

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/333677098>

# Journal of Development Economics Registered Report Stage 1 The impact of a multidimensional program on nutrition and poverty in Burkina Faso – PRE-ACCEPTED

Preprint · June 2019

CITATIONS

0

READS

471

2 authors, including:



**Adrien Bouguen**

University of California, Berkeley

17 PUBLICATIONS 30 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



L'assignation aléatoire comme méthode d'évaluation des politiques publiques [View project](#)



Early Child Care Cambodia [View project](#)

Journal of Development Economics Registered  
Report Stage 1: Proposal The impact of a  
multidimensional program on nutrition and  
poverty in Burkina Faso - REVISED

04/22/2019

**Abstract**

This study adapts BRAC's ultra-poverty program to focus on child nutrition, household food security and resilience, and evaluates its impact in two regions of Burkina Faso. In our study, we test the hypothesis that a targeted cash transfer, asset transfer, and nutrition-focused project can transform the early childhood environment to reduce poverty and malnutrition. We expect to test whether the program has effects on poverty and productivity as well as children's anemia level, stunting and, after two years of program implementation, young children's cognition. This cluster randomized-control trial also seeks to assess the independent contributions of assets and nutrition interventions, potential spillover of such policy on other individuals in the village and the cost-effectiveness of such policies in the long-run. The target populations are ultra-poor and poor households with a breastfeeding/pregnant or malnourished child under 5 years old. Villages were first randomly assigned to treatment and control branch (43 in control) and second, treatment villages were assigned to either unconditional cash transfer (T1), a cash and asset transfer branch (T2), or a cash, asset transfer as well as a nutritional focused interventions (T3).

**JEL classification:** I15, I32, I31

**Keywords:** Multi-faced programs, cash transfer, asset transfer, nutrition, ultra-poor, poverty alleviation programs

**Study pre-registration:** AEA registry, The impact of a multidimensional program on nutrition and poverty in Burkina Faso, registration is ongoing.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Proposed timeline</b>	<b>4</b>
<b>3</b>	<b>Research Design</b>	<b>5</b>
3.1	Interventions . . . . .	5
3.2	Basic methodological framework / Identification strategy . . . . .	6
3.3	Sampling . . . . .	7
3.3.1	Sampling protocol . . . . .	7
3.3.2	Sampling strategy and estimation . . . . .	8
3.4	Statistical Power . . . . .	9
<b>4</b>	<b>Instruments and Surveys</b>	<b>12</b>
4.1	Household Economic Assessment (HEA) . . . . .	12
4.1.1	Census . . . . .	13
4.1.2	Community Based Classification (CSE) . . . . .	13
4.1.3	Using census and CSE to determine eligibility . . . . .	14
4.2	Baseline . . . . .	14
4.3	Follow and Endline surveys . . . . .	15
4.4	Data Collection . . . . .	16
4.5	Data Processing . . . . .	16
4.6	Variations from the intended sample size . . . . .	16
<b>5</b>	<b>Empirical Analysis</b>	<b>17</b>
5.1	Program’s compliance . . . . .	17
5.2	Intent-to-treat . . . . .	19
5.3	Heterogeneous Effects . . . . .	22
5.4	Multiple outcome and multiple hypothesis testing . . . . .	24
5.5	Specification . . . . .	24
<b>6</b>	<b>Missing values, attrition and outliers</b>	<b>25</b>
<b>7</b>	<b>Declaration of interest</b>	<b>25</b>
<b>8</b>	<b>Acknowledgments</b>	<b>26</b>

# 1 Introduction

Burkina Faso has experienced strong economic growth, however, this growth has not benefited all households. The number of poor in the country remain very high and inequality has risen. Besides, The Sahel environment leaves households vulnerable to food insecurity and shocks.

Investing in regions in a situation of fragility and conflict has been one of the key priorities of European Union (EU) development policy over the last decades. In Burkina Faso, food and nutritional security is one of the sectors prioritized by the National Indicative Program (Programme Indicatif National) for 2014 - 2020 between the EU and Burkina Faso (European Union - Burkina Faso).

Evidence from multi-country studies, Ghana (Banerjee et al., 2015) and from Bangladesh (Bandiera et al., 2017) suggests that a multifaceted programs under a graduation model can lead to consistent and lasting impacts on consumption, food security, assets ownership, household income and revenues. The aim of this study funded by the EU Trust Fund, is to adapt BRAC's approach to design a multidimensional project focused on child nutrition, household food security and resilience. Our approach differs from previous studies as it offers an opportunity to appraise the impact of the intervention package on households with pregnant women or young children. The study therefore looks at the effect of cash, asset and nutrition (small garden+ fortified floor distribution) transfer program on poor households with children aged less than 5 years old or/and pregnant/breast-feeding women. We hypothesize that multifaceted programs such as the one developed by BRAC, could, if targeted to households with young children or pregnant women, fundamentally transform the early environment of the children, reducing poverty and malnutrition which could improve children's cognitive skills. To our knowledge, the available literature on multifaceted program has not looked carefully at this important causal pathway directly. An earlier literature that debated estimates of the calorie-income elasticity generally concluded that income growth alone does not address malnutrition.

The study aims at providing evidence on the following research questions:

- First, does a multidimensional nutrition-focused program targeting poor households with pregnant women or young children is instrumental in transforming the early environment of young children in their first 1000 days ? Although several articles have already shown the relevance of the graduation approach, we are here interested in one of the potentially overlooked aspects of graduation program: by providing at the same time, cash, asset and other health service, graduation programs are likely to fundamentally transform early investment of parents in nutrition, education or health. Relatedly, such a program is likely to improve resilience and increase food security.

- Second, what is the relative effect of each intervention in comparison to cash transfer? Conditional or unconditional cash transfer programs have already shown their effectiveness in increasing enterprise activities Blattman et al. (2014, 2018) or well-being (Haushofer and Shapiro, 2016). What is the value added of assets and nutrition interventions in this context? What is the income elasticity of calories? Of food with high nutritive content for young children and pregnant women (with high iron or vitamin)? Is cash enough to improve nutrition? What is the value of the additional nutrition intervention (with small garden interventions and fortified iron flour)? This study design looks at fortified commodity transfers to households as well as asset transfer relative to cash. Disentangling how household use the transfers and their impacts on poverty and nutrition will provide insight into the mechanisms through which households exit extreme poverty.
- Thirdly, what are the potential spillover effects on other households in the village? Positive and/or negative spillover may be envisioned. Previous studies (Bandiera et al., 2017) point out how such programs allowed the poorest to access occupations that were only available to wealthier individuals, while increasing the economic trade between rich and poor (increased amount of loans between wealthiest and poorest). Increase in market price of assets and consumption goods may be one of the potential negative sides of the intervention as reported recently (Filmer et al., 2018). Investigating precisely how such large transfers may impact the life of the non-eligible individuals in the eligible village is of prime concern.
- Fourthly, are these social investments sustainable in the long run? What is the cost-effectiveness of such policy?

