Centre of Excellence for Development Impact and Learning

March 2022

Design Paper 6

An Empirically-driven Theory of Poverty Reduction

Sudhanshu Handa

London International Development Centre

UK aid from the British people

Oxford Policy Management

About CEDIL

The Centre of Excellence for Development Impact and Learning (CEDIL) is an academic consortium supported by the UK government through UK Aid. The mission of the Centre is to test innovative methodologies in evaluation and evidence synthesis and promote evidence-informed development. CEDIL-supported projects fall into three programmes of work: evaluating complex interventions, enhancing evidence transferability, and increasing evidence use.

CEDIL design paper

CEDIL design papers present the protocols of studies supported by the CEDIL programme. They describe the objectives, the methodology and the implementation plan of impact evaluations, evidence syntheses, and methodological papers supported by CEDIL.

About this design paper

This design paper was submitted to CEDIL by the "An empirically driven theory of poverty reduction" Project L.191 team.

Please direct any comments or queries to the corresponding author, Sudhanshu Handa at Carolina Population Center, University of North Carolina at Chapel Hill.

Suggested citation: Handa, S. (2022) 'An Empirically-driven Theory of Poverty Reduction', CEDIL Design Paper 6. Centre of Excellence for Development Impact and Learning (CEDIL), London and Oxford.

This project was funded by the Centre of Excellence for Development Impact and Learning (CEDIL), supported by UK aid from the UK Government. The views expressed in this research project paper do not necessarily reflect the UK Government's official policies or CEDIL.

All content is the sole responsibility of the authors. CEDIL has not commissioned a professional copy-edit of this paper. Any errors and omissions are the sole responsibility of the authors.

Project title: An Empirically-driven Theory of Poverty Reduction

Lead organisation: Carolina Population Center, University of North Carolina at Chapel Hill

Principal investigator: Sudhanshu Handa

Project number: L.191

Date of draft: October 23, 2020, Revised February 4, 2021

REARCH DESIGN PAPER

Overview

We use secondary data from four countries to develop a 'middle-range' theory of poverty reduction. The data come from impact evaluations of four national unconditional cash transfer programs and are merged with a secondary variables on the microenvironment. We apply machine learning algorithms (MLA) to identify the variables that predict large gains in well-being among cash transfer recipients. By looking across four different settings, we can use the empirical results to construct a mid-range theory of graduation from ultra-poverty. The key innovation in this study is the use of a new MLA that is explicitly built to estimate heterogeneous treatment effects.

Policy relevance

The persistence of poverty in sub-Saharan Africa means that policy-makers in this region continue to debate the right mix of supply- versus demand-side interventions that can move large groups of households out of extreme or ultra-poverty. For several decades microfinance has been viewed as the holy grail of poverty reduction, the implicit assumption being that that the poor lack access to credit, which is the key barrier to underinvestment in productive activity, thus limiting their income growth. However, recent rigorous evidence on the microfinance model confirms that this approach has over-promised and under-performed in moving large numbers out of poverty.[1][2] A more recent innovation comes from Bangladesh through the NGO BRAC, who provide a big-push intervention designed to 'graduate' households from poverty. Proponents of these 'graduation programs' claim that a big push at the household level, comprising a cash transfer, asset transfer, skills training and coaching can graduate households permanently into the middle-class. Evidence from a six country impact evaluation suggests promising evidence of these graduation programs[3], but these interventions are expensive and complex, and have yet to be actually implemented by a national government, which is telling.

Others have argued that rather than context and skills, decision-making under scarcity and the psychological toll of poverty are key barriers to economic growth[4]. Specifically, the condition of poverty itself taxes mental bandwidth, reducing decision-making capacity and leading to sub-optimal outcomes that perpetuate poverty. Meanwhile, at the other end of the spectrum are proponents of unconditional cash transfers to the ultra-poor. A recent article in Foreign Affairs by Chris Blattman and Paul Niehaus argue that unconditional cash should be the new benchmark in foreign aid, and that very few interventions can beat the cost-effectiveness of providing the ultra-poor with plain cash, which allows them to spend money in the way that best allows them to satisfy their priorities[5]. Two recent studies provide head-to-head comparison of impacts of unconditional cash versus a bundled health and nutrition program and a youth training program respectively. The bundled health and nutrition program did better at improving savings (an explicit program objective), while the cost-equivalent cash transfer was better at reducing debt and improving a wider range of outcomes relative to the bundled sector specific program, though neither program was able to improve child health and nutrition outcomes[6]. In the second study, a cost-equivalent cash transfer delivered larger impacts on productive assets, savings and productive hours worked relative to the skills development intervention that provided three 10-week training sessions[7].

