WPS8412

Efficiency, Legitimacy and Impacts of Targeting Methods

Evidence from an Experiment in Niger

Patrick Premand
Pascale Schnitzer

Abstract

The methods to select safety net beneficiaries are the subject of frequent policy debates. This paper presents the results from a randomized experiment analyzing how efficiency, legitimacy, and short-term program effectiveness vary across widely used targeting methods. The experiment was embedded in the roll-out of a national cash transfer program in Niger. Eligible villages were randomly assigned to have beneficiary households selected through community-based targeting, a proxy-means test, or a formula designed to identify the food-insecure. Proxy-means testing is found to outperform other methods in identifying households with lower consumption per capita. The methods perform similarly against other welfare benchmarks. Legitimacy is high across all methods, but local populations have a slight preference for formula-based approaches. Manipulation and information imperfections are found to affect community-based targeting, although triangulation across multiple selection committees mitigates the related risks. Finally, short-term program impacts on food security are largest among households selected by proxy-means testing. Overall, the differences in performance across targeting methods are small relative to the overall level of exclusion stemming from limited funding for social programs.

Efficiency, Legitimacy and Impacts of Targeting Methods

Evidence from an Experiment in Niger[1]

Patrick Premand (World Bank)

Pascale Schnitzer (World Bank)

Introduction

One of the most frequent topics raised in policy debates on social programs pertains to the choice of methods to select beneficiaries. Although universal coverage of eligible populations is often a long-term aspiration, many countries face limited budgets, especially in low-income settings or in the early phases of setting-up social protection systems. Spreading benefits across a large population would involve substantial transaction costs. In this context, even if some areas may be prioritized through geographical targeting, methods to select beneficiary households are often required.

The challenge of how to select beneficiaries for social programs is particularly salient in low-income settings. Contexts such as the Sahel illustrate the issues. Poverty and extreme poverty are high, but budgets for social programs are small and safety nets have relatively limited coverage. Multiple non-state and humanitarian actors operate with a range of programs using different approaches to select beneficiaries. Governments have stated objectives to build national social protection systems. A better understanding of targeting methods is considered a core element in improving coordination and policy effectiveness.

We present results from a randomized experiment analyzing how efficiency, legitimacy and short-term program effectiveness vary across widely used targeting methods. We worked with the Government of Niger to embed the study in the roll-out of a national cash transfer program. Specifically, 318 eligible villages were randomly assigned to have beneficiary households selected either through community-based targeting (CBT), through a proxy-means test (PMT) formula designed to identify the poor, or through a formula designed to identify food-insecure households (FCS). These three methods were chosen as they are the most commonly used methods to select beneficiaries for safety nets in contexts where means-testing is not feasible, and self-selection would lead to large oversubscription. The experiment tests two of the targeting methods that are most commonly used around the world (CBT and PMT), as well as another method to identify food-insecure households akin to the Household Economic Analysis (HEA), which is commonly used among humanitarian actors in regions such as the Sahel.[2]

This study makes several contributions to the literature on the targeting of social programs. The study analyzes targeting performance of various targeting methods not only by analyzing their relative efficiency, but also their legitimacy among local populations and how they affect program effectiveness. It also assesses a range of determinants of targeting performance. Together with a paper from Indonesia (Alatas et al. 2012), this is the only randomized study on targeting that is based on actual implementation of the methods. The paper complements the literature in several ways.

First, few studies measure the legitimacy of targeting methods by looking at satisfaction among local populations (Alatas et al., 2012; Cameron and Shah, 2014). An important innovation of this paper is that it documents differences in legitimacy across various methods, but also between beneficiaries and non-beneficiaries.

Second, little is known about the actual impacts of targeting choice on social program effectiveness. Studies considering the effect of the choice of targeting method on poverty usually rely on simulations (such as in Brown et al (2017) or Alatas (2012)), or ex post analysis of impact heterogeneity along a range

---

[2] For more information on HEA, see Schnitzer (2016).

of characteristics. This study provides evidence on the *actual* impacts of targeting choice on cash transfer effectiveness.

Third, two of the main reviews on the targeting of social interventions in developing countries conclude that design and implementation (rather than choice of method) are the main factors explaining targeting performance (Coady, Grosh, and Hoddinott (2004) and Devereux et al. (2017)). While the literature has analyzed different mechanisms that can explain performance, in this paper, we consider the relative effect of a range of determinants of targeting performance, such as manipulation, imperfect information, and fatigue in CBT performance. We then show how specific implementation strategies can mitigate observed imperfections. We also assess whether implementation issues such as measurement error or external validity affect formula-based methods. We also examine whether results from the various methods capture the incidence of shocks.

In terms of efficiency, results show that PMT outperforms other methods in identifying households with lower consumption per capita and higher poverty rates. Households selected by PMT have poverty rates that are 8 percentage points higher and consume 15 percent less than households selected by CBT. However, the various targeting methods perform similarly based on other welfare measures such as food security, asset ownership, income per capita or malnutrition. All methods perform well in matching poverty perceptions from the community, particularly PMT and CBT.

Legitimacy is found to be relatively high, with little variations across methods on average. Individuals' perceptions of targeting accuracy show that inclusion and exclusion errors are reported, but with little differences across targeting methods. Importantly, legitimacy is much higher among beneficiaries than among non-beneficiaries, although perceptions of accuracy only differ slightly. Overall, non-beneficiaries tend to prefer the use of formula-based methods, despite perceptions of accuracy being marginally higher for CBT over PMT.

The study highlights a range of determinants that affect targeting performance and shows how various implementation issues can be addressed. Results show that some committee members attempt to manipulate the CBT process, and that there are substantial exclusion errors due to imperfect local knowledge. Still, both these risks are mitigated by triangulating rankings across several committees. Households are aware of manipulation risks, with almost half of respondents reporting that committee members may try to benefit themselves. Fatigue is not found to affect overall efficiency in CBT, although performance seems to slightly worsen for households ranked last. Overall, these results may explain why local populations tend to prefer formula-based methods. Finally, PMT tends to perform slightly worse than CBT in taking into account the effects of shocks, although the magnitude of the differences is not large. Measurement error and external validity are not found to be a major issue in formula-based methods.

The paper is organized as follows. Section 2 provides a brief overview of the literature on targeting performance. Section 3 describes the study design, including a description of the program, experimental design, methods tested, data used, and key sample characteristics. Section 4 presents the main results on the relative efficiency and legitimacy of targeting methods. Section 5 explores the determinants of targeting performance. Section 6 analyzes how the choice of targeting methods affects program effectiveness. Section 7 concludes. Tables are presented in annex I. A supplementary appendix available upon request contains additional tables.

1. Measuring and explaining targeting performance

The literature has often evaluated targeting efficiency based on inclusion or exclusion errors, which aim to measure the extent to which households have been wrongly included or excluded in a program. While informative, these measures are not always well-understood, and do not necessarily provide robust benchmarks as they can be defined and calculated in different ways.

By construction, targeting errors depend on poverty rates and program eligibility thresholds. Consider a country with a poverty rate of 20 percent where safety nets cover 20 percent of the population. If beneficiaries are randomly selected, 80 percent of the beneficiaries would be wrongly included. If the poverty rate changes to 40 percent, then the error under a random selection process would be 60 percent. Any change in either the poverty rate or program eligibility threshold would affect the levels of errors. The fact that inclusion and exclusion errors are a function of poverty rates and eligibility thresholds is often overlooked in the literature comparing targeting performance across programs.

In addition, program eligibility thresholds used to calculate targeting errors are oftentimes not the same as national poverty lines. Calculating inclusion and exclusion errors based on poverty lines may be more relevant if policy makers are interested in the coverage of the poor. Inclusion and exclusion errors calculated based on program eligibility thresholds are on the other hand often determined by budget constraints.[3]

Any analysis of targeting efficiency also needs a clear definition of a welfare benchmark, which often depends on program objectives. Targeting methods may relate to differing concepts of poverty as underlying benchmarks. While social protection policies often assume that targeting households with low levels of consumption will also imply reaching households facing other deprivations, this is not always the case. For example, based on a multi-country study in Sub-Saharan Africa, Brown et. al. (2017) show that malnourished individuals are not necessarily in households presenting low consumptions levels. Similarly, Schnitzer (2016) finds that households facing seasonal food insecurity are not necessarily households with the lowest per capita consumption levels. Besides, community definitions of poverty are not always consistent with income or consumption welfare proxies. Increasing evidence is pointing to more nuanced poverty definitions among local populations (Alatas (2012), Schnitzer (2016), Stoeffler et al. (2016)).

Inclusion and exclusion errors represent only one dimension of targeting performance, and other measures are important to consider. Legitimacy or social acceptability of targeting among local populations is often a concern of its own. In addition, targeting is not an end goal but rather a means to effectively achieve program objectives, such as reducing poverty. In this context, assessing the impact of targeting choice on program effectiveness is particularly relevant. Overall, considering targeting legitimacy or the impact of targeting choice on program effectiveness can provide a broader assessment of targeting performance than simple measures of efficiency such as inclusion or exclusion errors.

---

[3] While the literature is consistent in describing inclusion errors as those wrongly included, the denominator used in these calculations can differ. For example, Ravallion (2009) defines inclusion errors as the share of the ineligible that are wrongly included in the program, while Devereux et. al. (2017) express them as the share of beneficiaries.

There are few studies analyzing the legitimacy of targeting methods among local populations. Results from Indonesia suggest high legitimacy of CBT relative to PMT (Alatas (2012)). Again in Indonesia, Cameron and Shah (2014) find that a mixed targeting approach combining CBT and PMT resulted in significant increases in crime and declines in social capital within communities.

Studies on targeting that consider the ultimate impacts on household welfare have usually relied on simulations under different targeting choices.[4] For example, Brown et al. (2017) and Alatas (2012) have found that the choice of targeting method has limited impact on the effect of cash transfer programs on poverty reduction. Based on simulations using LSMS surveys in 9 Sub-Saharan African countries, Brown et al. (2017) also suggest that, with enough money to cover 20 percent of the population, perfect information and no transactions costs, a well calibrated PMT formula would reduce the Headcount Poverty Index by 23 percent.[5] This compares to 14.5 percent under Universal Basic Income (UBI).[6] Results from such simulations may provide different impacts than actual implementation of targeting methods depending on households' behavioral responses, among other factors.

The literature has highlighted different mechanisms that can explain targeting performance in community-based targeting, such as local knowledge, manipulation, and fatigue. Based on 14 studies on the performance of CBT methods, Coady et al. (2004) find large variations in efficiency, which they attribute to differences in implementation. While the consideration of *local knowledge* has been widely recognized as a key advantage of CBT, there is no direct evidence measuring how information failures at the local level may affect CBT outcomes. *Manipulation* is one of the main concerns with CBT. While the literature abounds with suggestive reports of corruption (McCord 2013), there are studies finding no evidence of such issues (such as in Alatas 2012). Features such as supervision, increased civic participation, sanctions, and institutionalized structures of competition between rival local leaders may minimize manipulation (Watkins, 2008 and Platteau and Gaspert, 2003). Finally, *community fatigue* may negatively impact targeting performance in CBT approaches, if the community is unable to sustain attention. Alatas et al. (2010) find that community effort declines as the ranking exercise progresses.