## 2 Proposed timeline

The main milestones of the study are the following:

Activities	Period
Selection of the 85 villages per Region	August 2017
Randomization (T/C)	August 2017
Household Economic Assessment	October 2017 – November 2017
Baseline	June 2018 – July 2018
Program’s implementation	July 2018 – September 2018
Midline	January 2019
Program’s implementation	February 2019 – September 2019
Endline	January 2020

The Household Economic Assessment (HEA) is a participatory targeting approach used to identify the households eligible in the villages. The aim of the HEA

is to identify the ultra-poor and poor households in each of the treatment villages. To this purpose, we first conducted a census in each village to collect socio-economic information on households and hard-to-hide information (e.g. animals, type of houses). Then, we implemented a village survey where we asked villagers to identify key poverty criteria. Villagers ranked each village member into poor, ultra-poor, average or wealthier based on their knowledge of the local context. Afterwards, we used census data to assert the quality of villagers ranking. The resulting ranking is finally used to target the interventions to the ultra-poor and poor households and identified the eligible households.

Baseline sample is conducted after the HEA essentially because we used the HEA to constitute both the baseline sample and the eligibles. Yet at baseline, none of the households knew about their actual treatment status. The only information household members may have inferred is whether or not they were classified as poor before baseline, which give them only imperfect information about their future treatment status.

The follow-up and the endline surveys will be consistent with the baseline survey instrument. The follow-up surveys will also include sections on the implementation of the program to verify compliance with the treatment and household composition changes.

## 3 Research Design

### 3.1 Interventions

The interventions are implemented by our two partners Action contre la Faim (ACF) and Terres des Hommes (TDH). The study comprises 168 villages in two regions (Boucle du Mouthon and East). In each region, villages are randomized to three treatment groups and one control group. The distribution for households and villages is presented in the table below:

Table 1: **Household and Village Distribution per type of intervention**

	T1/ F (42 villages)	T2/FA (41 villages)	T3/FAN (42 villages)	C (43 villages)
F (Finance/Cash)				
A (Asset)				
N (Nutrition)				

The treatment branches correspond to the following interventions:

- Finance: Cash transfer of 20,000 FCFA per month for the first year of intervention and 15,000 FCFA for the second year plus training on cash utilization. Cash is distributed four times per year
- Asset: Distribution of: 3 Ruminants/ 11 Chicken/5 pigs + Winter seeds kits

- Nutrition: Distribution of: Garden kit and Enriched Flour

Each village is randomly assigned to one of the treatment branches in three steps. First, the research team randomly assigns the villages in treatment and control. Second, a HEA is organized to determine the eligibility status. Third, the treatment group is divided into three treatment groups (F, FA and FAN).

Furthermore, some interventions implemented at the commune level will benefit all villages, control or not. These policies affect similarly all experimental groups and hence will not be analyzed independently from the other components. It includes the following interventions:

- Capacity Building: Training linked to the different interventions, e.g. Nutrition
- Early Alert: Information given to the eligibles that help prevent events with a negative impact, e.g. strong rains.
- Accountability: Training for local representatives to enhance management transparency and induce them to convey more information to their fellow citizens.
- Social Cohesion: Activities to avoid conflict within the community around certain sensitive issues, e.g. water sharing.
- Integration of Policies: Training for local representatives so they include the different interventions in their communities' management plans.
- Capacity Building UPA: Training for food production units (UPA) so they can increase their output.
- Husband school: A space to generate husband's interest in the treatment received by their wives.
- Facility to wash your hands: wash hand basin in the health care centers
- Health care centers: Facilities that provide information/advice on health and nutrition related issues to breastfeeding or pregnant women.
- Warrantage: Rural financial system consisting in storing harvests in warehouses for them to be sold at a later point in time when prices are high.

The design presented above in table 1 allows identifying the independent effect of the cash intervention, the marginal effect of the Assets intervention when combined with the cash interventions and the overall effect of the program when the three interventions are combined.

### **3.2 Basic methodological framework / Identification strategy**

The multi-dimensional poverty alleviation program will be evaluated based on a cluster-randomized controlled trial (RCT). In this methodology, villages are allo-

cated, before the beginning of the interventions, in different experimental groups, one of them being a pure control group. Given that many interventions will not, for financial or practical reasons, benefit to all target villages in a region, randomizing the interventions is a transparent way to allocate the eligibles. The randomization is conducted in two phases:

1. Before HEA, we assign the villages into the treatment and the control group. We had to assign randomly in T and in C before conducting the HEA because the HEA was conducted by different institutions in the treatment and in the control group. IPA implemented the HEA in the control group while the partners ACF or TDH implemented the HEA in the treatment group. The result of the randomization was however not communicated to the villages.
2. After the HEA, we used the HEA data to randomize between treatment group (T1, T2, T3).

### 3.3 Sampling

#### 3.3.1 Sampling protocol

The study will be conducted in two regions of Burkina Faso: the Boucle du Mouhoun region and the East region. Primary subjects of the study are children and households across the two selected regions. The partners (ACF TDH) are in financial capacity to treat an average of 21 households per villages (about 3500 households in the treatment locations). We hence try as much as possible to identify 21 eligible households per treatment villages. To measure spill over, we also include 21 households non-eligible to the program. Finally, we included 5 additional eligible and 5 additional non-eligible households in each village in a reserved list. These additional households are surveyed only in the cases when the households refuse the treatment, the households could not be found or the households left the village during the programs' implementation.