---

[1] Abhijit Banerjee, Dean Karlan and Jonathan Zinman (2015). "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Journal: Applied Economics* 7(1): 1–21.

[2] Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. "The Miracle of Microfinance? Evidence from a Randomized Evaluation", *American Economic Journal: Applied Economics*, 2015, 7(1): 22-53.

[3] Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Pariente, J. Shapiro, B. Thuysbaert, and C. Udry (2015) "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." Science 348, no. 6236 (May 14, 2015): 1260799–1260799.

[4] Haushofer, J., & Fehr, E. (2014). "On the Psychology of Poverty." Science 6186: 862-867.

[5] https://www.foreignaffairs.com/articles/show-them-money

[6] Craig McIntosh & Andrew Zeitlin. Benchmarking a child nutrition program against cash: Experimental evidence from Rwanda. 2018. https://gps.ucsd.edu/_files/faculty/mcintosh/mcintosh_research_child_nutrition.pdf

[7] Craig McIntosh & Andrew Zeitlin. Using household grants to benchmark the cost effectiveness of a USAID workforce readiness program. 2020. https://arxiv.org/pdf/2009.01749.pdf

This project will use secondary data to test and develop an empirically driven 'middle-range' theory of poverty traps that can provide key programming guidance to governments on how to move large group of households out of poverty permanently. The current policy innovations in sub-Saharan Africa in the area of inclusive growth include initiatives (primarily by NGOs) to implement graduation-style programs or sector-specific livelihood interventions (livestock or introduction of new crops), large-scale unconditional cash transfer programs with primarily protective objectives, and a handful of behavioral interventions mostly done by independent researchers. There is published literature on the impacts of graduation programs (see above), unconditional cash transfers[8], and the potential effects of psychological states and internal constraints on poverty[9]. We bring all these ideas together into one framework in order to ultimately provide recommendations about the appropriate mix and context to deliver sustained reductions in poverty, or 'graduation'. The fact is that not all poor household are the same, and the required interventions and supports will vary by household capacity, environmental factors and their interaction. Our objective is to explicitly model these demand and supply side factors and their interactions across four countries to look for patterns and inform poverty reduction policy.

This work is not estimating the impacts of four national unconditional cash transfer programs on a series of outcomes, as mentioned above there is a nice existing literature on this already. And evaluations often look at heterogeneous effects along prespecified characteristics (such as gender of head, pre-program consumption, or schooling of the recipient). Our approach is to let the data identify the heterogeneity in the pattern of impacts, since not all possible interactions of characteristics can be prespecified. We seek patterns in the identification of characteristics by looking across four national programs. Once we identify groups of households that tend to be high-flyers, we then examine the behaviors they adopted with the cash, The specific policy recommendations of the work would be to provide governments with the types of activities that households engage in with the cash that lead to large gains in consumptions, and the profile of households that engage in these activities. Complementary programs that support these activities can then be directed to households that fit the profile of likely households to engage in those activities. This is in fact exactly what governments are trying to do now. Members of the research team have already participated in dialogues with the Governments of Ghana, Zambia and Malawi to assist them in trying to identify a suite of relevant complementary activities to be targeted to different types of cash transfer beneficiaries to improve overall well-being.

---

[8] Sudhanshu Handa, Luisa Natali, David Seidenfeld, Gelson Tembo and Benjamin Davis, "Can unconditional cash transfers raise long-term living standards? Evidence from Zambia, "Journal of Development Economics, Vol 133(July): 42-65, 2018.
[9] Mani Anandi, S. Mullainathan, E. Shafir, and J. Zhao. Poverty Impedes Cognitive Function. Science 341, no. 6149:976–980. 2013.

With respect to the stated categories of policy relevance provided by CEDIL, this work makes contributions in three categories: 1) methodological innovation by using newly a developed MLA designed explicitly for heterogeneous treatment effects; 2) addressing evidence gaps on what works for poverty reduction; 3) providing evidence that can lead directly to policy actions.