Various issues can also affect the performance of formula-based methods, such as survey accuracy, external validity, and the incidence of shocks. *Survey accuracy* issues may arise from simple measurement error or manipulation by respondents. In fact, manipulation has been shown to happen over time in some systems relying on PMT (Camacho and Conover, 2011). *External validity* may also be a concern as formulas are developed from samples that may not necessarily represent well program areas or may be outdated. Brown et al (2017) find that lags in implementation considerably matter. Finally, formulas relying on variables that change slowly over time (for example, household demographics and dwelling materials)

---

[4] Targeted households are assumed to receive a fixed transfer amount. Poverty reduction is then measured by looking at the post-transfer consumption levels.

[5] Estimations based on tables 11 and 13 from the paper. Poverty is assumed to be 20% and it is important to note that in reality the poverty rates in the countries studied are usually considerably higher. PMT usually works poorly in predicting the levels of living of the poorest and one may expect that raising the poverty line to a more realistic one may result in better performance. Countries analyzed included Burkina Faso, Ethiopia, Ghana, Malawi, Mali, Niger, Nigeria, Tanzania and Uganda.

[6] While the improvement of PMT over UBI looks small, in a country like Niger, it would still imply covering approximately an additional 350,000 people (of an estimated population size of 20.67 million) if PMT were to be used as opposed to UBI. Also, if using distribution-sensitive poverty measures, the difference in the relative performance of PMT and UBI is substantial – using PMT would decrease the Watts index by 48 percent as opposed to 22 percent under UBI.

may not properly identify households affected by shocks. Del Ninno and Mills (2014) conclude that PMT efficacy remains an open issue with respect to the identification of beneficiaries vulnerable to poverty.

2. Study Design

3.1 Niger Safety Nets Program and Targeting Experiment Design

Niger is among the poorest countries in the world, and resources for social programs are limited. Poverty incidence was estimated at 44.5 percent in 2014, or 51.4 percent in rural areas (World Bank, 2017). This means that 8.17 million people lived in poverty as of 2014, 7.91 million in rural areas. In addition to persistent poverty, there are also recurrent food crises, and even in good agricultural years, between four and five million people experience food shortages.[7]

Social protection programs have traditionally been thin and fragmented in Niger. A large share of resources has been spent on emergency response and humanitarian interventions. These interventions are implemented by various institutions, including a range of non-state actors, with relatively weak government coordination. In 2011, the Government of Niger, with support from the World Bank, set up a national safety nets project. Its objective was to establish an effective and adaptive safety net system that would increase access of poor and vulnerable people to cash transfer and cash for work programs. The cash transfer program provides small, regular transfers or 10,000 FCFA per month (or approximately $20), for 24 months, to poor households. The cash transfer program followed a pilot initiated in 2010 (see Stoeffler et al., 2016), and has expanded over time to reach 100,000 households by 2018.

In a context of overwhelming need but very thin administrative data, there are recurring questions on how to efficiently allocate resources from social programs. Poverty is typically widespread, and the level of inequalities between households is relatively low in rural areas. At the same time, food insecurity affects an important share of households and is largely seasonal (Schnitzer 2016). Since available resources only allow covering a small share of the population in need of support, there are important questions on how to select safety net beneficiaries.

Working with the Government of Niger, we set up a large-scale randomized experiment to test alternative targeting methods. The experiment was embedded in the second phase of the Niger Safety Nets Program implemented from 2016 to 2018. This second phase operated in 5 regions (out of 8). The cash transfer program used geographical targeting to identify 38 high-poverty rural communes.[8] Within these communes, there is little data to determine which villages are poorest. The number of villages to be served by the program in each commune was determined so that on average 40 percent of households would be covered by the cash transfer program. All villages were considered eligible and, to ensure transparency, beneficiary villages were selected through public lotteries in the presence of village authorities (Premand

---

[7] https://reliefweb.int/sites/reliefweb.int/files/resources/niger_2018-02-23.pdf
[8] A range of indicators available at the commune level, together with population data, were used to select and allocate beneficiary quotas for regions, departments and communes.

et al., 2016). For the second phase of the cash transfer program, 558 villages were selected to participate. The geographical targeting process was applied consistently and is not the focus of this paper.

The targeting experiment was implemented in 318 villages selected to benefit from the second phase of the safety nets programs across 18 communes.[9] Once beneficiary villages were randomly identified through public lotteries, a registry survey based on a short questionnaire was collected among all households. The questionnaire contained information to compute PMT and FCS scores for each household. Based on the registry survey, the PMT formula was then applied to determine a quota of beneficiaries per village.[10]

The 318 eligible villages were then randomly assigned to have beneficiary households selected through one of three methods: community-based targeting (CBT), a proxy-means test (PMT) formula designed to identify the poor, or an alternative formula designed to identify food-insecure households (FCS). Specifically, within each commune, each of the 318 villages was randomly assigned to one of the three targeting methods (CBT, FCS or PMT). To ensure balance, within each commune, randomization was stratified using a PCA index based on village population, village average PMT score and Gini coefficient.

After beneficiaries were selected by each of the 3 methods, a community validation exercise was organized. While the community validation exercise aimed to correct for any mistargeting, in practice this final step did not change results much. In fact, over 99% of households identified through each of the methods were validated by the community. This outcome does not necessarily mean that community poverty perceptions are the same as those resulting from the different methods. Anecdotal evidence from the field suggests that during the validation process communities tend to refrain from disagreeing with the presented lists of beneficiaries.

We now describe the application of each method in greater detail, before presenting the experimental design and its implementation.

3.2 Alternative Methods to Select Beneficiary Households for Safety Nets

Proxy Means Testing (PMT)

The PMT method relies on a formula approximating household consumption based on a limited set of household characteristics (Grosh and Baker, 1995). The characteristics include variables such as demographics, asset and or housing information. The information for the formula is both quicker to collect and easier to verify than a consumption aggregate, which requires listing all items purchased or self-consumed over a reference period ranging from 7 days for food to 12 months for non-food expenditures (such as clothes or household durables).

---

[9] The 18 communes were chosen by the implementing agency in order to reflect the diversity of project areas.
[10] A minimum share of beneficiaries per village was set at 10%, and a maximum share at 70% per village. These were rarely binding in practice, however.

In Niger, the PMT method was used systematically during the first phase of the program. Then, during the second phase of the program, it was applied in a subset of beneficiary villages as part of the present experiment. The PMT formula was developed based on a national household survey (LSMS) collected in 2011. The survey was representative at the national level and at the livelihood zone level.[11] The survey included two rounds of data collection, the first during the lean season, and the second during the harvest season. The formula was designed to predict average consumption between the two rounds. Separate models were estimated for each livelihood zone. The formula was derived from a step-wise regression model and relied on a set of variables including household demographics, dwelling characteristics, land ownership, livestock, and assets.

Once beneficiary villages were selected, private firms were hired to collect a registry survey in all households residing in the selected villages. The survey was based on a two-page questionnaire and lasted approximately 30 minutes. Based on these data, the formula was applied, and a PMT score was estimated for each household. Next, within each village, households were ranked according to their score, which together with the village quotas, and after community validation, determined their beneficiary status.

Community-Based Targeting (CBT)

The CBT method relies on village members to select beneficiaries. The process to implement CBT can vary widely in practice, ranging from simple nomination by local leaders to participatory wealth ranking approaches, whereby villagers are asked to rank households according to their poverty level (Coady et. al 2004; McCord 2013).

In the context of the Niger safety nets project, application of the CBT approach followed a well-defined participatory wealth ranking process, with detailed and strict protocols. A village assembly was organized to provide information about the objectives and process of the exercise, and form three selection committees. To ensure that key information was correctly explained throughout, scripts with key messages were developed in the different local languages. Specifically, key messages aimed to (i) recall the objectives and key features of the cash transfer program; (ii) explain the objectives and process to identify program beneficiaries; (iii) guide the community in selecting three committees. Protocols were put in place to maximize participation during the assembly. With guidance from a field agent, all villagers together were asked to form three committees of five members each. Committee members were chosen on the basis that they knew most residents and were trusted by the community. The three committees include a group of leaders (with at least two women), a group of non-leader women, and a mixed group of non-leaders (with at least two women). Finally, within each group and to the extent possible, diversity was encouraged in dimensions such as household location, education, wealth levels, or ethnicity.

Once the three committees were in place, the ranking exercise was conducted separately by each committee. The discussion started by asking members to define the usual characteristics of households that have a low, medium, and high capacity to satisfy basic needs during the entire year. Next, the committees were given a set of pre-printed cards (each representing a household from the registry survey)

---

[11] Niger is subdivided in various livelihood zones, and the survey is representative of the pastoral, agro-pastoral and agricultural zones.

which were read one by one, in a randomly provided order. As each card was read, committee members were asked to place it on a long board containing 3 different colors (red, yellow and green), according to the capacity of each household to satisfy needs during the entire year.[12] When disagreements were observed, the facilitator would recall the characteristics defined during the discussion to help members find a consensus.

Once the process was completed, ranks from each committee were standardized and an average was obtained. If a score for any given household deviated by more than 2 standard deviations from the other two ranks, it was not used. Finally, within each village, households were ranked according to their average score, which together with the village quotas, and after community validation, determined their beneficiary status.

Food Consumption Score Formula (FCS)

The FCS method relies on a formula to identify households at risk of food insecurity during the lean season based on a limited set of household characteristics. In Niger, food insecure households during the lean season are not necessarily those with low consumption levels throughout the year. For instance, based on nationally representative data in Niger, Schnitzer (2016) highlights the relatively weak correlation between consumption and food insecurity during the lean season. Trade-offs may arise in the design of targeting mechanisms for social policy if the poor differ from the seasonally food insecure. The consideration of transient vulnerabilities and shocks in targeting remains a challenge (Del Ninno and Mills, 2014). In the Sahel, many humanitarian actors use a method called the Household Economist Analysis (HEA), which tends to identify households at risk of being seasonally food insecure. Schnitzer (2016) shows that the formula on which the HEA method is based can be further improved to identify food-insecure households. The third method we use here is thus a formula that aims to estimate the food consumption score indicator based on a limited number of easily verifiable variables. The food consumption score is widely used by agencies addressing food crises and aims to reflect both the quantity and quality of food consumed. Specifically, the food consumption score is based on a set of questions asking about the consumption frequency of different food groups during the past seven days (WFP, 2008).[13]

The FCS formula was developed based on data from the lean season of the 2011 LSMS survey. A separate model was estimated for each of the three livelihood zones. The formula was derived from a step-wise regression model and relied on a set of variables including household demographics, dwelling characteristics, land ownership, livestock, and assets.

---

[12] In practice, based on the door-to-door registry survey data, around 7 percent of households in the top part of the PMT distribution were filtered out from the selection process. This step aimed to reduce the burden of the ranking process (by reducing the number of households to be ranked), while at the same time improve targeting efficiency and minimize manipulation. As shown in Schnitzer (2016), PMT can be highly effective at excluding households in the top part of the consumption distribution. Section 5.3 below discusses that this step does not affect the results on targeting performance.

[13] Food groups included cereals and cereal products, tubers and plantains, legumes and seeds, vegetables, fish and meat, fruits, milk and milk products, oil and grease, sugar products, spices and condiments.

As part of the safety nets project, the formula was applied to data from the registry survey. An FCS score approximating the level of food insecurity was assigned to each household.[14] The ranking of households based on this score, together with the village quotas, and after community validation, determined beneficiary status.

3.3 Data and Timeline

Field activities for the study took place between August 2015 and December 2016. Figure 1 illustrates the timing of events.

The first source of data from the experiment comes from a short, large-scale door-to-door registry survey collected in all households in the 318 villages selected to participate in the cash transfer program in the 18 study communes. The survey collected data using a limited set of questions that allowed the application of both the PMT and FCS formulas. In total, 54,051 households were surveyed between August and November 2015.