To benefit from the program (see section 4.1 for more detailed on the HEA), a household is either Ultra-Poor (UP) or Poor (P) and must be a Nutritional Vulnerable household (NV) as determined during the HEA.<sup>1</sup> We define as NV every households that have a breastfeeding mother, a pregnant mother or a child under 5 year. Therefore we can define the following household types:

---

<sup>1</sup>If there are more than 21 eligible households in the village, we selection the 21 poorest (using a wealth index). If there are less than 21 eligible households, we report the available spots to the villages that have eligible households. Hence, the number of eligible households per village fluctuates. This is to insure that the partners treat enough households in total



Table 2: **Household Categories**

	<b>NV households</b>	<b>non-NV households</b>
Ultra-Poor	(1) UP-NV	(2) UP-nNV
Poor	(3)PNV	(4) PnNV
Medium	(5) MNV	(6) MnNV
Wealthy	(7)WNV	(8)WnW

The eligible households (as well as the eligibles on reserved list) are taken from groups (1) and (2), while the non-eligibles (as well as the non-eligibles on reserved list) are taken from groups 2 and 4. To sum up the sampling will respect the following structure:

- 21 of the poorest UP and P households will be sampled in treatment and control to participate in the survey. Depending on the result of the HEA, this population should essentially be composed of ultra poor and nutritional vulnerable households and maybe of some poor and nutritional vulnerable households. Five households selected among the 5 next poorest in the ultra-poor and nutrition group will be put in the reserved list.
- 21 non-eligibles taken from the poorest of the remaining households not included in the eligible group (or its reserved list). This group will include NV and non-NV households and may include middle or well off households. As for the eligibles, the 5 next poorest will be included in a non-eligible reserved list.

### 3.3.2 Sampling strategy and estimation

The sampling strategy we adopted here is different from the one adopted in most of the multifaceted studied before (Bandiera et al., 2017; Banerjee et al., 2015). In deed, in most previous research, authors used the HEA (see below section 4.1 HEA) to identified the eligibles (usually poor and/or ultra poor) and then randomly selected the non-eligible among the remaining categories (Poor, medium or wealthy). There are important difference in our program that made that approach non viable. Conversely to previous research, the objective of the project is to improve early environment, nutrition and resilience (to shocks). While previous research were particularly interested in Ultra poor, our research target the nutrition vulnerable households. Given this program’s constraints we had two alternative choices to select the non-eligibles: (i) We could have randomly selected the non-eligibles and eligible among the NV households (poor and ultra poor with a pregnant or breastfeeding women or a child below 5) (ii) we could have randomly selected the non-eligibles among all the remaining households (i.e. the ones not selected in the eligible). We discard both alternatives for the following reasons.

**About (i):** We actually discussed with our partners (and within the research team) of the possibility to randomly assigned eligibles within villages but rapidly

discarded this possibility. This for two important reasons. First, since our partners, in absence of the program's evaluation, would have selected the poorest households among the Nutritional Vulnerable households categorized as Poor and Ultra, randomly assigning eligible households within villages would have raised important external validity issues. Second, given the (small) size of the village and the (large) size of the household, in the majority of villages, less than 42 NV households were available. As a result, if we had selected the 21 eligible randomly, we would have significantly reduced the size of the non-eligible sample. To insure that we could have an average of 21 non-eligible household per village, we had to include Poor and Ultra poor households that were not NV.

**About (ii):** We also discarded the possibility to randomly assigned non eligible among the remaining households (without any poverty restriction). Conversely to Bandiera et al. (2017), we are not primary interested in economic spill-overs: our research questions focus on nutrition and early environment. Since nutrition and early environment is our keys outcome, we believe that measuring the spill-overs on wealthy households would have been of lesser interest: on them, we do not expect much impact simply because children in wealthier households are probably already benefiting from a favorable early environment. What we are interested in is how a nutrition program such as this one may indirectly impact the early environment of families which are not too dissimilar from the eligible.

In summary, our approach consisted in selecting the peer group that was the closest to our group of eligible in terms of characteristics: we believe that this group is the most likely to be indirectly affected by the program through negative spillovers (price of food) or positive spillovers (information, positive general equilibrium effect through labor demand...)

### 3.4 Statistical Power

As explained in section 3.3, the sample is composed of four different categories of households determined by the HEA: the eligible households, the non eligible households, the eligible household on reserved list, non eligible household on reserved list. The overall sample size is hence : 21 eligibles+5 additional eligibles+21 non-eligibles + 5 additional eligibles = 52 households per village i.e.  $52 \times 168 = 8736$  households.

The calculation is based on a recent baseline study about agriculture conducted last January in Burkina Faso. From that survey, we extracted the above statistics:

- M corresponds to the mean
- SD the standard deviation of the variable
- ICC ( $\rho$ ) the intra-cluster correlation (village level as well).

Table 3: Minimum detectable effects, by region and group

	M	SD	All T versus C			One branch versus control		
			No baseline	Baseline	% change	No baseline	Baseline	% change
1 region								
Standardized effect	0.00	1.00	0.243	0.217		0.305	0.273	
Binary variable (no wage revenu)	0.50	0.50	0.122	0.109	21.65%	0.152	0.136	27.14%
Binary variable (secondary educatio	0.06	0.24	0.059	0.053	83.41%	0.074	0.067	104.56%
Food security index	3.32	1.87	0.455	0.407	12.24%	0.570	0.510	15.34%
Number of animal	30.94	35.79	8.699	7.781	25.15%	10.905	9.754	31.52%
2 regions								
Standardized effect	0.00	1.00	0.170	0.152		0.214	0.192	
Binary variable (no wage revenu)	0.50	0.50	0.085	0.076	15.13%	0.107	0.096	19.08%
Binary variable (secondary educatio	0.06	0.24	0.041	0.037	58.29%	0.052	0.047	73.51%
Food security index	3.32	1.87	0.318	0.284	8.55%	0.401	0.358	10.78%
Number of animal	30.94	35.79	6.079	5.438	17.57%	7.666	6.857	22.16%
2 regions - 21 households								
Standardized effect	0.00	1.00	0.189	0.169		0.239	0.214	
Binary variable (no wage revenu)	0.50	0.50	0.095	0.085	16.87%	0.119	0.107	21.27%
Binary variable (secondary educatio	0.06	0.24	0.046	0.041	64.98%	0.058	0.052	81.94%
Food security index	3.32	1.87	0.354	0.317	9.53%	0.446	0.399	12.02%
Number of animal	30.94	35.79	6.777	6.061	19.59%	8.546	7.644	24.70%

The table above gives the Minimum detectable effect (MDE) expressed in a standardized fashion (first row) or using the baseline variables (wage, education, food security index or number of animal). To calculate MDE, we use a type I error of 5% ( $\alpha$ ) a type II error of 80% ( $\beta$ ) and the following formula:

$$MDE = \frac{t_{1-\alpha/2} + t_{1-\beta}}{\sqrt{(P(1-P)N)}} * \sqrt{\rho + \frac{1-\rho}{n}} * \frac{1}{c} * (1 - R^2)\sigma$$

With P the treatment probability, t the t-value from a Student distribution, c compliance rate (the share of treatment households who are treated minus the share of control households who are treated). We do not believe that any control households will be treated and we believe most people will accept cash/asset or nutrition intervention. We hence set the compliance rate at a high level of 90%.  $R^2$  corresponds to the R square of the regression using baseline characteristics (set at 20% again from a recent study conducted in Burkina Faso),  $\sigma$  is 1 for the standardized case and replaced by the standard deviation of the variables in the other cases and  $\rho$  is set to 10% (again an average value from a previous study in Burkina Faso).