We will use impact evaluation data from four national unconditional cash transfer programs. The countries and programs are: 1) Malawi Social Cash Transfer Program SCTP); 2) Zimbabwe Harmonized Social Cash Transfer Program (HSCT); 3) Ghana Livelihood Empowerment Against Poverty (LEAP); 4) Zambia Child Grant Program (CGP). All evaluations include one baseline and multiple follow-ups (except for Ghana with just one follow-up). In Malawi and Zambia the design is an RCT; in Ghana it is a discontinuity design using the proxy means test cut-off; in Zimbabwe the design is matched Wards (administrative units below the district) followed by the application of household targeting by the program so all households in the comparison Wards are future eligible households. In all four cases, treatment can be considered exogenous, and extensive baseline balance tests confirm the fidelity of the original designs. Table 1 summarizes the data sets and survey years. All initial surveys are baseline. All impact evaluation reports are available at the Transfer Project website (https://transfer.cpc.unc.edu/). These reports provide impact estimates across all major productive, social and family domains, typically over 100 indicators.

Table 1: Sample sizes and survey years (first year is pre-treatment)

	Treatment	Control	Survey years
Ghana LEAP 1000	1,262	1,235	2015, 2017
Malawi SCTP	1,730	1,800	2013, 2014, 2015
Zambia CGP	1,260	1,252	2010, 2012, 2013, 2014, 2017
Zimbabwe HSCT	1,029	1,034	2013, 2014, 2017
Total	5,281	5,321	

The four programs have some core common elements that make them good case studies to use together, but also have some distinctive features that might affect the behavioral response to cash and the possibility of graduation. In terms of common elements, all four are unconditional cash transfer programs, with payments in cash provided bimonthly, and all households are rural and ultra-poor in the sense that their consumption falls below the extreme or ultra-poverty lines in the respective countries (an explicit targeting criterion). The notable differences have to do with the size and structure of the transfers and the demographic composition of the target population.

In Ghana, Malawi and Zimbabwe, transfer values vary with the number of members in the household while in Zambia the transfer is independent of household size. The relative value of the transfer differs considerably across programs, the transfer as a share of baseline consumption is 14 in Ghana, 18 in Malawi, 26 in Zambia and 21 percent in Zimbabwe. Experience from the Transfer Project across a dozen country programs indicates that a transfer value that is at least 20 percent of baseline consumption is likely to lead to larger, more transformative impacts, that is, impacts that move beyond food security and into broader consumption, savings and productive activity (Davis & Handa 2015)[10]. Based on this experience we would expect a larger possibility of graduation in Zambia and Zimbabwe relative to the other two countries. One key difference between the Ghana LEAP and the other programs is that LEAP beneficiaries are automatically eligible for a fee waiver for the Ghana National Health Insurance Scheme (NHIS). However, this linkage and enrollment is not automatic, so that beneficiaries still have to go to a health clinic with their LEAP card to enroll and obtain their NHIS card, which in practice poses a barrier for many households (Palermo et al 2018)[11].