Shortly after the registry survey, and before any beneficiary selection process was implemented, a baseline survey was collected in a random sample including 12 households in each of the study villages. The registry survey was used as the sampling frame. With 12 households per village, and 318 villages in total, the targeted number of households to be interviewed was 3,816. Of the 3,816 sampled households, 3,479 (or 91%) were interviewed. Among those not interviewed, 70% were households not recognized. This was due to errors in the registry survey, including errors in village identifiers. The remaining were absent. The survey contained detailed information on a number of dimensions, including consumption, food insecurity, anthropometrics (for children under the age of 5), assets, household demographics, education and health status of household members. The baseline survey took place between September and December 2015.

Following the baseline survey, targeting protocols were implemented and households selected for the cash transfer program. The CBT process happened between December 2015 and January 2016. The PMT and FCS formulas were applied in parallel. These steps were followed by the community validation (February to March 2016), registration (March to April 2016) and first cash transfer payments (June to July 2016). The program was scheduled to last for 24 months, until May 2018.

Three to 5 months into the cash transfer program, a follow-up survey was collected to analyze the legitimacy of the various targeting methods, as well as to document how the transfers were being used and assess short-term impacts on a limited set of outcomes. The follow-up survey took place in all 5 regions and 18 communes from the baseline survey, and about half the villages in the baseline were included in the sample.[15] In total, the survey covered 162 villages, in each of which 8 households were

---

[14] As for CBT, around 7 percent of households in the top part of the PMT distribution were filtered out from the selection process based on the door-to-door registry survey data. Section 5.3 below discusses that this step does not affect the results on targeting performance.

[15] Sample size was determined based on power calculations to capture differences in legitimacy outcomes.

selected, including 4 beneficiary and 4 non-beneficiary households.[16] The follow-up survey sample included 1,296 households, of which 94% were successfully interviewed. The follow-up survey was collected between October 2016 and January 2017.

Program administrative data complement the survey data. These include costs, household-level information on registration and payment status, as well as monitoring data from the CBT implementation process.

Figure 1: Study timeline
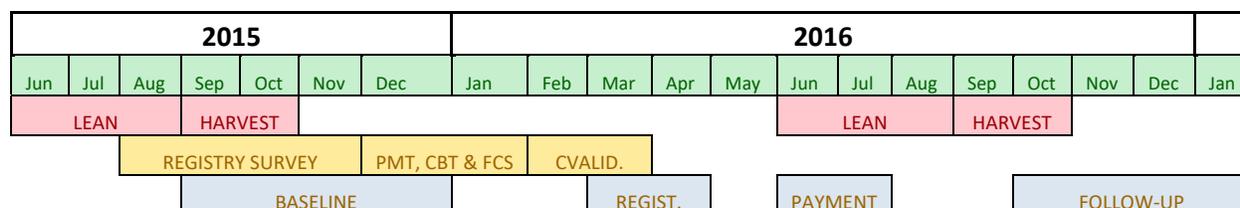
The study uses different measures of welfare, including consumption per capita, food consumption score, assets and livestock indices, income per capita, the food stock coverage, stunting, wasting, and underweight. Consumption includes annual expenditure on food, non-food (including health and schooling) and durables. The asset and livestock indices are constructed using Principal Component Analysis (PCA) based on different asset and livestock variables. Annex II contains detailed definitions of each of the welfare measures.

3.4 Key sample characteristics and balance

Table 1 presents general household-level descriptive statistics of the baseline sample, as well as tests to document balance in characteristics across villages. This is a poor, homogeneous, and uneducated population. Of the sample households, 85 percent are found to be below the national poverty line, compared to a national poverty rate of 48.2 percent for Niger as a whole (INS 2016). The level of inequality between households in project villages is low, with a Gini coefficient of 0.30.[17] Only 1 in 5 households has at least one member who has any formal education. Beneficiary households represent 43 percent of the sample.

Although food insecurity was measured during the harvest season, when it is usually at its lowest levels, it still affected about 13 percent of households at baseline. Households reported that their own agricultural production covered food needs for an average 3.4 months during the year prior to the survey. Households are large with a mean size of 7. Twenty-two percent of all households are polygamous. Nine percent of households are headed by a woman.

---

[16] In some villages, due to sampling variation, the total number of beneficiary or non-beneficiary households sampled at baseline was lower than 4. In these cases, additional non-baseline households were added to the sample. Of the 1,296 households included in the sample, 86% were also sampled at baseline.

[17] If project areas represented a country, it would be among the most equal in the world.

The population is young with an average age of 19.4 years. Children under the age of 12 represent a remarkably large 44% of all individuals.[18] The data highlight that malnutrition is a serious problem. Roughly 3 in 4 children under the age of 5 (72%) are mildly stunted, while close to one in four (23%) is severely stunted. Seventy percent are mildly underweight, and 16 percent severely underweight. Forty-two percent are mildly wasted and 4 percent severely so.

The balance tests suggest that randomization worked as intended. No statistically significant differences are found across villages assigned to different targeting methods across a large range of characteristics.[19]

3. Efficiency and Legitimacy of targeting methods

3.1 Targeting Efficiency

To measure relative targeting efficiency, we first compare the mean characteristics of beneficiary households selected by each of the three methods. As shown in Table 2, we focus on key welfare measures categorized into household indicators (panel A), children indicators (panel B), and subjective well-being indicators (panel C). Specifically, panel A displays poverty status, consumption (per capita, adjusted for economies of scale, and per adult equivalent), food insecurity, an asset index, a livestock index, income per capita and food stock coverage. Panel B contains incidence of severe stunting, wasting and underweight of children under the age of 5. Panel C presents proxies for (i) poverty perceptions by other community members and (ii) self-perception of basic needs. For the first indicator, baseline households were asked to rank other sampled households. For each household, we construct a dummy variable that equals one if the average of rankings provided by other baseline respondents falls below the program cut-off line. The second indicator is measured through a dummy variable that equals one if the ranking based on self-assessed basic needs falls below the program cut-off line.[20] For each of these welfare indicators, columns 1-3 present the mean characteristics for households selected by CBT, FCS, and PMT methods, respectively. Columns 4-6 show the differences in means across the different methods. The statistical significance of the difference in means is estimated through an OLS regression for each welfare measure, which includes explanatory variables capturing the assigned targeting method for each village, as well as commune fixed effects. Robust standard errors are clustered at the village level.

Results show that PMT outperforms the FCS and CBT methods in identifying households with lower consumption per capita and higher poverty rates. Households selected by PMT consume 15 percent less of what households selected by CBT consume, and 11 percent less than what households selected by FCS consume. Similarly, households selected by PMT have a poverty rate higher by 8 percentage points than households selected by CBT, and higher by 5 percentage points than households selected by FCS. The

---

[18] Table A.1 in the supplementary appendix contains additional descriptive statistics for children under 12.

[19] Exceptions include small differences in food consumption score and presence of covariate shocks between CBT and PMT villages (significant at 10%). Small differences in severe wasting are also found between CBT and FCS villages (significant at 5%).

[20] This measure is based on the average response across 6 questions that ask respondents how satisfied they feel (on a scale from 1 to 3), with respect to 6 basic needs, including food, housing conditions, health, education, clothing, and revenues.

differences remain significant, though somewhat smaller in magnitude, when consumption is adjusted by economies of scale, or adult equivalency scales. FCS and CBT methods perform similarly relative to poverty and consumption welfare benchmarks.

Figure 2 provides another illustration of the relative efficiency of targeting methods across the consumption distribution. For each method, Figure 2 plots the likelihood of a household to be selected as beneficiary across the consumption distribution. The PMT is relatively more efficient at including the poorest and excluding the richest, based on consumption per capita. On the other hand, the performance of CBT and FCS is rather flat along the consumption distribution.

Figure 2: Targeting efficiency across consumption distribution
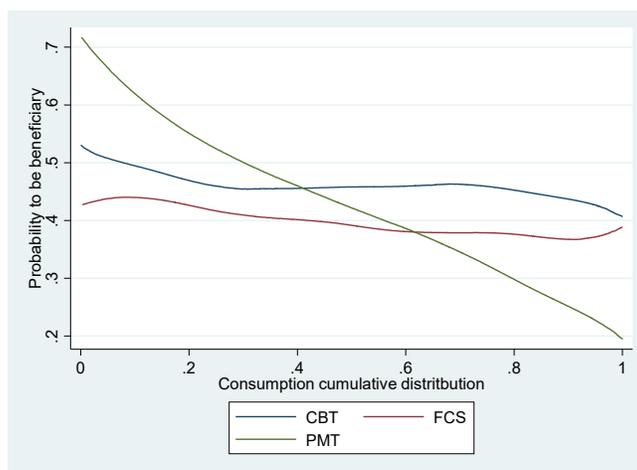
Note: Graph based on semiparametric regression model where commune and village fixed effects enters linearly in the functional form.

Figure 3 summarizes the inclusion errors for each method, compared to a universal program covering everyone. Inclusion errors are calculated based on the consumption per capita benchmark, and separately for two eligibility thresholds, first based on the national poverty line, second based on the program eligibility cut-off.[21] Compared to a benchmark scenario of universal coverage, the PMT method reduces inclusion errors by 60 percent (from 15 percent to 6 percent) when using the national poverty line, and by 32 percent (from 57 percent to 39 percent) when using the program eligibility threshold. The CBT and FCS methods lead to smaller reductions in inclusion errors. The relative performance of methods varies depending on which eligibility threshold is used, because their performances differ along the consumption distribution.

---

[21] Inclusion errors are estimated by dividing the number beneficiaries with consumption per capita above the cut-off by the total number of beneficiaries.

Figure 3: Inclusion errors
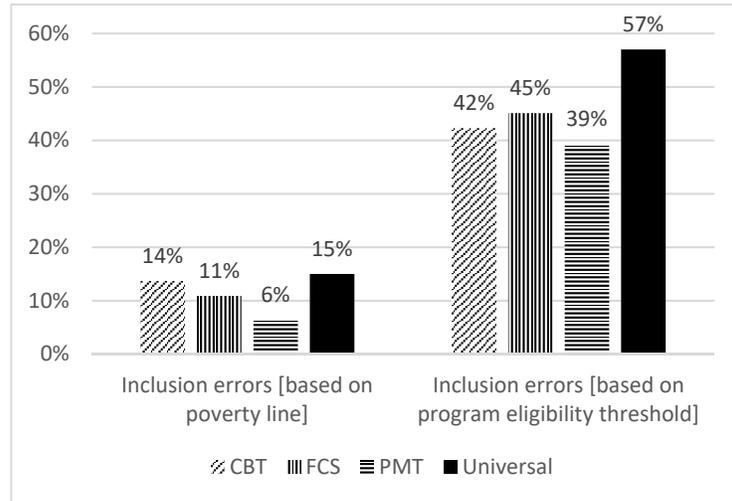
Note: Table A.2 in supplementary appendix summarizes the results.