The calculations are based on a sample of 42 households per village, except for the last row that looks at the effect on the sub-sample of eligible (21 households). Minimum detectable effects are given for a standardized index as well as for four variables taken from a recent baseline survey implemented in Burkina Faso. Table 2 above indicates that when both regions are analyzed jointly, the sample size allows detecting small effects, generally below 20% SD, even for the sample of eligible (last columns, last rows). For the eligible sample of 21 eligible households and when

looking at all treatment branches together versus control, the MDE corresponds (with covariates) to an effect of about 9 pp (a 17% effect change) on a binary variable such as “having no wage revenue” or a 4pp on a less likely binary variable such as having a secondary education (+65% increase). On continuous outcome, such as the food security index or the number of animal owned, the MDE is about +10% change (food security) or + 20% change (or + 6 animals). The precision is however much lower when both regions are analyzed independently, with effect size closer to 30% SD. We hence cannot guarantee that our design will allow separating the effect by regions.<sup>2</sup>

Our statistical precision lies within the typical impacts found for multifaceted programs (Banerjee et al., 2015): For instance, with a MDE of 15.2 % SD we are easily capable of detecting the asset (25.8 % SD), income and revenues (38.3 % SD) average effects detected in the six multifaceted studies analyzed by Banerjee et al. (2015). We are also well powered to detect effect on financial inclusion (36.7 % SD). We are however less powered to detect consumption per capita which only reaches 12.2 % SD or the food security index (10.7 % SD). We are not powered to detect Mental health index, political involvement indexes or women’s empowerment indexes, although these are not outcomes of interest in our study. Note however that the impact size presented in (Banerjee et al., 2015) are indexes, aggregate measures: aggregate measures are generally much more precisely estimated than individual parameters. As a result our MDE is a high bound of our expected detection power when we will create our own indexes.

Other important outcomes of interest like anemia, anthropometrics measure were not covered neither in (Banerjee et al., 2015) nor in (Bandiera et al., 2017). To compare our MDE with available results on early childcare programs, we look at the cash transfer literature. In a very recent re-analysis of the available literature on cash transfer (Bouguen et al., 2018), we identified the studies that report impacts on anthropometrics measures, for instance Height-For-Age (HAZ). Our MDE of 15.2% would allow to detect many of the reported impacts, although not all of them. For instance, Akresh et al. (2016) reports a 23% effect of an Unconditional cash intervention on HAZ and arm circumference while Kandpal et al. (2016) reports an effect 28.6% SD. There are also some smaller impacts like in (Macours et al., 2012) who find a CCT effect of only 7.2 % SD or more recently (McIntosh and Zeitlin, 2018) who finds an effect of 9.3% on HAZ. While we cannot guarantee that all positive impacts will be significantly detected, we are confident that we will have enough power for reasonable small impact sizes.

---

<sup>2</sup>Since ACF and TDH each cover one region, it would in any case be hard to distinguish the regional effect from the NGO effect. As a result, separating the effect by region has never been an objective

Table 4: **Height-for-Age Impacts in previous research**

Study	Project	Country	Impact HAZ (SD)	HAZ baseline (SD)	Age children	Treat.
McIntosh and Zeitlin (2018)	Benchmarking	Rwanda	9.3	-1.92	0-5 y	UCT
Macours et al. (2012)	AAC	Nicaragua	7.1	-1.08	3-4 y	CCT
Kandpal et al. (2016)	Pantawid	Indonesia	28.4	-1.93	6-36 m	CCT
Evans et al. (2014)	Tasaf	Tanzania	0	-1.45	0-4y	CCT
Akresh et al. (2016)	NCTPP	Burkina	33.4	-1.469	0-5 y	CCT
Paxson and Schady (2010)	BDH	Ecuador	0	-1.2	3-7 y	UCT
Galiani and McEwan (2013)	PRAF II	Honduras	+	?	6	CCT

## 4 Instruments and Surveys

Data collection instruments are primarily the questionnaires for the different stages of the study: a census survey, a targeting survey, baseline survey, a midline survey and an endline survey. Each stage or questionnaire includes several modules.

### 4.1 Household Economic Assessment (HEA)

As mentioned, the Household Economic Assessment (HEA) is a participatory targeting approach used to identify the households eligible in the villages. The HEA is conducted in three phases: (i) a census during which the survey team collects some basic information on all households of the villages (ii) a community based classification (CSE) in which the village classifies the households into four socioeconomic categories (iii) a quantitative analysis using both the CSE and the census to determine the list of eligible households.

Importantly, the HEA was not conducted by the same survey teams: TDH’s survey team conducted the HEA in the treatment villages in their region (Boucle du Mouhoun), ACF conducted the HEA in the treatment villages in their region (East) while the research team conducted the HEA in the control villages. The reason for this discrepancy comes from the financial structure of the project. Our donor included the HEA as part of the treatment and was included in the implementation partners budget, not in the research budget. Despite our effort, we could not figure out a legal way to assigned the HEA budget to the research budget.

Well aware of the problem, the research team took the lead on how the HEA was conducted everywhere. The research team first agreed on a set of survey tools that will be used for the HEA census and CSE. Second, we agreed on the data input technology used: we impose the implementation partners to use survey CTO which was our standard input method. The research team also set up and managed the server on which data were stored. Third, the research team conducted the training of field staff in a similar fashion in the three regions (control group, ACF treatment group and TDH treatment group). This way, the research team had total control

over the data, the way it was collected and the quality of the collection. The only real difference was the recruitment process of the field staffs.

To verify how such approach impacted the quality of the targeting, we will simply use the baseline - that will be collected by the same survey team (see baseline section 4.2)- and check if we can observe any differences in the composition of the eligible and non eligible in the treatment and control group: if the HEA was conducted differently in the treatment groups and in the control group we should see significant differences at baseline.

#### **4.1.1 Census**

The household census is administered to every single households in the village. It includes seven sections:

1. household identification,
2. characteristics of household head,
3. Access to safe water and sanitation and the nature of house
4. Livestock and land access,
5. household productive and non-productive assets,
6. access to basic social services
7. targeting group, ex. number of children under 5 years, pregnant women.

#### **4.1.2 Community Based Classification (CSE)**

The village survey is administered during a village committee where all the villagers are asked to participate. During the committee, villagers collectively decide:

- the Relevant poverty criteria to determine ultra-poor (UP), poor (P) average (A) and wealthier households (w).
- for each group (UP, P, A, W) and each criteria, the number of items/asset/amount each categories are supposed to have,
- Weights associated to each relevant criteria, which defines the importance of the criteria.
- Percentage of poor, ultra-poor, average and wealthier household in the village.