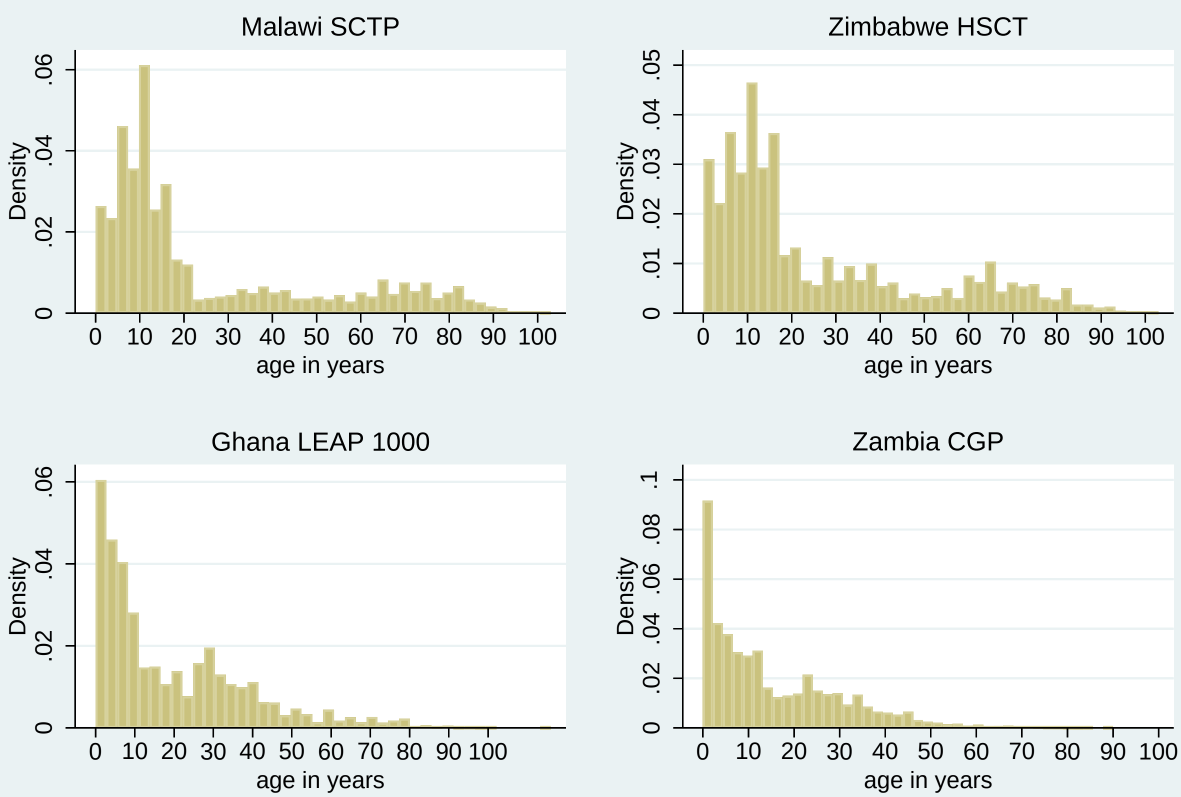
The other key difference among the four programs is the demographic eligibility criterion. The Malawi and Zimbabwe programs target labor-constrained households, which is operationalized as having a dependency ratio above 3, where the dependents are defined as those age 0-18 and 60+ years of age and anyone who is disabled. The Zambia CGP targets households with children under age 5 years, and the LEAP evaluation is of the 'LEAP 1000' window--the specific component of LEAP that targets households with a pregnant woman or a child under 1 year of age. Figure 1 highlights these demographic features by showing the age profile of beneficiaries from the baseline data of the four program evaluations. The Malawi SCTP and Zimbabwe HSCT use almost identical demographic eligibility rules to select beneficiaries. Targeted households tend to have very few members of prime working age, have elderly heads of household, and a large proportion of the adolescent children are orphaned grandchildren or nieces/nephews of the head of household. In the Ghana LEAP 1000 window and the Zambia CGP, by contrast, tend to have households that are younger, with many young children, and more able-bodied, prime-age members. These differences in demographic composition will undoubtedly influence the response to the transfer. Labor-constrained households--those with fewer prime-age members--will find it harder to use cash transfers for productive activity that require complementary labor input. They may instead use the cash to hire labor, opt for productive

[10] Davis, B., and S. Handa. 2015. "How Much do Programmes Pay? Transfer Size in Selected National Cash Transfer Programmes in Africa." The Transfer Project Research Brief 2015-09, Carolina Population Center, University of North Carolina, Chapel Hill.

[11] Palermo TM, Valli E, Ángeles-Tagliaferro G, LEAP 1000 Evaluation Team, et al. "Impact evaluation of a social protection programme paired with fee waivers on enrolment in Ghana's National Health Insurance Scheme. " *BMJ Open* 2019;9:e028726. doi: 10.1136/bmjopen-2018-028726

activity that is not labor-intensive, or choose options that can be easily done by available labor (i.e. children) such as livestock rearing[12].

Figure 1: Age structure of members in beneficiary households



Innovation

CEDIL's stated objective is to support research "*…to innovate and build capacity in the field of international development impact evaluation*." Our contribution to this objective is to apply new methodological approaches to a classical question, thus contributing directly to the 'innovate' component of the mission statement. Specifically, we adopt recent developments at the intersection of machine learning (ML) and causal inference[13] to build a 'middle-range' theory around poverty traps, which in turn can help governments understand the mix of interventions necessary under different conditions to enable households to raise their

---

[12] Jacob De Hoop, Valeria Groppo and Sudhanshu Handa, "Cash transfers, entrepreneurial activity, and child work: Evidence from Malawi and Zambia," World Bank Economic Review, Vol. 34(3): 670-697, 2020. https://academic.oup.com/wber/article/34/3/670/5611144