Importantly, exclusion errors based on the national poverty line are large. This is driven by the overall program coverage due to its limited funding (as opposed to targeting performance). While 85% of households in the project areas are poor, the program can only cover about half of them (43%). This means that, in the best-case scenario, if the program were to correctly identify all the poor, exclusion errors would still be 49%. Given the limited program coverage relative to needs, none of the targeting methods can substantially reduce exclusion errors.[22]

The various methods have similar efficiency against welfare measures other than consumption and poverty. Table 2 shows that there are no robust statistical differences in average characteristics of beneficiary households selected by each method across other welfare benchmarks such as food insecurity, an asset index, a livestock index, income per capita, food stock coverage, as well as stunting, wasting and underweight of children under the age of 5. There are no statistical differences between households selected by CBT and FCS, aside from a slightly higher share of severely wasted children in households selected by CBT. There are no statistical differences in the profile of households selected by PMT and CBT either, aside from a slightly higher share of severely underweight selected by CBT. There are only a couple of small differences between households selected by FCS and PMT, with FCS households slightly less food insecure, and having slightly more food stock. The few observed differences are small in magnitude, only significant at the 10 percent level, and with no clear or robust patterns.

Based on subjective well-being indicators, PMT and CBT perform equally well. For instance, 62 percent of households selected by CBT are considered to belong to the bottom 43 percent of the welfare distribution based on poverty perceptions from other households. Results are largely similar when considering the distribution based on the self-perception of basic needs, with 64 percent of CBT selected households belonging to the bottom 43%. Similar results are observed for PMT: 59 percent of households selected by PMT are considered to belong to the bottom 43% based on the two subjective well-being indicators. These

---

[22] Table A.2 in the supplementary appendix provides the results.

results contrast with those from Indonesia, where CBT performed considerably better than PMT in matching community perceptions of poverty (Alatas et. al. 2012).

Table 3 summarizes the differences in the profile of beneficiary households, compared to non-beneficiaries, for each method. The CBT process leads to the selection of households with fewer assets, fewer livestock, lower income per adult equivalent, and lower food stock. It also leads to the selection of households with slightly less consumption, but only based on the adult-equivalent measure. The FCS method leads to the selection of households with slightly lower consumption, and fewer household assets. The PMT method leads to the selection of households with higher poverty, lower consumption, as well as lower income per adult equivalent and food stock. PMT also selects households with a slightly lower share of underweight children. None of the methods is able to identify food insecure households (as measured by the food consumption score during the harvest season). We return to this issue below. All methods are well-aligned with subjective well-being indicators, and this is especially the case for CBT and PMT. For instance, poverty (based on perceptions of others) of CBT and PMT beneficiaries is 28 and 25 percentage points higher than non-beneficiaries.

3.2 Legitimacy among local populations

We now turn to one of the innovations of the study, which is to document whether the legitimacy of the various targeting methods varies among local populations. To analyze legitimacy, we focus on four main questions from the follow-up survey. The first two questions capture legitimacy based on (i) whether respondents are satisfied with the overall targeting method,[23] and (ii) whether respondents would like the same method to be used again in similar programs in the future. The following two questions capture perceptions of accuracy, namely (i) whether respondents believe that *any* households were wrongly included; and (iv) whether respondents believe that *any* households were wrongly excluded from the program. For each of these variables, we estimate an OLS regression that includes indicators for the targeting method assigned to the village (using CBT as the excluded category), as well as commune fixed effects. Robust standard errors are clustered at the village level. We first assess differences across methods by pooling all respondents, and then disaggregate the results between beneficiaries and non-beneficiaries.

Legitimacy is found to be relatively high across targeting methods (Table 4). Overall, only 2 percent of households reported to be "unsatisfied", 10 percent a "little satisfied", 40 percent "satisfied", and 49 percent "very satisfied" with the targeting approach. There are no statistical differences in satisfaction across methods. Also, 83 percent of respondents stated that they would like the same approach to be used again in the future. Small differences are observed, suggesting that respondents tend to prefer the formula-based method to be repeated, when compared to CBT.

Perceptions of the accuracy of each method show that inclusion and exclusion errors are reported, but again with little differences across targeting methods. On average, 36.4 percent of respondents report that at least one non-poor household has been selected. Exclusion errors are more prevalent, with 55.5 percent of respondents reporting that at least one poor household has been excluded. Note that it is

---

[23] This is measured on a scale from 1 to 4 (1, unsatisfied; 2 a little satisfied, 3 satisfied, and 4 very satisfied).

expected that exclusion is more frequent, since the program operates in high-poverty areas, and does not have resources to cover all the poor in selected villages. There are no statistical differences in perceptions of accuracy across PMT and FCS, or FCS and CBT. The only small difference observed is that respondents tend to report slightly more inclusion issues with PMT compared to CBT.

There are important differences in legitimacy and perception of accuracy across beneficiaries (Table 5) and non-beneficiaries (Table 6). In general, legitimacy is much higher among beneficiaries than among non-beneficiaries, although perceptions of accuracy differ less. Most beneficiaries are "very satisfied", with an average satisfaction of 3.7 (on a scale of 4). This is contrasted with an average satisfaction of 2.9 for non-beneficiaries. Also, 94 percent of beneficiaries wish the same targeting approach to be used in the future as opposed to 71 percent of non-beneficiaries. Beneficiaries are also 5 and 12 percentage points less likely to report that there is at least one poor or rich household wrongly identified by the program respectively.

While differences in legitimacy and perceived accuracy across methods are limited for the overall sample, stronger differences are found among non-beneficiaries. Non-beneficiaries are particularly unlikely to want the CBT approach to be repeated: in fact, they are 7 and 11 percent more likely to want to repeat the PMT and FCS approach respectively. This is despite the fact that non-beneficiaries in villages where the PMT approach was applied report more inclusion issues: non-beneficiaries in PMT villages are 12 percent more likely to say that there was a non-poor household included in the program than non-beneficiaries in villages where the CBT approach was used. Overall, non-beneficiaries tend to prefer the use of formula-based methods, despite perceptions of accuracy marginally favoring CBT over PMT. In contrast, for beneficiaries, almost no significant differences are found across methods in terms of legitimacy and perceived accuracy. The only exception is that beneficiaries selected by FCS are less satisfied, both compared to CBT and PMT, but this is only marginally significant.

In the next sections, we turn to analyze a range of factors that can explain differences in efficiency and legitimacy.

4. Determinants of targeting performance

We now turn to analyzing a range of factors that can potentially affect the performance of each targeting method. First, we assess whether imperfect information, manipulation and fatigue affect CBT performance. Second, we examine the extent to which the incidence of shocks may affect the performance of methods. Lastly, we assess whether survey accuracy, external validity and the application of filters affect formula-based methods.

4.1 Imperfect Information, Manipulation and Fatigue in Community-Based Targeting

While it is often assumed that the consideration of local knowledge is a key advantage of CBT, there is limited evidence on the role of local information failures in targeting efficiency. Households that are not known by other village members may be more likely to be wrongly excluded in the CBT process. We test

this hypothesis by creating a proxy indicator of local knowledge. Specifically, the baseline survey asked each household to provide a welfare ranking of the other (11) sampled households in the same village. In some cases, respondents would report not knowing another household. We use this information to create a proxy variable for "local knowledge" that is the share of sampled households in a village knowing a given household. On average, we find that households are known by 73% of other households residing in the same village. We run a regression of exclusion errors on the local knowledge dummy. We do this separately for rankings from each of the 3 committees, and for the combined ranking. Results are presented in Table 7.

Results show substantial exclusion errors due to imperfect local knowledge in rankings from individual committees. However, the triangulation of results across the three committees fully corrects for this imperfect information. A household that is known by a higher share of households in the village is substantially less likely to be excluded by error by individual committees. The effects are substantial, with exclusion probabilities increasing between 32 and 38 percent depending on the committee. Figure 4 illustrates that the exclusion error linearly decreases with local knowledge. However, when selection results are based on the *combined* rankings from the three committees, households less well-known are not more likely to be excluded. Therefore, while each committee taken individually has imperfect information, the creation of multiple committees effectively helps address information asymmetries.

Figure 4: Exclusion error (leaders' committee) across the "local knowledge" distribution
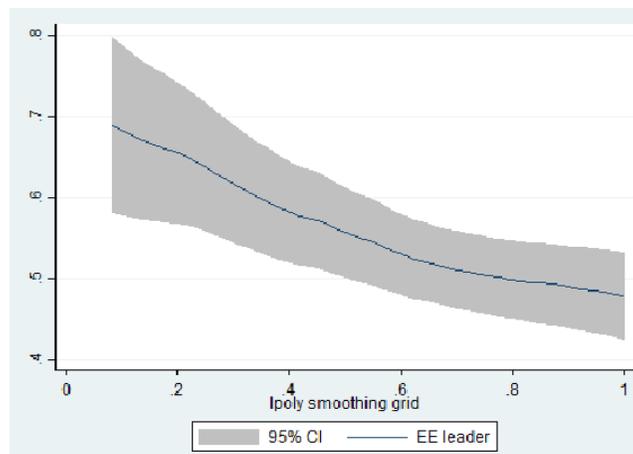
Note: graph based on semiparametric regression model. *Local knowledge* is defined as the share of sampled households in a village knowing a given household.

Another concern with CBT is that committee members may manipulate the selection process. If committee members come from better-off households and try to benefit themselves, the performance of CBT will worsen. As mentioned above, when the CBT was applied in the context of the Niger safety nets project, three independent committees were set up in each village. CBT committee members represent a median of 12 percent of village households. To assess whether any manipulation occurs, we first look at the ranking provided by each of the three independent committees and analyze whether committee

members' households have a higher chance to be selected as beneficiaries. Estimation controls for household welfare measures in Table 2.[24]

Results show that some committee members attempt to manipulate the process, but that the risk of manipulation is partly mitigated by triangulating rankings across several committees (Table 8). After controlling for welfare measures, the probability that committee members' households are selected is higher than for other households. This difference is statistically significant for two of the three committees. There is no significant manipulation in the women's committee, but households of the mixed non-leader and leader group have a 20.4 and 14 percentage point higher chance to be selected as beneficiary, respectively. At the same time, the triangulation of rankings from the three committees largely reduces the risk of manipulation; based on the final ranking, committee members have an 8.2 percentage point higher chance to be selected as beneficiaries, and this is only marginally significant (at the 10% level). In sum, results show that the triangulation of results and careful consideration of committee composition can play an important role in mitigating manipulation. Results are consistent with the perceptions of 48 percent of the survey respondents, who declare that committee members may try to benefit themselves, including 37 percent who believe this very strongly.[25]

A similar analysis was used to assess whether committee members' relatives also had a higher chance to be selected as beneficiaries, and no significant differences were found, suggesting that risks of manipulation primarily stem from committee members trying to favor their own household.[26] Finally, we also looked at how inclusion errors change when restricting the sample to non-committee members, and find no significant difference.[27] This is in part be due to committee members representing only a small fraction of households. Ultimately, even if manipulation occurs, it does not significantly affect the overall targeting efficiency of CBT.

Another concern with CBT targeting is that committees may become fatigued. To assess the role of fatigue, we randomized the order in which households were ranked by committees within each village. Therefore, we can test whether being considered late during the committee deliberation affects the precision of the ranking. Figure 5 shows that inclusion errors start to slightly increase at around the 50th household ranked. Table 9 shows that, although households ranked within the first 50 households present inclusion errors that are 5 percentage points lower than households ranked later, this difference is only marginally insignificant. While Alatas et al. (2012) also find a fatigue effect in Indonesia, in this study the overall effects of fatigue on performance are more muted. Results suggest that committees are less efficient ranking households considered last, the effects are not statistically significant.

---

[24] Controls include consumption per capita, the food consumption score, assets, livestock, food stock, and poverty perceptions by other community members.
[25] Table A.3 in the supplementary appendix displays these descriptive statistics. In addition, there is a correlation between respondents who believe that committee members may try to benefit themselves, and who report not wanting the community-based targeting approach to be repeated.
[26] Table A.4 in the supplementary appendix shows these results.
[27] See Table A.5 in the supplementary appendix.