Then based on the criteria, the villagers elect three committees: two targeting committee and a reporting committee. Committee members are elected on the basis of their local knowledge of villagers and their standing in the community. If a

household does not agree with his classification, he could report directly to reporting committee which has the final decision. The two targeting committees consult the household census and use the data from the census to assign independently each household in one of the 4 categories. In case of discrepancy between the two committees, a meeting is organized to determine the final category of the household.

### **4.1.3 Using census and CSE to determine eligibility**

The research team finally uses the results from the CSE and the census to establish a list of eligible households. The criteria to be eligible were the following:

- The CSE assigned the household to the P or UP groups
- the census indicates that the households is composed of a pregnant women, a breastfeeding women or a child below five (NV)
- Since there were a limited number of eligible per village (max. 21), among the NV households categorized UP or P, census data were used to select the 21 poorest (see sampling)

As a result, being classified as UP or P does not guarantee eligibility: Since only 21 households per village could be selected, in some large villages, some Poor and Ultra poor households have been excluded (see sampling).

## **4.2 Baseline**

The multi-topic baseline study focuses on food security and nutrition using recently validated survey modules from the “Enquête Démographique et de Santé” (EDSN) and MICS survey instruments. This survey contains broad modules on nutrition and food security (to obtain child-specific food consumption scores), parenting practices, as well as anthropometric measures. Specifically, three indicators will be used to measure the nutritional status of children: i) height-to-age ratio to measure stunting, ii) weight-to-age ratio for measuring wasting, and iii) report height / weight to measure underweight. For these measures, an oral consent will be requested from parents / caregivers. The table below presents all modules of the baseline survey instrument:

<b>Module</b>	<b>Brief description</b>
Introduction and tracking	General information, informed consent,
Other programs participation	Identify if participants have benefited from other programs/interventions in the last 6 months.
Household roster	Collect information on household composition and demographics. Identifies number of children of age 0 to 5 years old and caregiver anthropometrics, test score and anemia measure
Agricultural revenue	Revenues from agricultural activities: crop production, land rents and expenses (hired labor, fertilizer costs, seed costs...).
Revenues from Income Generating Activities (not agriculture)	Collect information on revenues generating by small activities that are not small enterprises.
Social Network	Collect info on social relations that each member entertain with 10 other households in the village.
Finance	Formal borrowing, informal borrowing; total amount deposited into savings, total savings balance.
Transfers	In-cash in kind transfers to or from other households
Health & Education	Health and education expenditures in the last year
Expenses	Household's Expenses in the last 30 days, 6 months and 12 months.
Consumption	Total per capita consumption (value and quantity), Dietary diversity, Food consumption.
Food security	Does everyone gets enough food every day? Do adults skips meal? ...
Shocks	Incidence and response to (agricultural) shocks
Criminality	Have you been robbed, victim of physical violence in the past 12 months?
Aspirations	Aspirations, self-reported measures of own perspective wrt other people in the village
Child's Anthropometric & health, 0-5 years old)	Vaccines, Weight, Height and MUAC, hemoglobin rate per ml of blood for below 5 years old children, Sickness, visits to the nearest health centers, exclusive breastfeeding.
Child's development, 0-5 years old	Caregiver Reported Early Childhood Development modules (CREDI) caregivers declared

### 4.3 Follow and Endline surveys

The follow-up and the endline surveys will be drafted based on the baseline survey instrument. Specifically, the follow-up surveys will include sections on the implementation of the program to verify compliance with the treatment and provide feedback to partners on their performance on the field with respect to the research protocols.



## 4.4 Data Collection

The entire data collection process is conducted between August 2017 and January 2020. This period might slightly change due to unexpected events in the field, as activities could last longer than planned. Data from the survey is collected using tablets. All tablets are password protected to enhance data security, even in case of loss. All surveyors working in the process signed a non-disclosing of information document. Only the research team get access to data and data are de-identified before use. The confidentiality is strengthened by the consent form whereby it is clearly stated that information will be kept respecting confidentiality rule. Data will be collected using SurveyCTO, which includes a number of data-security features such as transport encryption (all internet communications are always encrypted with a Secure Sockets Layer (SSL) certificate, securing the data from a data-collection device to the server, or from the server down to the researcher) and survey data encryption. Sensitive data stored on the server will be encrypted and accessible only to the research team.

## 4.5 Data Processing

The project includes a targeting study, a baseline, a midline and an endline survey. The data will be analyzed at the end of each collection, normally within 3 months, but this may vary depending on the amount of information to be analyzed. Data processing involves data cleaning (making sure the dataset is complete, deleting duplicates, identifying and correcting extreme values, relabeling certain variables), and producing descriptive statistics and charts using Stata. First, the data with Personally Identifying Information (PII) will only be shared with the people mentioned in the IRB, including PIs. In addition, the PII will be replaced by missing values before sharing the data. The processed data will be made available to the researchers for empirical analysis. PIs will have access to encrypted data using Boxcryptor . A version of unencrypted data without PII could be shared with PIs who do not have a Boxcryptor account.

## 4.6 Variations from the intended sample size

We will disaggregate sources of attrition related to migration, complete refusal or partial refusal of survey questions. Based on IPA 5 years' experience conducting field surveys in rural areas, we anticipate a small attrition from the sample ranging between 3 and 10 per cent. Considering the study area proximity with national borders, this rate could go up to 10% due to migration. We can expect higher rates at the beginning of the rainy season because households are busy in agricultural activities.

We can limit attrition by improving data collection (survey design, administration and tracking). Our survey instruments and our field protocols will be pre-tested during their development phase and piloted with enumerators several times, before and after training. We will insure that during data collection, we will collect information about participants' phone number and about their neighbors. Enumerators are also asked to write directions to the household's compound, so we can better track them at midline and endline. We also ask if the participant can provide us with the contact of a point person who would know if they moved and where they moved if they did so. At baseline, we will also collect information at the village level with the chief of the village or the CVD, as it has been proven that it often works better to use several sources of information rather than asking only one person.

Small incentives can have large impacts on behavior to reduce attrition. Enumerators will distribute small packs of powder laundry or soap bars (for a total value less than \$1) to households participating to the study. It compensates households and especially respondents for the free time they are willing to give up in answering our questions.

## 5 Empirical Analysis

We here describe the empirical specifications, the main outcomes of interest and the hypothesis we will test. We begin by investigating program's compliance and the Intent to treat. Then we list the heterogeneity analyzes we will implement, the way we will deal with multiple hypothesis testing and finally describe the three specifications that we will use.