[13] Susan Athey, 2019, The Impact of Machine Learning on Economics, Ch.21 in The Economics of Artificial Intelligence, University of Chicago Press.

economic productivity. ML techniques are appropriate for our problem because we have a very large set of potential indicators that can influence the treatment effect of a cash transfer. ML allows the data to decide which characteristics lead to heterogeneous treatment effects, either large or small (or even negative). These characteristics can then be used to build a theory of graduation from poverty, indicating which households are more likely to benefit from a cash transfer versus other types of interventions, and what the pathways are to sustained increases in consumption.

Further details of the methodological innovation are described in the Technical Design section below. The methodological innovation contributes to CEDIL's research agenda in the area of evidence transferability, because it will allow us to build a middle range theory that can be applied to other settings. As discussed in more detail below, we will partition our data into trial and test sets in order to validate the predictive ability of the results derived from the method. We will make our code for running the MLA available to other researchers who are tackling similar questions, which will support the diffusion of the method.

Technical design

The cash transfers are an external liquidity injection, which affect psychological states and household behavioral responses. These may interact with each other, and with the microenvironment to produce different effects among households with different characteristics. The typical approach to heterogeneous treatment effects is to specify, using theory, why an intervention may have a differential impact on one group over another, and then to test this hypothesis by comparing treatment effects across groups. Our approach is different. We first identify households for whom there is a large treatment effect, where consumption is the key outcome of interest. Our survey instruments (available on the Transfer Project website) include the full consumption module taken from the respective national living conditions surveys in each country, so we have an excellent, comprehensive consumption measure. We have a large set of pre-treatment variables, including geovariables (see below for a discussion of variables); MLAs allow us to identify from this very large set of pre-treatment variables the key subset that predict large gains in consumption.

A key challenge we face in this exercise is how to classify households into 'high achievers'--those who were able to convert the cash transfer into large gains in consumption--and moderate and low (or perhaps negative) achievers. A simple but naïve approach would be to look at the endline consumption of treatment households and pick those in the top 10 percent (say) of the distribution as high-achievers, and then describe their pre-treatment characteristics. This approach has at least three weaknesses. First, we do not know the counterfactual of those in the top 10^{th} percentile of consumption, perhaps they were always high achievers even before the program (even if we were to use consumption growth to identify high achievers, the issue of the missing counterfactual still exists). Second, we have a large potential pool of characteristics to use to describe these households—which are the most important or salient? Third, we do not how these characteristics may interact with each other to produce high achievers.

We resolve these issues using new methodological results from the intersection of the literatures in ML and causal inference. Athey & Imbens (2016)[14] lay out an approach that allows the data to define the relevant sub-groups for which heterogeneous treatment effects exist within a randomized control trial or other design where the treatment can be considered exogenous, as in our case. The approach uses classification regression trees (CART), a method that partitions or classifies the data based on the degree of homogeneity (or 'purity') of an outcome variable or target. For example, say the target is income and the characteristic is sex (male, female). The CART approach would partition the data into two groups (male, female) if the degree of homogeneity or purity of the outcome (income) in each of the sub-groups was greater than the degree of homogeneity in the overall sample, where purity is measured using the sum of the mean squared deviations for example. This is the intuition behind the CART, and the typical application is when there are a large number of characteristics or features in the data, the outcome or predictor is known, and we want to classify or partition the sample based on groups of characteristics that have the same value of the predictor. This is almost exactly our problem, except that our predictor is not the level of consumption but rather *the level of the treatment effect on consumption*. In other words, we need to solve the problem of the counterfactual by identifying households where the treatment effect is largest—this is the innovation introduced by Athey & Imbens.

Specifically, they develop an approach that allows the data to be partitioned using a CART based on the *treatment effect* rather than the level of the predictor, using an 'honest' estimation approach where half the sample is used to estimate the tree and perform cross-validation, and the other half used to compute the actual treatment effects for each terminal leaf.[15] We describe the sequence of steps below.

1. Data is divided into two (50% split), one half is used to estimate and prune the tree (trial data) and the other to estimate the actual treatment effects (test data). This is referred to 'honest' estimation.

