-men-liberia>
- Graduating the Ultra-Poor in Ghana (IPA): <https://www.povertyactionlab.org/evaluation/graduating-ultra-poor-ghana>

Tabita, Kenya

Rabiya, India

Stephanie, Madagascar

Agatha, Madagascar

Dalima, Kenya

Ann-Sophia, Kenya

Valerie, Madagascar



*Thank you!*