Figure 5: Inclusion errors along the rank order distribution
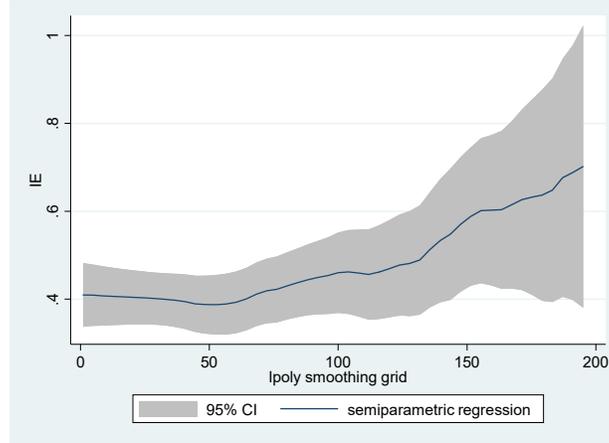
Note: Graph based on semiparametric regression model. The rank order corresponds to the order in which a given household was ranked during the ranking process.

4.2 Shocks and seasonality

Some targeting methods may be more or less effective in accounting for transient poverty factors, for instance resulting from the incidence of shocks. PMT is based on relatively time-invariant factors that can predict consumption and chronic poverty, but may lack capacity to identify households that are transient poor following a shock (Del Ninno and Mills, 2014). In contrast, communities may be better able to identify transient poor households affected by negative shocks. The FCS method also aims to estimate risks of food insecurity and as such to better capture fluctuations in welfare. To explore these considerations, we analyze the extent to which the incidence of shocks explains exclusion errors across targeting methods. To do so, we estimate regressions where the dependent variable is the probability that a household has been wrongly excluded, and the explanatory variables include the interaction between a shock index and the various targeting methods. The shock index is defined as the standardized number of shocks a household reported suffering from during the past 12 months (from when the baseline survey took place). Shocks considered included droughts, floods, pests, as well as livestock, price, health, or household member death related shocks. Results are presented in Table 10.

We find that shock-affected households selected through the PMT have a somewhat higher chance to be excluded than shock-affected households selected through the CBT. Specifically, households in PMT villages are 5 percentage points more likely than households in CBT villages to be excluded when the intensity of shocks increases by one standard deviation. While these estimations are indicative on the role of transient factors in explaining targeting performance, the approach has some limitations. Despite recent evidence from Niger suggesting an important correlation between self-reported and actual shocks (Gao and Mills 2017), self-reported shocks may still suffer from reporting bias, and already capture households' risk-management abilities (Macours et al., 2012).

Finally, the FCS formula is not found to be very effective in identifying households that are food insecure, at least based on measures during the harvest season. This may be explained by different reasons. The FCS formula was developed to measure food insecurity during the lean season, while the current indicator was measured during the harvest season (because of the project timing, data could not be collected during the lean season). The correlation of the food consumption score between the lean and the harvest seasons is limited, and during the harvest season the food insecurity prevalence is usually much lower (Schnitzer 2016).

4.3 Implementation issues in formula-based methods

Finally, we turn to potential implementation issues with formula-based methods. This includes (i) measurement errors in the registry survey, (ii) external validity of the formulas, and (iii) the use of exclusion filters. Overall, results show that these implementation aspects have little effects on targeting method performance.

Measurement errors may affect the performance of formula-based targeting methods. Errors are expected to happen. Registry surveys are usually collected at a very large scale, often under tight time pressures and with less quality control than for more advanced surveys. A particular source of concern is that respondents may manipulate their answers to increase their chances of being selected, creating a strategic response bias. This risk may be higher when the selection formula is widely understood, and has been used for some time (Camacho and Conover, 2011). To assess the extent of measurement error, we use the fact that the baseline survey included similar variables as those in the registry survey used to apply the PMT and FCS formulas. We can thus compute a PMT and FCS score both from the registry survey and baseline survey. We can then compare the two PMT and the two FCS scores to (i) assess the size of the measurement error; (ii) assess how these errors affect targeting performance; and (iii) check if these errors are correlated with poverty. In interpreting these differences, we assume that baseline survey responses are those reflecting the real status of the household, and any deviations with the registry survey are errors. Of course, this assumption does not necessarily hold, and in practice there can be various reasons for observed differences, including real changes in household characteristics, or errors in the baseline survey itself. Still, the comparisons provide a conservative, upper bound estimate of the effect of measurement error on targeting efficiency. In reality, the administrative survey errors are likely to be lower.

We find that measurement errors in the registry survey affect targeting performance, but that the effects are small. Figure 6 shows that there is substantial variation between the scores estimated from the registry survey and the baseline survey, and this is especially the case for the FCS score. The fact that the FCS score fluctuates more may be due to different reasons. First, variables contributing to the food consumption score may be more prone to measurement errors. Alternatively, as demonstrated in Schnitzer (2016), unlike consumption, food security can fluctuate more between seasons. Part of the registry data was collected during the lean season, while the baseline was mostly collected during the harvest season. As such, differences may be due to seasonal changes as opposed to measurement errors. The errors also seem randomly distributed around zero along the consumption and food consumption score distribution (Figure 7). This suggests that, even if there are strategic responses, these are not correlated with the welfare measures used. Table 11 and Table 12 illustrate how measurement errors

affect inclusion errors. Survey errors in PMT and FCS scores significantly affect inclusion errors by 6 percentage points based on consumption and food security, respectively. This magnitude is small, and again should be considered an upper bound attributing all measurement error to the registry survey.

Figure 6: Percentage change in FCS and PMT scores between registry and baseline survey
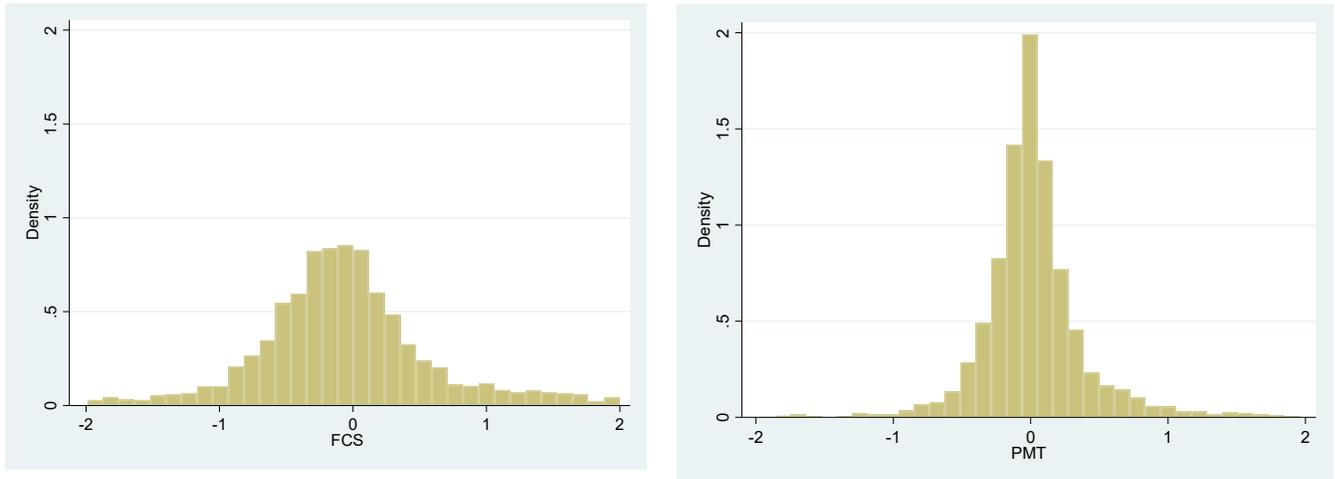
Note: Figures presents the distribution of percentage changes in FCS and PMT scores between the registry and baseline surveys.

Figure 7: Differences in FCS and PMT scores along the consumption distribution
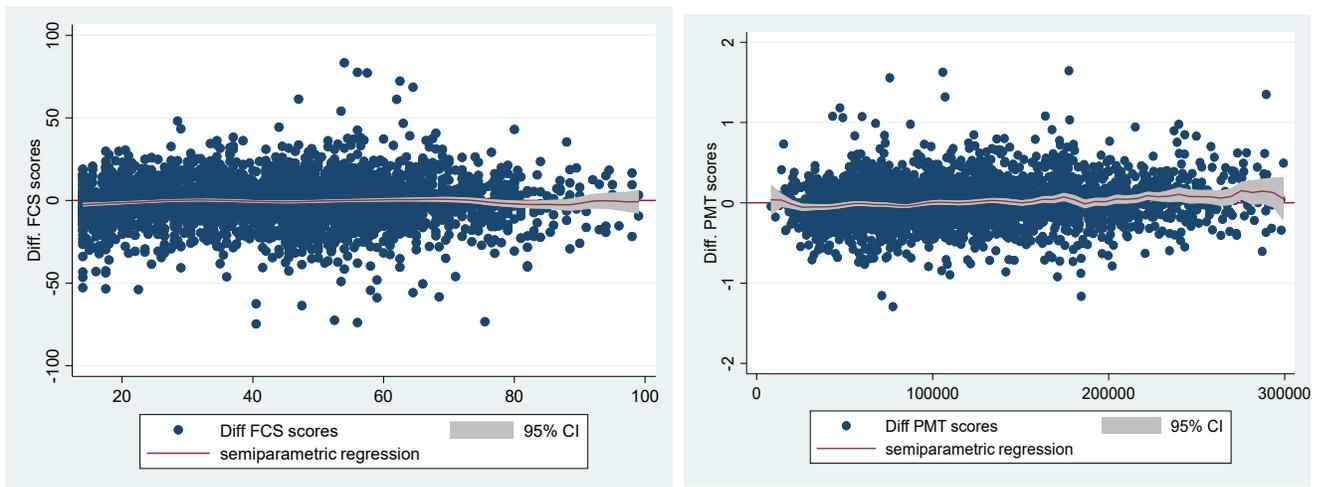
Note: Dots represent data points corresponding to differences in FCS and PMT scores between the registry and baseline surveys. Red lines represent a semiparametric regression of the differences in scores on consumption. Grey areas represent the 95% confidence interval of such regressions.

We further assess whether formulas based on the nationally representative survey are valid for the project area. We find that external validity of the formulas does not affect targeting efficiency. The PMT and FCS methods are based on a formula developed from a nationally representative survey. The survey was collected a few years earlier, and as such some of the determinants of poverty may have changed. At the same time, there may be specificities in the poor communes where the project operates, and the formula may not be as effective at differentiating the poor from the non-poor in these areas. To assess whether external validity of the formula is an issue, we develop new weights for the formula based on the baseline survey, and check if performance changes. Tables 13 and 14 provide the results, and show that external validity does not seem to be a concern. In fact, if the recalibrated model were to be used, inclusion errors would not be significantly different for any of the models.

Finally, we also tested whether the exclusion criteria that were used prior to applying the CBT and CFS formulas explain part of their performance. The filter does not seem to matter for any of the key results.[28]

5. Short-term impacts of cash transfers across targeting methods

In this section, we analyze whether short-term impacts of the cash transfer program vary across beneficiaries selected by different targeting methods. This analysis complements earlier analysis of targeting performance related to efficiency and legitimacy. It tests whether the choice of targeting method has implications for program effectiveness. The analysis is based on a short-term follow-up survey collected a few months after the start of the cash transfer program. The survey also included a few short-term outcome indicators. It did not include longer-term measures or a full consumption aggregate. The results do not seek to provide a full picture of program effectiveness (see Stoeffler et al. (2016) or Premand et al. (2016) for analysis of the impacts of earlier phases of the program), but rather to highlight if there are some differences in short-term impacts across targeting methods.