[14] Susan Athey & Guido Imbens, 2016, Recursive Partitioning for Heterogeneous Causal Effects, PNAS, Vol.113 (27): 7353-7360.

[15] Athey & Imbens derive an estimate of the mean square deviation for the counterfactual outcome based on the characteristics of the household (those used to partition the data) and its treatment status, which allows the algorithm to decide whether further partitions are appropriate. They also show that the 'honest' approach performs better in estimating the treatment effects for each terminal leaf.

2. The tree is estimated using the treatment effect (rather than the level of the target) using the trial data. In our initial runs we have used ten variables—see below on variable selection—this list is still being finalized and represents an important part of the analysis. The algorithm is that the mean squared error of the target (the treatment effect) is minimized within leaves and maximized across leaves.

3. The tree is pruned using the other half of the trial data using a parameter that penalizes additional leaves (cross-validation). We can also pre-specify the minimum sample size for each terminal leaf to achieve well-balanced trees, so between the minimum leaf size and the penalty for additional leaves versus the gain in within leaf homogeneity and across leaf heterogeneity we arrive at the final tree.

4. The actual treatment effects for each terminal leaf are then estimated using the 50% of the sample that was withheld in step 1—this is the honest estimation proposed by Athey & Imbens.

5. Since any 50% sample split in step 1 can generate a slightly different tree, we repeat the process to generate a random forest, which is a set of trees. We use the bootstrap aggregation method or 'bagging', where we sample with replacement from the original data a sample of the same size as the original data set, and run the tree on this data set. We repeat the process N times (say 100) and generate an average of the characteristics predicted as important based on the 100 trees—this is the random forest.

6. We will estimate four different random forests, one for each of the four countries.

Our first output is to characterize these high achievers, and look for patterns that might lead to a generalizable theory about the types of households that are able to translate small cash transfers into large consumption gains (or alternatively, the types of households that need more support). Our next step is to understand the pathway through which the high achievers realized their gains. Here again we resort to ML because the possible pathways and the combinations of behaviors is quite large, though not as large as in the first phase of the analysis since we are limiting the features to those that are under the control of the household. We use post-treatment variables and will exclude the exogenous, secondary information on context.

The ML approach we propose for the second part of the analysis (the post-intervention behavioral choices) is K-means or hierarchical clustering[16], which identifies discrete

[16] Max Kuhn & Kjell Johnson, 2013, Applied Predictive Modeling, Springer.

groupings based on a set of covariates. In our application, the covariates are behavioral choices the household could make in response to the cash transfer, such as livestock production, non-farm enterprise, crop diversification, overall crop production, use of fertilizer or improved inputs, wage work, crop yield or migration. We could potentially include psychological states if we believed that these were a barrier to decision-making and so represented a constraint that was relaxed by the cash transfer—this would clearly have important policy implications. The algorithm groups together cases for which the 'distance' in their values across all the covariates is minimized. Typically, the researcher specifies the number of clusters or provides a starting value. Hierarchical clustering approaches allow the data to define the optimal number of clusters. In our application the number of clusters depends crucially on the number of high achievers we identify in step one. Note that this is an 'unsupervised' approach because there is no outcome or predictor *per se*; rather units are being grouped based on their values over a set of covariates. We will also conduct this same exercise for 'low achievers' and medium-achievers, which would strengthen our claim that there are distinctive responses made by households that lead to large consumption gains. In other words, a test of whether what we have found is indeed a generalizable theory is if observe that the behavioral responses to the cash transfer were very different for high- versus low-achievers. Note that these choices might be explained by their initial characteristics—this is why we look at both the initial characteristics and post-treatment choices.

The output from these two statistical exercises will provide us with rich insights on consumption growth among the poor and the role of income support. Specifically, we will identify a set of key characteristics (contextual and household) that identify who benefits the most and the least from income support. Second, we will potentially identify the pathways through which successful households achieved large consumption gains (relative to their counterfactual), and the choices made by less successful households. This is the key information we will use to build a theory of graduation. For example, in step one we may identify market access as a key feature of high achievers, and in step two, non-farm enterprise as an important behavioral response among high achievers. Or, we might instead identify climate and land cover or land use as important features in step one (all households are rural in our samples), and crop diversification or cash cropping as the behavioral choice in step two. As another example, and following the framework of Ghatak[17], market friction might be high among both low- and high-achievers, but high-achievers may have high psychological outcomes at baseline, which led to large productive behaviors in response to the cash transfer. The actionable implication is that productive interventions that do not address scarcity-induced behavioral constraints may not be effective. In practice, there are a range of combinations of actions, psychological states and

[17] Ghatak, Maitreesh. 2015. Theory of Poverty Traps and Anti-Poverty Policies. The World Bank Economic Review, Vol. 29, Supplement, pp. S77–S105.

their interactions with the microenvironment that are possible. These combinations of characteristics and behavioral choices, driven by the data, can provide rich insights on the pathways out of poverty, and the role of public policy.