### 5.1 Program's compliance

We worked with the implementing partners to formalize the initial targeting approach as well as a monitoring system for transfers. As transfers are regular and relatively frequent, information on noncompliance or more precisely, non-receipt of transfers, will be collected and validated by comparing administrative data provided by partners and IPA collected survey data.

The program design is organized to transfer cash, assets, or food to very poor households. For this reason, we do not believe that poor households will refuse these transfers and not participate in the program. Yet, program (incomplete) implementation poses two potential noncompliance issues:

**Within Village non-compliance:** The partners do not perfectly respect the HEA ranking and distribute some interventions to non-eligibles or are unable to identify some eligibles. Although we provided the partners with very precise list of eligibles we cannot entirely rule out some non-compliance within the same villages.

**Between Village non-compliance:** The partners may not be able to treat all treatment villages. This risk is not null especially given the actual context in Burkina Faso. Numerous terrorist attacks were reported in eastern part of the country. Flooding is also a cause of concerns especially regarding cash distribution which is usually distrusted during the lean period (which corresponds to the rainy) season.

Both type of non-compliance (within and between) have different consequences. Village level non-compliance can be corrected using an instrumental variable. Yet, since we expect spill-overs, within village non-compliance usually cannot be easily corrected using an instrumental variables as it would violate the SUTVA assumption. We start by a village level first stage of the form:

$$D_v^1 = \beta_0 + \beta_1 F_v + \beta_2 F A_v + \beta_3 F A N_v + \epsilon_v \quad (1)$$

Where  $F_v$ ,  $F A_v$  and  $F A N_v$  are village level treatment variable and  $D_v^1$  is composed of village level compliance variable i.e. the village received at least partially some intervention .  $D_v^1$  will be created using data given by our partners and our own data. Discrepancy between partner (administrative data) and own report (survey data) will be analyzed carefully to determine which one is true. Therefore  $D_v^1$  will include the following six measures (a village is deemed treated if the partners treated at least one household in the village) :

- The village has been treated with the finance program according to the partners and according to our survey
- The village has been treated with the Asset program according to the partners and according to our survey
- The village has been treated with the Nutrition program according to the partners and according to our survey

Our assumption **(H1)** is that  $\beta_1, \beta_2, \beta_3$  are all equal to 1 i.e perfect compliance. If H1 is violated, impacts in sub-section 2 will be analyzed using a LATE estimator where (1) (plus potentially some control variables) is the first stage.

We then, move to individual level first stage of the form:

$$D_{iv}^2 = \alpha_0 + \alpha_1 F_v + \alpha_2 F A_v + \alpha_3 F A N_v + \alpha_4 F_v * B_{iv} + \alpha_5 F A_v * B_{iv} + \alpha_6 F A N_v * B_{iv} + \alpha_7 B_{iv} + \mu_v + \epsilon_{iv} \quad (2)$$

Where  $B_{iv}$  is the variable indicating if the household i was identified as eligible during the HEA.  $D_{iv}^2$  is a set of variable indicating whether each household received each of the intervention and will include six measures:

- The household was treated with the finance program according to the partners;

- The household was treated with the finance program according to our follow-up data;
- The household was treated with the Asset program according to the partners;
- The household was treated with the Asset program according to our follow-up data;
- The household was treated with the Nutrition program according to the partners;
- The household was treated with the Nutrition program according to our follow-up data

Discrepancy may exist between partners and survey as partners may improperly register attendance or household improperly remember program's participation. Since there will be no easy way to know what is the truth, we will present both measures. If (1) show perfect village compliance, (2) measures the take-up rate. We expect:

**(H2)**  $\alpha_1 = \alpha_2 = \alpha_3 = 0$  non-eligibles in the treatment group receive no interventions;

**(H3)**  $\alpha_7 = 0$ ; eligibles in the control group receive no program;

**(H4)**  $\alpha_4 = \alpha_5 = \alpha_6 = 1$  all eligibles in the treatment groups receive the program.

H2 corresponds to zero intervention among the non-eligibles, (H3) to zero interventions in the control village and (H4) full compliance among the eligibles in the treatment group. As mentioned, if any of these hypothesis is violated, LATE might be bias as the SUTVA assumption may probably be violated. We would therefore provide ITT results and first stage, analysis LATE only in the situation where we can be certain that no spill-overs happened.

## 5.2 Intent-to-treat

The overall intention to treat empirical will be estimated using a comparison between all eligibles and non-eligibles in the treatment and control groups:

$$y_{iv}^1 = \beta_0 + \beta_1 F + \beta_2 FA + \beta_3 FAN_v + \mu_v + \epsilon_{iv} \quad (3)$$

where  $y_{iv}^1$  is the outcome follow-up (midline or endline). The intent-to-treat (ITT) overall effect of cash transfers (F), cash and assets (FA), and cash, assets and nutrition interventions ( $FAN_v$ ) is given by  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  respectively. The individual components of the multidimensional program are given by  $\beta_1$  for the cash component,  $\beta_2 - \beta_1$  for the asset transfer (plus interaction between F and A), and  $\beta_3 - \beta_2 - \beta_1$  for the nutrition program component (plus interactions between F A and N).  $y_{iv}^1$  will include at least the following outcomes:

- Income (10 outcomes)
  - Income generated by external activities (wages, employment outside of family business or family farm). Number of days and hours per week spend in these activities. (4 outcomes)
  - Income generated by family business activity, time and days spend in these activities, capital invested in these activities (4 outcomes)
  - Income generated by agricultural farming (1 outcome)
  - Income generated by livestock farming (1 outcome)
- Household spending and consumption (7 Outcomes):
  - health and education spending (2 outcomes)
  - Household food consumption (1 outcomes)
  - non-durable good (<30 days) (1 outcome) average durable goods (<6 months) (1 outcome)
  - durable goods (<12 months) (1 outcome)
  - Household event spending (1 outcome)
- Household food security index and diversity consumption index (2 outcomes)
- Assets: (3 outcomes)
  - Animals
  - Farming equipment
  - House equipment
- Saving/finance (3 outcomes)
  - Belong to a “tontine”, amount invested in tontine
  - Amount save elsewhere
  - Credits
- Amount of transfer/remittances (2 outcomes)
  - Number of shocks (negative 1 outcomes)
  - Number of thefts (negative 1 outcome)
- Aspiration questions: number of hectares, number of cattle, number of years: own position – (max position -min position) (3 outcomes)
- Network structure (4 outcomes):
  - Number and intensity of connections with other eligibles in village

- Number and intensity of the connection with other non-eligibles in village.
- For kids (0-36 months) (six outcomes):
  - vaccinations,
  - anthropometrics (WAZ, HAZ, WHZ),
  - anemia (Hg/L)
  - CREDI score.