We start by analyzing whether households selected based on alternative targeting methods report using cash transfers differently, and find no statistically significant differences. Table 15 documents (i) the share of households declaring spending cash transfers across various categories, (ii) average spending amount, and (iii) average spending shares by categories. Households report spending between 65 and 69 percent of the transfers on food. Savings account for an additional 11-12 percent. This is followed by smaller expenditures for the purchase of durable goods, health expenditures, transfer to other households and education expenditures. Importantly, there are no significant differences in the use of transfers across the groups selected by alternative targeting methods.

We proceed to analyze short-term impacts of the cash transfer program across groups selected by the various targeting methods. We estimate program impacts on beneficiaries only. While targeting methods were randomized at the village level, only on average 43 percent of households per village were selected as beneficiaries. Although average differences at the village level could be estimated relying on the randomized design, they would not be sufficiently powered to detect meaningful differences across groups since only a sub-set of each village is covered by the program. Therefore, we focus on analyzing

---

[28] See tables A.6 and A.7 in the supplementary appendix.

impacts on beneficiary households. Doing so requires using quasi-experimental methods since there are differences in the profile of households selected by the various targeting methods, as documented in section 3. We rely on a difference-in-differences approach, focusing on four key outcomes observed at both baseline and follow-up. Outcomes include the food consumption score, food coverage (in months) using own agricultural production, a hunger scale, and a question about whether any conflicts took place between households. See annex II for a detailed description of each indicator.

We run a regression of the outcome variables on a dummy capturing the targeting method assigned to a village, a dummy for each of the two survey rounds, and interactions between the round dummy and the targeting method dummies. The latter parameter provides an estimate of the difference in program impacts across groups selected by different targeting methods. Table 15 presents the results. As a placebo test, we run the same analysis on non-beneficiaries only. We find no statistical differences,[29] which provides reassurance about the equal trend assumption on which the difference-in-differences estimation procedure relies.

Results suggest that short-term program effectiveness is larger along some dimensions in the group selected by the PMT method (Table 16). Specifically, households selected by PMT display a significantly higher program impact on food security compared to households selected by CBT. This difference is significant at the 10 percent level, but represents a substantial magnitude, namely an 11 percent increase in impacts on the food security scale. Households selected by the PMT also display a significantly higher program impact on food stock compared to both households selected by CBT and FCS. The magnitude of this difference is also quite substantial, as it reveals a 23 to 37 percent increase in impacts on food stock.[30] No significant differences in impacts on the hunger scale, or a variable capturing the frequency of conflict, are observed across beneficiaries selected by alternative targeting methods.

6. Conclusion

This study provides a comprehensive analysis of the performance of three targeting methods based on their actual implementation in a randomized experiment in Niger. The methods include community-based targeting (CBT) and proxy-means testing (PMT), two of the most commonly used targeting methods for safety nets around the world, as well as a food insecurity proxy formula (FCS) seeking to identify households at risk of food insecurity, akin to a method used by humanitarian actors in areas such as the Sahel. The paper analyzes targeting efficiency by looking at characteristics of beneficiaries selected by each method along a broad range of welfare benchmarks and subjective well-being indicators. Since the experiment is based on actual implementation of the methods, we can measure the legitimacy of each selection approach among local populations. We also analyze how program effectiveness varies among sub-groups selected by each method.

---

[29] See Table A.8 in the supplementary appendix for results.

[30] These results are consistent with the project timeline. Payments started during the lean season, and the follow-up survey occurred towards the end and after the harvest season when households may have accumulated food stocks.

The results show that in the context of the study PMT is more efficient than other methods in identifying households with lower consumption per capita. The methods perform similarly with respect to other welfare benchmarks, including community perceptions of poverty. Legitimacy is high across all methods. Formula-based methods are not found to be less legitimate that community-based targeting among local populations, but in fact seem to be preferred. The analysis of a range of determinants of targeting performance contributes to explain why. While community-based targeting can incorporate information available at the local level, for instance related to the incidence of shocks, it is affected by imperfect information, manipulation, and, to a lower extent, fatigue. Local populations recognize some of these issues. While concerns about manipulation and imperfect information in community-based targeting are in part justified, the associated risks can be largely addressed through careful implementation, including the triangulation of results across several independent selection committees. Implementation issues and measurement error are not found to strongly affect the performance of formula-based methods, but these methods have limited ability to take into account determinants of transient poverty and seasonality.

In the end, the decision on which targeting method to use should be driven by a cost-benefit analysis. In the context of the study, the results suggest that PMT is the method that may maximize program effectiveness, at least in the short term and based on a limited set of outcomes. Households selected by PMT tend to display significantly higher short-term program impact on food security and food stock than households selected through CBT, and, to a lesser extent, FCS. On the other hand, the costs of methods studied in this paper are largely similar; the cost of PMT or FCS per household is roughly $6.8 USD, while the cost of CBT per household is $5.4 USD, which represents lower bound estimate.[31]

The results obtained in the context of Niger contrast with those from other settings, most notably from Indonesia (Alatas et al., 2012). While both studies do not find large differences in targeting efficiency between methods, the results on legitimacy among local populations differ. In Indonesia, communities seemed to have a preference for CBT over PMT, which is not the case in Niger. The factors driving targeting performance can help explain the differences. For instance, in the context of Niger, imperfect information at the local level and manipulation are shown to affect results from community-based targeting, while these factors were not observed in Indonesia. Other factors may also play a role, including the prevalence of poverty, local inequalities or the degree of trust among communities, in addition to the way targeting methods are communicated to local communities and implemented. As such, the results highlight the importance of carefully considering factors that may affect the relative performance of alternative targeting methods in a given context, as well as taking a broad view in analyzing targeting performance beyond efficiency.

Many low-income countries, like Niger, have limited funding for social programs and low coverage of safety nets, despite high prevalence of poverty. Although PMT can be seen as performing slightly better than CBT and FCS in the context of the study, differences in efficiency, legitimacy and effectiveness remain limited. The results can thus also be interpreted as suggesting that the choice of methods to select safety net beneficiaries does not make a large difference. Variations in performance across targeting methods are small relative to the overall level of exclusion stemming from limited funding for social programs. In

---

[31] This is a lower bound estimate as household listings were based on a prior registry survey, and any CBT process would also require collecting complementary data on households to be registered. Estimates for PMT, FCS, and CBT are based on variable costs (including costs related to field staff and logistics), and exclude fixed costs related to program implementation (other than targeting), such as program staffing costs.

this sense, policy discussions at early stages of social protection systems may not need to focus on identifying one optimal targeting method. Instead, they could focus on ensuring a sufficient level of coverage while facilitating coordination and consolidation of data across programs through mechanisms such as social registries, which can support the application of alternative targeting methods.

References

Alatas, Vivi, et al. "Targeting the poor: evidence from a field experiment in Indonesia." *The American economic review* 102.4 (2012): 1206-1240.

Brown, Caitlin, Martin Ravallion, and Dominique Van de Walle. *A poor means test? Econometric targeting in Africa.* No. w22919. National Bureau of Economic Research, 2016.

Camacho, Adriana, and Emily Conover. "Manipulation of social program eligibility." *American Economic Journal: Economic Policy* 3.2 (2011): 41-65.

Cameron, Lisa, and Manisha Shah. "Mistargeting of cash transfers, social capital destruction, and crime in Indonesia." *Victoria: Monash University* (2011).

Coady, David, Margaret E. Grosh, and John Hoddinott. *Targeting of transfers in developing countries: Review of lessons and experience.* Vol. 1. World Bank Publications, 2004.

Del Ninno, Carlo, and Bradford Mills. *Safety Nets in Africa: Effective Mechanisms to Reach the Poor and Most Vulnerable.* Washington, DC: World Bank; and Agence Française de Développement, 2015.

Devereux, Stephen, et al. "The targeting effectiveness of social transfers." *Journal of Development Effectiveness* 9.2 (2017): 162-211.

Gao, J. and Mills, B. *Analysis of Household Resilience Strategies in Niger.* Draft, June 2017.

Grosh, Margaret, and Judy L. Baker. "Proxy means tests for targeting social programs." *Living standards measurement study working paper* 118 (1995): 1-49.

INS. Agriculture et Conditions de vie des menages au Niger. Institut National de la Statistique, Niger.

Kennedy, Gina, Terri Ballard, and Marie Claude Dop. "Guidelines for measuring household and individual dietary diversity. Nutrition and consumer protection division, food and agriculture organization of the United Nations. 2011." (2015).

Macours, Karen, Patrick Premand, and Renos Vakis. "Transfers, diversification and household risk strategies: experimental evidence with lessons for climate change adaptation." (2012).

McCord, A. *Community-based targeting in the Social Protection sector.* London: ODI Working paper 514. (2013).

Platteau, Jean-Philippe, and Frederic Gaspart. "Disciplining local leaders in community-based development." *Centre for Research on the Economics of Develpoment http://www. worldbank. org/research/wdr/WDR2004/papers/platteau. Pdf* (2003).

Ravallion, Martin. "How relevant is targeting to the success of an antipoverty program?." *The World Bank Research Observer* 24.2 (2009): 205-231.

Schnitzer, Pascale. *How to Target Households in Adaptive Social Protection Systems?: Relative Efficiency of Proxy Means Test and Household Economy Analysis in Niger.* World Bank, 2016.

Watkins, Ben. "Alternative methods for targeting social assistance to highly vulnerable groups." *A report prepared by Kimetrica for Care International* (2008).