What is the difference between our approach and an approach that pre-specifies characteristics and pathways beforehand and then tests whether they are important, as we described in the beginning of this section? The key differences are that we begin with a large list of potential characteristics, and allow them to interact non-parametrically. There may be unique combinations of characteristics and behaviors that we do not envision, and that we might miss if we pre-specified. The strength of our approach is that the results are driven by the data, and may lead to unexpected combinations that we may not have previously considered. While this is highly innovative, as with all highly innovative approaches, it is also risky as we may not be able to identify any clear patterns in the data. For example, we might discover that high-achievers and low-achievers engage in the same post-treatment behaviors. Or, the CART may not identify sufficiently large groups of households with similar characteristics (leafs are too small). These are key indicators of whether we are indeed able to successfully identify a middle range theory of poverty reduction.

Theoretical Framework

There are important features of our study context that have implications for the existence of poverty traps. Households are extremely poor, many do not engage (or engage minimally) in the cash economy, access to informal credit is available but costly, so most households are liquidity constrained. Imperfect capital markets, and friction in the labor market due to the isolation of households means that the classic separation of production and consumption decisions does not hold. Ghatak (2015) provides a framework for assessing the likely existence of poverty traps. In his model, and with the features of the environment described above, poverty traps are possible, and even more so if there are non-convexities in the production technology. Non-convexities are unlikely for subsistence farming and small-scale non-farm enterprise (NFE), and more likely for cash crop production and slightly larger scale NFE activity. These are key pathways for sustained improvements in consumption, and a predictable cash transfer can directly alleviate the liquidity constraint that propels household out of the trap. Note that scarcity-induced behavioral biases (what Ghatak refers to as internal friction) will exacerbate the potential for poverty traps. To the extent that a cash transfer directly removes or reduces scarcity, this is an additional pathway to graduation. Indeed, the conclusion of Ghatak is that an unconditional cash transfer is the *only single intervention* that can address both external and internal frictions and allow households to escape from a poverty trap. Other interventions discussed require other complementary interventions to allow households to emerge from a trap.

Data and variables

Our survey instruments are comprehensive multi-topic questionnaires including a full consumption module, and also contain psychological and behavioral measures such as affect, subjective well-being, self-reported stress, optimism of the future, and time-discounting, as well as livelihoods and productive activity, crop production, soil quality, livestock, credit, and productive and domestic assets. Table 2 presents a preliminary list of pre-treatment and post-treatment (behavioral) variables that we have constructed for the analysis.

While ML is advertised as a 'data-driven' approach, the researcher still has considerable discretion in selecting the initial set of variables for the analysis. Our strategy in selecting these variables is based on the general theory described above. This theory points to external frictions (contextual factors such as relative isolation and market access), internal frictions (time preference, psychological states), and relative starting points (poverty traps induced by non-convexities). Our variables in Table 2 fit into these three broad categories.

We collected GPS points on households in each country. These will be used to bring in secondary data to provide a rich characterization of the microenvironment of the household, which describe external or market frictions, but also production technology (which is affected by climate, land cover). Table 3 presents a preliminary list of variables we have merged into the data, and which will be used to predict high and low achievers. These variables can also interact with post-intervention behavioral choices—we are still considering whether it makes sense to add these to the k-means clustering analysis. In addition to the variables in Table 3 merged in via geocodes, we also collected detailed prices in each cluster and existence of social infrastructure (schools, health facility) that will use to develop a rich characterization of the context facing each household.