A total of 36 outcomes at the household level + 6 outcomes at the child level will be tested in priority. A hypothesis is validated if this hypothesis is valid for a majority of outcomes. The following hypotheses are hence deemed valid if the majority of the 42 outcomes are positive (at least at 10%) and significant (see multi-hypothesis testing below). From (1) we expect the following impacts:

**(H5):**  $\beta_1, \beta_2, \beta_3 > 0$  overall program's effect is positive (F, FA, FAN)

**(H6):**  $\beta_1 < \beta_2 < \beta_3$  the program's impact monotonously increases with resources/interventions.

Corollary to H5 and H6, we can test the following  $\beta_1, \beta_2 - \beta_1, \beta_3 - \beta_2 - \beta_1 > 0$  i.e. each program component (F, A, N) has a positive contribution.

**(H7):**  $\beta_1 > \beta_2 - \beta_1 > \beta_3 - \beta_2 - \beta_1 > 0$  will test whether cash is the relatively largest program component of the poverty reduction. Asset the second largest and nutrition is marginal.

Then, to estimate the main intent to treat effects on the eligibles, we will use the following:

$$Y_{iv}^1 = \alpha_0 + \alpha_1 F + \alpha_2 FA + \alpha_3 FAN_v + \mu_v + \epsilon_{iv} \text{ if } E = 1 \quad (4)$$

where E is an indicator of the eligible, determined before the beginning of the intervention during HEA (see infra). We compare eligibles in the treatment and control villages to each other.  $\alpha_1, \alpha_2, \alpha_3$  measure the impact of the program and as before the individual components of the multidimensional program are given by  $\alpha_1$  for the cash component,  $\alpha_2 - \alpha_1$  for the asset transfer, and  $\alpha_3 - \alpha_2 - \alpha_1$  for the nutrition program component. We will test the following hypothesis:

**(H8)**  $\alpha_1, \alpha_2, \alpha_3 > 0$  Positive effect of the program on the eligibles;

**(H9)**  $\alpha_1 < \alpha_2 < \alpha_3$  the program's impact monotonously increases with resources/interventions on the eligibles.

To estimate spillover effects, we restrict the sample to the non-eligibles in the treatment and control villages such that:

$$Y_{iv}^1 = \gamma_0 + \gamma_1 F + \gamma_2 FA + \gamma_3 FAN_v + \mu_v + \epsilon_i v \text{ if } E = 0 \quad (5)$$

(H10)  $\gamma_1, \gamma_2, \gamma_3 > 0$  Positive effect of the program on the non-eligibles

(H11)  $\gamma_1 < \gamma_2 < \gamma_3$  effect of the multifaceted programs is positive

### 5.3 Heterogeneous Effects

In addition to eligibles versus non-eligibles, we will use a wealth score to estimate heterogeneous wealth effects. The relative distribution of wealth may vary within or between villages. Note that we will not rely on the HEA cut off because the HEA was conducted by different organizations (IPA, TDH and ACF). We prefer to rely on the baseline conducted homogeneously by IPA in T and C branches. The wealth score will be based on the following baseline aggregate indices:

- Education index based on the average number of people who can read and write, the average highest level of education and the average enrollment of below 16 year old
- Housing index
- Equipment index
- Animal index
- Overall income
- Total spending
- Total food consumption
- Shock index
- Baseline food security index

These are the same variables used as control variable in specification (3). This wealth score will be used to create eligible quartile (Nq1, Nq2, Nq3, Nq4) and non-eligible quartile (NBq1, NBq2, NBq3, NBq4) We will then estimate the following model:

$$\begin{aligned} Y_{iv}^1 = & \alpha_0 + \alpha_1 F + \alpha_2 FA + \alpha_3 FAN_v \\ & + \alpha_4 F * Bq1 + \alpha_5 FA * Bq1 + \alpha_6 FAN_v * Bq1 \\ & + \alpha_7 Bq1 + \mu_v + \epsilon_{iv} \text{ if } E = 1 \text{ \& } Bq1 = 1 \text{ or } Bq4 = 1 \end{aligned} \quad (6)$$

(6) compares the treatment effects in the quartile 4 and in the quartile 1 of the eligibles. We will test the assumption that : (H13)  $\alpha_4 > 0, \alpha_5 > 0, \alpha_6 > 0$  i.e. the poorest benefit more from the program. (H14)  $\alpha_4 < \alpha_5 < \alpha_6$  i.e. the more multifaceted the program is the larger the impact on the poorest. Similarly for the

non-eligibles:

$$\begin{aligned}
Y_{iv}^1 = & \alpha_0 + \alpha_1 F + \alpha_2 FA + \alpha_3 FAN_v + \\
& + \alpha_4 F * NBq1 + \alpha_5 FA * NBq1 + \alpha_6 FAN_v * NBq1 \\
& + \alpha_7 NBq1 + \mu_v + \epsilon_{iv} \text{ if } E = 1 \ \& \ NBq1 = 1 \text{ or } NBq4 = 1
\end{aligned} \tag{7}$$

**(H15)**  $\alpha_4 > 0, \alpha_5 > 0, \alpha_6 > 0$  i.e. the poorest non-eligibles benefit more from the program.

**(H16)**  $\alpha_4 < \alpha_5 < \alpha_6$  i.e. the more multifaceted the program is the larger the impact on the poorest non eligibles.

Other heterogeneous variable will rely on the social proximity index. Each non-eligibles will be given a score based on the number of eligibles he knows and the average social proximity with the eligibles. Then the non-eligible will be ranked and place in quantile groups SPq1, SPq2, SPq3, SPq4. Then we can measure the spill over using:

$$\begin{aligned}
Y_{iv}^1 = & \alpha_0 + \alpha_1 F + \alpha_2 FA + \alpha_3 FAN_v + \\
& \alpha_4 F * NBq1 + \alpha_5 FA * SPq1 + \alpha_6 FAN_v * SPq1 \\
& + \alpha_7 SPq1 + \mu_v + \epsilon_{iv} \text{ if } NE = 1 \ \& \ SPq1 = 1 \text{ or } SPq4 = 1
\end{aligned} \tag{8}$$

where NE are the non eligibles.

**(H17)**  $\alpha_4 < 0, \alpha_5 < 0, \alpha_6 < 0$  i.e. the social proximity index determine the size of the spill over.