Table 1: Balance tests (Differences between average baseline characteristics of households in villages assigned to different targeting methods)

VARIABLES	(1) All	(2) CBT	(3) FCS	(4) PMT	(5) CBT-FCS	(6) CBT-PMT	(7) FCS-PMT
PANEL A: HOUSEHOLD INDICATORS							
Poor	85.4%	84.7%	86.0%	85.6%	-1.3%	-0.9%	0.4%
	(0.35)	(0.36)	(0.35)	(0.35)	(0.02)	(0.02)	(0.02)
Consumption PC	120	121	119	121	2	0	-2
	(66)	(65)	(66)	(65)	(3)	(3)	(3)
Consumption AE	168	170	166	167	3	3	-1
	(92)	(93)	(92)	(91)	(5)	(5)	(5)
Consumption EOS	143	144	142	144	2	0	-2
	(76)	(76)	(77)	(76)	(4)	(4)	(4)
Food insecure	13.40%	13.30%	12.70%	14.20%	0.60%	-0.90%	-1.50%
	(0.34)	(0.34)	(0.33)	(0.35)	(0.02)	(0.01)	(0.02)
Food consumption score	44.11	44.65	44.52	43.14	0.13	1.51 *	1.38
	(18.67)	(18.77)	(19.22)	(17.94)	(0.89)	(0.90)	(0.88)
Asset index	0	-0.04	0.06	-0.02	-0.1	-0.02	0.08
	(1.58)	(1.41)		(1.59)	(0.07)	(0.08)	(0.07)
Livestock index	0	0.08	-0.04	-0.03	0.11	0.11	-0.01
	(1.46)	(1.54)	(1.45)	(1.40)	(0.09)	(0.10)	(0.07)
Income per AE	111	110	114	109	-4	1	4
	(157)	(106)	(222)	(109)	(9)	(6)	(9)
Food stock coverage	3.39	3.34	3.46	3.35	-0.12	0	0.12
	(2.45)	(2.65)	(2.43)	(2.26)	(0.13)	(0.12)	(0.12)
Household size	7.01	7.07	7.13	6.81	-0.05	0.26	0.31
	(3.65)	(3.73)	(3.69)	(3.52)	(0.17)	(0.17)	(0.17)
Polygamous	22.30%	21.70%	24.10%	20.90%	-2.40%	0.80%	3.20%
	(0.42)	(0.41)	(0.43)	(0.41)	(0.02)	(0.02)	(0.02)
Female headed	9.00%	8.80%	8.10%	10.20%	0.60%	-1.40%	-2.00%
	(0.29)	(0.28)	(0.27)	(0.30)	(0.01)	(0.01)	(0.01)
Share of dependents	26.60%	26.40%	26.60%	26.80%	-0.30%	-0.40%	-0.10%
	(0.17)	(0.17)	(0.17)	(0.18)	(0.01)	(0.01)	(0.01)
Any handicap	7.50%	6.80%	8.00%	7.70%	-1.10%	-0.80%	0.30%
	(0.26)	(0.25)	(0.27)	(0.27)	(0.01)	(0.01)	(0.01)
Education head	20.20%	19.10%	20.10%	21.40%	-1.00%	-2.30%	-1.30%
	(0.40)	(0.39)	(0.40)	(0.41)	(0.02)	(0.02)	(0.02)
Land ownership	86.00%	85.90%	85.00%	87.00%	0.90%	-1.20%	-2.00%
	(0.35)	(0.35)	(0.36)	(0.34)	(0.02)	(0.02)	(0.02)

Plots cultivated	1.4	1.38	1.4	1.41	-0.02	-0.03	-0.01
	(1.02)	(1.00)	(1.05)	(1.02)	(0.05)	(0.06)	(0.05)
Cultivated land (hectares)	4.52	4.46	4.6	4.49	-0.15	-0.04	0.11
	(7.08)	(7.27)	(7.03)	(6.94)	(0.33)	(0.37)	(0.36)
Covariate shocks	27.50%	26.80%	27.20%	28.40%	-0.50%	-1.60% *	-1.10%
	(0.20)	(0.20)	(0.20)	(0.20)	(0.01)	(0.01)	(0.01)
Idiosyncratic shocks	13.50%	13.40%	13.30%	13.60%	0.10%	-0.20%	-0.30%
	(0.13)	(0.13)	(0.13)	(0.13)	(0.01)	(0.01)	(0.01)
Gini coefficient	0.3	0.29	0.3	0.29			
Observations	3,490	1,126	1,215	1,149			
PANEL B: CHILDREN INDICATORS							
Severe stunting	22.70%	23.20%	21.90%	23.20%	1.30%	-0.10%	-1.30%
	(0.36)	(0.37)	(0.35)	(0.37)	(0.02)	(0.02)	(0.02)
Severe wasting	4.60%	5.70%	3.60%	4.70%	2.10% **	0.90%	-1.20%
	(0.18)	(0.20)	(0.16)	(0.18)	(0.01)	(0.01)	(0.01)
Severe underweight	16.00%	15.60%	16.00%	16.40%	-0.40%	-0.80%	-0.40%
	(0.32)	(0.31)	(0.32)	(0.32)	(0.02)	(0.02)	(0.02)
Observations	1,970	645	705	620			

Data source: Baseline Survey. Consumption PC, AE and EOS correspond to consumption per capita, per adult equivalent, and per capita adjusted for economies of scales, respectively. See annex II for a detailed definition of variables. Regressions include commune fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 2: Average baseline characteristics of beneficiaries selected by each method

VARIABLES	(1) CBT	(2) FCS	(3) PMT	(4) CBT-FCS		(5) CBT-PMT		(6) FCS-PMT	
PANEL A: HOUSEHOLD INDICATORS									
Poor	0.85	0.88	0.93	-0.03		-0.08	***	-0.05	**
	(0.36)	(0.33)	(0.26)	(0.02)		(0.02)		(0.02)	
Consumption PC below eligibility cut-off	0.42	0.48	0.50	-0.07	**	-0.09	***	-0.02	
	(0.49)	(0.50)	(0.50)	(0.03)		(0.03)		(0.03)	
Consumption PC	118	114	103	4		15	***	11	**
	(65)	(65)	(56)	(4)		(4)		(4)	
Consumption AE	165	159	147	7		19	***	12	*
	(92)	(88)	(82)	(6)		(6)		(6)	
Consumption EOS	141	136	126	6		16	***	10	*
	(76)	(75)	(67)	(5)		(5)		(5)	
Food-insecure	0.14	0.11	0.15	0.03		-0.01		-0.04	**
	(0.35)	(0.32)	(0.36)	(0.02)		(0.02)		(0.02)	
Asset index	-0.15	-0.16	-0.11	0.00		-0.05		-0.05	
	(1.35)	(1.46)	(1.49)	(0.08)		(0.10)		(0.10)	
Livestock index	-0.02	-0.07	-0.08	0.05		0.06		0.01	
	(1.44)	(1.71)	(1.22)	(0.12)		(0.10)		(0.10)	
Income per AE	99	112	93	-13		5		19	
	(94)	(196)	(119)	(10)		(7)		(11)	
Food stock coverage	3.12	3.34	2.90	-0.23		0.22		0.44	*
	(2.10)	(2.51)	(2.03)	(0.17)		(0.14)		(0.15)	
Observations	518	483	487	1,001		1,005		970	
PANEL B: CHILDREN INDICATORS									
Severe stunting	0.22	0.24	0.26	-0.02		-0.04		-0.02	
	(0.35)	(0.37)	(0.37)	(0.03)		(0.03)		(0.03)	
Severe wasting	0.07	0.04	0.05	0.03	*	0.02		-0.01	
	(0.23)	(0.16)	(0.18)	(0.02)		(0.02)		(0.02)	
Severe underweight	0.15	0.18	0.20	-0.03		-0.05	*	-0.01	
	(0.31)	(0.35)	(0.35)	(0.03)		(0.03)		(0.03)	
Observations	305	275	287	580		592		562	
PANEL C: SUBJECTIVE WELL-BEING INDICATORS									
Ranked by others below eligibility cut-off	0.62	0.52	0.59	0.09	***	0.03		-0.06	**
	(0.49)	(0.50)	(0.49)	(0.03)		(0.03)		(0.03)	
Self-perception basic needs ranked below eligibility cut-off	0.64	0.57	0.59	0.07	**	0.04		-0.02	
	(0.48)	(0.50)	(0.49)	(0.03)		(0.04)		(0.04)	
Observations	518	483	487	1,001		1,005		970	

Data source: Baseline Survey (beneficiary sub-sample). Consumption PC, AE and EOS correspond to consumption per capita, per adult equivalent, and per capita adjusted for economies of scales, respectively. *Ranked by others below eligibility cut-off* corresponds to the share of households that were categorized below the program eligibility cut-off (43%) based on poverty rankings provided by other community members. *Self-assessed basic needs ranked below program eligibility cut-off* corresponds to the share of households that were categorized below the program eligibility cut-off (43%) based on self-assessed basic needs rankings. See annex II for a detailed definition of variables. Regressions include commune fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 3: Differences in characteristics between beneficiaries and non-beneficiaries selected by each method

VARIABLES	CBT				FCS				PMT			
	(1) NB	**(2)** B	**(3)** Diff. 1-2		**(4)** NB	**(5)** B	**(6)** Diff. 4-5		**(7)** NB	**(8)** B	**(9)** Diff. 7-8	
PANEL A: SURVEY-BASED HOUSEHOLD BENCHMARKS												
Poor	0.85	0.85	0.00		0.85	0.88	-0.03		0.80	0.93	-0.13	***
Consumption PC below eligibility cut-off	0.41	0.42	0.01		0.45	0.48	0.04		0.35	0.50	0.15	***
Consumption PC	123	118	4		122	114	8	*	133	103	30	***
Consumption AE	173	165	8	*	171	159	13	**	182	147	35	***
Consumption EOS	146	141	4		146	136	10	*	157	126	31	***
Food-insecure	0.11	0.14	-0.03		0.12	0.11	0.01		0.11	0.15	-0.04	
Asset index	0.05	-0.15	0.20	***	0.21	-0.16	0.37	***	0.04	-0.11	0.14	
Livestock index	0.15	-0.02	0.17	**	-0.02	-0.07	0.04		0.00	-0.08	0.08	
Income per AE	120	99	21	***	115	112	3		121	93	28	***
Food stock coverage	3.53	3.12	0.41	**	3.54	3.34	0.19		3.67	2.90	0.77	***
Observations	650	518	1,168		730	483	1,213		660	487	1,147	
PANEL B: SURVEY-BASED CHILDREN MEASURES												
Severe Stunting	0.24	0.22	0.03		0.21	0.24	-0.03		0.21	0.26	-0.05	
Severe Wasting	0.04	0.07	-0.03		0.03	0.04	0.00		0.05	0.05	0.00	
Severe Underweight	0.16	0.15	0.01		0.14	0.18	-0.04		0.13	0.20	-0.06	*
Observations	341	305	646		431	275	706		334	287	621	
PANEL C: COMMUNITY BASED MEASURES												
Ranked by others below eligibility cut-off	0.34	0.62	0.28	***	0.33	0.52	0.19	***	0.33	0.59	0.25	***
Self-assessed basic needs ranked below eligibility cut-off	0.38	0.64	0.26	***	0.35	0.57	0.22	***	0.38	0.59	0.22	***
Observations	650	518	1,168		730	483	1,213		660	487	1,147	

Data source: Baseline Survey. NB and B stand for beneficiaries and non-beneficiaries respectively. Consumption PC, AE and EOS correspond to consumption per capita, per adult equivalent, and per capita adjusted for economies of scales, respectively. *Ranked by others below eligibility cut-off* corresponds to the share of households that were categorized below the program eligibility cut-off (43%) based on poverty rankings provided by other community members. *Self-assessed basic needs ranked below program eligibility cut-off* corresponds to the share of households that were categorized below the program eligibility cut-off (43%) based on self-assessed basic needs rankings. See annex II for a detailed definition of variables. Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Village used as cluster. Regressions include commune fixed effects.

Table 4: Legitimacy of targeting methods

VARIABLES	(1) Satisfaction approach (1-4)	(2) Repeat same approach?	(3) Any non-poor included?	(4) Any poor excluded?
FCS	0.0278	0.0504**	0.0266	0.0427
	(0.0419)	(0.0230)	(0.0273)	(0.0270)
PMT	0.00974	0.0291	0.0689**	0.0215
	(0.0463)	(0.0230)	(0.0270)	(0.0274)
Observations	1,250	1,250	1,250	1,250
R-squared	0.041	0.029	0.073	0.111
FCS-PMT	0.0180	0.0213	-0.0423	0.0212
SE	(0.0450)	(0.0192)	(0.0284)	(0.0261)
CBT mean	3.273	0.793	0.267	0.786

Data source: Follow-up Survey. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Village used as cluster. All models include commune fixed effects.