Table 2: Preliminary list of household indicators

Pre-treatment household variables	Post-treatment (behavioural) variables
Demographics	
# of members in different age groups	
Characteristics of the head (age, sex, education)	
Highest education of other adult members	
Productive assets	Productive assets
Axe, hoe, machete, shovel, sickle, watering can	Axe, hoe, machete, shovel, sickle, watering can
Domestic assets	

Bed, table, chairs, radio, bicycle, mobile phone, lantern	
Livestock	Livestock
Chickens, goats, cows, pigs, other fowl	Chickens, goats, cows, pigs, other fowl
Food security	
Worry about food	
Number of meals per day	
Agricultural activity	Agricultural activity
Value of harvest	Value of harvest
Spending on agricultural inputs (fertilizer, labor, seeds, hiring equipment)	Spending on agricultural inputs (fertilizer, labor, seeds, hiring equipment)
Whether sold any produce	Whether sold any produce
Crop diversification (# of different crops produced)	Crop diversification (# of different crops produced)
Non-farm activity	Non-farm activity
Any wage employment	Any wage employment
Revenue from non-farm enterprise	Revenue from non-farm enterprise
Finance and debt	Finance and debt
Whether holds any cash savings	Whether holds any cash savings
Value of cash savings	Value of cash savings
Amount of loan debt outstanding	Amount of loan debt outstanding
Amount of credit outstanding	Amount of credit outstanding
Psychological state and preferences	
Happy with life	
Future will be better in 1 year	
Time preference (wait for future money)	
Stress	

A key issue is the actual measure to use to estimate heterogeneous treatment effects—the high flyers. Microeconomic theory points to consumption as being the ultimate yardstick for living standards and we have a comprehensive, detailed measure of consumption that we will use. However, in the presence of liquidity and borrowing constraints, consumption may closely track income and smoothing will not be possible. In this case consumption is not necessarily a good measure of permanent changes in well-being. An alternative measure would be assets (or wealth), comprising livestock, land, durable goods and perhaps productive assets (agricultural tools). The comprehensiveness of our asset measure within each of these domains varies across surveys, but we do capture all domains in all instruments, and so we will also use assets as an alternative way to capture the phenomenon of high-flyers.

Table 3: Variables to be merged to data via geocodes

Proximity

- Distance to nearest major road
- Distance to nearest population center
- Distance to the nearest market
- Distance to the nearest border crossing

Land Use

- Majority landcover class within 1 km buffer
- Percent agriculture within 1 km buffer
- Percent forest within 1 km buffer
- Land use heterogeneity

Climate

- Annual mean temperature
- Mean temperature of wettest quarter
- Mean temperature of wettest month
- Annual precipitation
- Precipitation of wettest quarter
- Precipitation of wettest month
- Precipitation of driest quarter
- Precipitation of driest month
- Potential wetness index
- Soil moisture
- Drought potential (based on Palmer Drought Severity Index)

Topographic

- Elevation
- Terrain roughness

Qualitative Validation of the Theory

We will conduct qualitative interviews with 15 households in each country to confirm our quantitative findings and provide additional context and understanding of the results. As we do not know the results of the quantitative analysis, we provide the guiding principles behind the design of the qualitative study. Ideally the quantitative results will identify one or two types of high-achieving households and one or two types of low achieving households. By type we mean a set of households that either have a common set of initial (pre-treatment) characteristics or a commons set of post-treatment behaviours or both. We select five households from each of these three groups and conduct in-depth interviews to understand how they used the cash transfer, and how their well-being has evolved over the years since receiving the transfer, and their own opinion about the reasons for the evolution of their well-being. We will probe on themes that come out of the quantitative analysis. For example, if a key pre-treatment characteristic is identified as proximity to a town (Boma), we will probe on whether the household views the Boma has important to its income generation activity. Similarly, if the K-means clustering identifies non-farm enterprise as a key behavioural choice among high achievers, we will probe on the nature of the enterprise, how and why it was started, why it became successful, and so on. The precise design of this component is highly dependent on the results of the quantitative work, so it is hard to provide more specifics at this point. However, the overarching purpose of the component is to validate and provide more detail to support the quantitative results.

The team did consider whether it would be better to begin with qualitative work first to help guide the initial selection of variables for the ML analysis. However, we believe we are able to already capture the likely variables of interest for the quantitative analysis based on theory and our own local knowledge of the context, and of course we are also limited by what is contained in the survey instruments. On the other hand, ground-truthing the quantitative results and understanding exactly how specific behavioural choices improved well-being, and why they were selected, will help us flesh out a more accurate middle-range theory.

Centre of
Excellence
for Development
Impact and Learning

www.CEDILprogramme.org