If (6), (7), (8) produces imprecise results, we can also test the same strategy but separating groups at the median and testing the same hypothesis (the poorest benefit more than the richest). This has the advantage of using all the sample size (instead of half) but has the disadvantage to use a less precise measure of poverty/social proximity.

To summarize, we propose to estimate the following regression:

Hypothesis	Specifications	Outcomes	# of regression
H1-H4	1	6	24
H5-H17	3	36 (household)	1326 (household level)
		42 (children level)	1638 (children level)
Total			1350

A total of 1350 regressions will be conducted for the household level regression, 1638 if we take into account children level variables.



## 5.4 Multiple outcome and multiple hypothesis testing

The first strategy used will be to create an index for the 11 aggregated outcome categories listed above (income, spending, asset, food and security, saving, transfer, shocks, network, aspiration, criminality, children performance) using the Kling Liebmann strategy (Kling et al., 2007) and then use the Anderson (2008) (family wise correction) to correct for 11 outcome hypothesis testing. Second, inside each of the 11 categories, we employ false discovery rate to study each specific outcomes within each categories. Finally we can test the 36 (or 42 outcomes) using the false discovery rate.

## 5.5 Specification

For equations 4,5,6,7,8 , we will estimate three specifications:

- One including treatment variables as well as an dummy indicating the region (TDH or ACF regions);
- One including in addition the variables S used during randomization. S includes the village level variables used during the randomization (T versus C, before HEA):
  - a commune fixed effect (the commune variable used at randomization),
  - village GPS coordinates (2 continuous variables),
  - distance to the closest health center,
  - total number of households in village,
  - share of female in village,
  - whether or not TDH/ACF ever intervene in this village,
  - whether the village has low-lying ground (“bas fond”),
  - whether the village is amenable to soil restoration.
- The last one will include in addition to 2, the following baseline variables  $y_{iv}^0$ :
  - The  $y_{iv}^1$  equivalent at baseline, if collected.
  - All  $y_{iv}^0$  that were significantly unbalanced at baseline
  - If  $y_{iv}^1$  is at the household level, we further control for:
    - \* Overall measure of household Income
    - \* Household size;
    - \* Overall measure of social proximity index ;
    - \* Total baseline saving minus credit;
    - \* Total baseline household general spending based on health, education, durable and non durable good and event spending

- \* Total baseline food consumption
- \* Baseline Food security index ;
- \* Index of baseline shocks
- \* Housing/equipment/animal index
- \* Education index: aggregate measure that includes a read/write score, ratio of school <16 year old, average highest education >16 year old
- If  $y_{iv}^1$  is at the individual level (not a child 0-36months), in addition to household level characteristics:
  - \* Age
  - \* Gender;
  - \* Highest education;
  - \* Read/write.
- If  $y_{iv}^1$  is a child aged between 0-36months, in addition to household level characteristics, we will include: Baseline CERDI test score ;
  - \* trimester and year of birth (fixed effect);
  - \* gender;
  - \* baseline weight;
  - \* baseline height.
  - \* Anemia level.

(1) (2) and (3) will be estimated and interpreted jointly. According to Bruhn and McKenzie (2009), SE in (1) are biased downward and therefore (2) and (3) will be given priority in the interpretation of our results. Standard error will be estimated using a cluster at the village level and p value will be adjusted for multiple-hypothesis testing (see supra).

## 6 Missing values, attrition and outliers

We always impute baseline control variable missing value, never the follow-up outcome variables. If a baseline variable use in the specification has a missing value, we set its value to 0 and control for a dummy variable taking the value 1 when the variable is missing and 0 otherwise. Attrition, not significantly different between experimental branches, is not directly address. Differential attrition is addressed using Lee Bounds. Each income, saving, and spending variables can be corrected for outliers. Each value above p99 is set at p99.

## 7 Declaration of interest

The PIs have no conflict of interest with both partners ACF and TDH. PIs are not remunerated by the partners, directly or indirectly.

The European Union funds this research is funded as well as the interventions (ACF and TDH).

## **8 Acknowledgments**

All IPA Burkina staff that have worked on the project including Aliou Diallo, Estelle Plat, Pablo Cordoba-Bullens.

## References

- Akresh, R., De Walque, D., and Kazianga, H. (2016). Evidence from a randomized evaluation of the household welfare impacts of conditional and unconditional cash transfers given to mothers or fathers. Technical report, The World Bank.
- Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.
- Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., and Sulaiman, M. (2017). Labor markets and poverty in village economies. *The Quarterly Journal of Economics*, 132(2):811–870.
- Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B., and Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236):1260799.
- Blattman, C., Fiala, N., and Martinez, S. (2014). Generating skilled self-employment in developing countries: Experimental evidence from uganda \*. *The Quarterly Journal of Economics*, 129(2):697–752.
- Blattman, C., Fiala, N., and Martinez, S. (2018). The long term impacts of grants on poverty: 9-year evidence from uganda’s youth opportunities program. *Working Paper*.
- Bouguen, A., Huang, Y., Kremer, M., and Miguel, E. (2018). Using rcts to estimate long-run impacts in development economics. Working Paper 25356, National Bureau of Economic Research.
- Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American economic journal: applied economics*, 1(4):200–232.
- Evans, D., Hausladen, S., Kosec, K., and Reese, N. (2014). *Community-based conditional cash transfers in Tanzania: results from a randomized trial*. The World Bank.
- Filmer, D., Friedman, J., Kandpal, E., and Onishi, J. (2018). General equilibrium effects of targeted cash transfers: nutrition impacts on non-beneficiary children. Technical report, The World Bank.
- Galiani, S. and McEwan, P. J. (2013). The heterogeneous impact of conditional cash transfers. *Journal of Public Economics*, 103:85–96.
- Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimentalevidence from kenya. *The Quarterly Journal of Economics*, 131(4):1973–2042.
- Kandpal, E., Filmer, D., Friedman, J., Onishi, J., Alderman, H., and Avalos, J. (2016). A Conditional Cash Transfer Program in the Philippines Reduces Severe Stunting. *The Journal of Nutrition*, 146(9):1793–1800.

- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.
- Macours, K., Schady, N., and Vakis, R. (2012). Cash transfers, behavioral changes, and cognitive development in early childhood: Evidence from a randomized experiment. *American Economic Journal: Applied Economics*, 4(2):247–73.
- McIntosh, C. and Zeitlin, A. (2018). Benchmarking a child nutrition program against cash: Experimental evidence from rwanda. Technical report.
- Paxson, C. and Schady, N. (2010). Does Money Matter? The Effects of Cash Transfers on Child Development in Rural Ecuador. *Economic Development and Cultural Change*, 59(1):187–229.