Table 5: Legitimacy of targeting methods – beneficiaries only

VARIABLES	(1) Satisfaction approach (1-4)	(2) Repeat same approach?	(3) Any poor included?	(4) Any poor excluded?
FCS	-0.0794*	-0.0217	-0.00221	0.0427
	(0.0408)	(0.0197)	-0.036	-0.0377
PMT	-0.00949	-0.0176	0.0116	0.0413
	(0.0392)	(0.0205)	-0.0393	-0.0356
Observations	611	611	611	611
R-squared	0.073	0.051	0.088	0.121
FCS-PMT	-0.0699*	-0.00416	-0.0138	0.00134
SE	0.0388	0.0211	0.0397	0.0342
CBT mean	3.635	0.950	0.233	0.755

Data source: Follow-up Survey (beneficiary sub-sample). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Village used as cluster. All models include commune fixed effects.

Table 6: Legitimacy of targeting methods – non-beneficiaries only

VARIABLES	(1) Satisfaction approach (1-4)	(2) Repeat same approach?	(3) Any poor included?	(4) Any poor excluded?
FCS	0.113	0.113***	0.0558	0.0455
	(0.0719)	(0.0422)	-0.0393	(0.0293)
PMT	0.0272	0.0734*	0.124***	0.00359
	(0.0815)	(0.0427)	-0.0399	(0.0325)
Observations	639	639	639	639
R-squared	0.078	0.061	0.084	0.124
FCS-PMT	0.0854	0.0401	-0.0685*	0.0419
SE	0.0801	0.0342	0.0412	0.0313
CBT mean	2.933	0.646	0.298	0.816

Data source: Follow-up Survey (non-beneficiary sub-sample). Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Village used as cluster. All models include commune fixed effects.

Table 7: Effect of local knowledge on exclusion errors

VARIABLES	(1) Leader	(2) Women	(3) Mix	(4) Actual
Local knowledge	-0.380***	-0.363***	-0.319***	0.00658
	(0.109)	(0.0937)	(0.111)	(0.0948)
Observations	491	491	491	491
R-squared	0.111	0.116	0.101	0.099
Exclusion error mean	0.526	0.536	0.549	0.423

Data source: Baseline Survey (CBT sub-sample below eligibility cut-off of consumption distribution). Local knowledge is defined as the share of sampled households in a village knowing a given household. Robust standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Village used as cluster. Regressions include commune fixed effects.

Table 8: Likelihood of being a beneficiary based on CBT committee membership

VARIABLES	(1) Women	(2) Leader	(3) Mix	(4) Actual
Committee member	0.0280	0.140**	0.204***	0.0824*
	(0.0751)	(0.0678)	(0.0666)	(0.0423)
Constant	0.299***	0.244**	0.163*	0.255***
	(0.0984)	(0.0979)	(0.0959)	(0.0964)
Observations	998	999	998	1,006
R-squared	0.084	0.098	0.131	0.121

Data source: Baseline Survey (CBT sub-sample), and project monitoring data. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Village used as cluster. All models control for consumption per capita, food consumption score, assets, livestock, food stock, and subjective well-being indicators.

Table 9: Differences in overall errors with and without "fatigue effect"

VARIABLES	Obs	Mean	Std. Err.	[95% Conf.	Interval]
(1) IE actual	518	42%	0.02	0.38	0.47
(2) IE for "non-fatigued" group only	223	39%	0.03	0.33	0.45
(3) IE for "fatigued" group only	271	44%	0.03	0.38	0.49
Diff. (1) - (2)		-3.3%	0.04	-0.11	0.04
Diff. (2) - (3)		-4.5%	0.04	-0.13	0.04

Data source: Baseline Survey (CBT and beneficiary sub-sample), and project monitoring data. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Sample restricted to CBT villages.

Table 10: Effects of targeting choice and shocks on exclusion errors

VARIABLES	(1) Exclusion error
Number of shocks (standardized) X FCS	0.0147
	(0.0310)
Number of shocks (standardized) X PMT	0.0508*
	(0.0306)
Observations	1,488
R-squared	0.053
Mean CBT	-0.0190
FCS-PMT	-0.0361
	(0.0293)

Data source: Baseline Survey (sub-sample below eligibility cut-off of consumption distribution). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Village used as cluster. Regressions include commune fixed effects. Regressions include dummies for FCS and PMT as well as the shock indicator, defined as the standardized number of shocks affecting a household. Shocks considered included droughts, flood, pest, livestock, price, health, death. Excluded group is CBT.

Table 11: Inclusion errors from PMT formula, administrative versus baseline survey data

Inclusion errors	Obs	Mean	Std. Err.	[95% Conf.	Interval]
PMT score admin. survey	1,488	42%	0.01	0.45	0.40
PMT score baseline survey	1,487	36%	0.01	0.39	0.34
diff		6.1%***	0.02	0.03	0.10

Data source: Baseline (beneficiary sub-sample) and Registry Survey. Inclusion error based on consumption per capita.

Table 12: Inclusion errors from FCS formula, administrative versus baseline survey data

Overall errors	Obs	Mean	Std. Err.	[95% Conf.	Interval]
FCS score admin. survey	1,643	45%	0.01	0.42	0.47
FCS score baseline survey	1,625	39%	0.01	0.36	0.41
diff		6.0%***	0.02	0.03	0.09

Data source: Baseline (beneficiary sub-sample) and Registry Survey. Inclusion error based on the food consumption score.

Table 13: Inclusion error for actual versus recalibrated PMT model

Inclusion errors	Obs	Mean	Std. Err.	[95% Conf.	Interval]
PMT old	1,488	42%	0.01	0.40	0.45
PMT recalibrated	1,484	41%	0.01	0.39	0.44
		0.8%	0.02	-0.03	0.04

Data source: Baseline Survey (beneficiary sub-sample). Inclusion error based on consumption per capita.

Table 14: Inclusion error for actual versus recalibrated FCS model

	Obs	Mean	Std. Err.	[95% Conf.	Interval]
FCS old	1,643	45%	0.01	0.42	0.47
FCS recalibrated	1,641	43%	0.01	0.41	0.45
		1.7%	0.02	-0.02	0.05

Data source: Baseline Survey (beneficiary subsample). Inclusion error based on the food consumption score.

Table 15: Use of transfers

Expenditure item	CBT	FCS	PMT	CBT	FCS	PMT	CBT	FCS	PMT
	Spent on item?			**Amount spent from last payment**			**Share of total**		
Food	95%	93%	92%	10,985	11,779	10,444	68%	69%	65%
Durables	15%	14%	17%	959	1,098	1,145	6%	6%	7%
Transfers	39%	36%	42%	547	647	582	3%	4%	4%
Education	67%	71%	74%	472	414	459	3%	2%	3%
Health	23%	29%	29%	1,146	884	1,291	7%	5%	8%
Saving	25%	32%	28%	1,904	1,905	1,927	12%	11%	12%
Other	6%	12%	9%	184	563	337	1%	3%	2%

Data source: Follow-up Survey (beneficiary sub-sample). The monthly transfer provided amounted 10,000 CFA. In some cases, a household missed a payment and would subsequently receive several payments at the same time. This is why average amounts in the table can be greater than 10,000 CFA.

Table 16: Short term impacts of targeting choice – beneficiary sample

VARIABLES	(1) Food insecurity	(2) Food Stock	(3) Hunger	(4) Conflict
PMT X Round	4.651*	0.765**	-0.00853	-0.0152
	(2.734)	(0.360)	(0.0467)	(0.0834)
FCS X Round	2.870	-0.519	-0.0253	0.00455
	(2.725)	(0.330)	(0.0479)	(0.0797)
Round	1.345	0.750***	0.0210	0.374***
	(1.771)	(0.243)	(0.0319)	(0.0612)
FCS	-0.506	0.283	-0.0339	-0.0102
	(1.606)	(0.208)	(0.0250)	(0.0600)
PMT	-2.759*	-0.165	-0.0132	-0.0368
	(1.435)	(0.203)	(0.0238)	(0.0594)
Observations	1,743	1,713	1,714	1,738
R-squared	0.012	0.033	0.003	0.047
Diff. effects PMT & FCS	1.781	1.283***	0.0167	-0.0198
SE	2.938	0.346	0.0493	0.0763
CBT mean	43.94	3.397	0.861	3.615

Data source: Baseline and Follow-up Survey (beneficiary sub-sample). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Village used as cluster. Regressions include commune fixed effects.

Table 17: Definition of variables used in the analysis

Variable	Description
Poor	A household is considered poor if its aggregate annual consumption is below the national poverty line. We use the latest available 2014 national poverty line at CFA 189,233.2.
Consumption per capita (PC)	Consumption includes annual expenditure on food, non-food (incl. health and schooling) and durables items. Consumption per capita is the total consumption divided by the household size.
Consumption adjusted for Economies of Scale (ECOS)	Consumption includes annual expenditure on food, non-food (incl. health and schooling) and durables items. Consumption adjusted for economies of scale is the total consumption divided by the household size raised to the power of 0.9. The EOS adjustment takes account of the fact that larger households can typically purchase larger quantities of items, at a lower average unit price.
Consumption per adult equivalent (AE)	Consumption includes annual expenditure on food, non-food (incl. health and schooling) and durables items. Consumption adjusted for adult equivalence takes account of the fact that children or old persons do not consume as much food as adults. The standard FAO scale was used to make gender-age specific adjustments for adult equivalence.
Food-insecure	A household is considered food insecure if its Food Consumption Score is below 21.5 points.
Asset index	The asset index is created using principal component analysis (PCA) based on quantities owned of the following productive and non-productive assets: bicycle, iron, sewing machine, bed, table, attic, hilaire, mattress, simple mattress, motorbike, cell phone, radio cassette, cart, plow, seed drill. 0.156 of the overall variation is explained by the first component.
Livestock index	The livestock index is created using principal component analysis (PCA) based on the quantities of the following livestock items: veal, cow, cattle, chicken, sheep, goats, donkey. 0.2367 of the overall variation is explained by the first component.
Income per capita	Income is the total of all reported incomes from crops, self-employment, wage-employment, livestock and received transfers. Income per capita is the total income divided by the household size.
Food stock coverage	Duration of food coverage using own agricultural production (in months) in past year.
Severe Stunting	A child is severely stunted if its height-age z-score is 3 z-scores below the age-gender specific WHO world medians.
Severe Wasting	A child is severely wasted if its height-weight z-score is 3 z-scores below the age-gender specific WHO world medians.
Severe Underweight	A child is severely underweight if its weight-age z-score is 3 z-scores below the age-gender specific WHO world medians.

Food Consumption Score	Standard FAO Food Consumption Score which measures a household's dietary diversity (Kennedy et al. 2015).
Share of dependents	Number of dependents (aged 14 or younger or 60 or older) divided by the household size.
Handicap	Any individual affected by a handicap, i.e. is blind, deaf, mute, mentally challenged or has leprosy, or other.
Education	An individual that has *any* level of education
Land ownership	A household that owns any cultivated land
Plots cultivated	Total number of plots cultivate by household
Covariate shocks	Over the past 12 months household was affected by any covariate shocks (including drought, flooding, price increase for food)
Idiosyncratic shocks	Over the past 12 months household was affected by any idiosyncratic shock (including death of a household member, divorce/separation, theft of money, goods or harvest, crop disease, livestock disease, price increase in agricultural inputs, price drop in agricultural outputs).
Hunger	This variable is based on a question that asks, whether in the past 12 months, the household has experienced a situation where he didn't have enough food to feed the entire household.
Conflict	This variable is based on a question that asks, if in the past 12 months, there has been tension or conflict between members of the village. Response options are always, sometimes, rarely, never, which are coded as 1, 2, 3, and 4 respectively.