The World Bank Economic Review, 0(0), 2020, 1–29
doi: 10.1093/wber/lhaa019
Article

Efficiency, Legitimacy, and Impacts of Targeting Methods: Evidence from an Experiment in Niger

Patrick Premand and Pascale Schnitzer

Abstract

The methods to select safety net beneficiaries are the subject of frequent debates. Targeting assessments usually focus on efficiency by documenting the pre-program profile of selected beneficiaries. This study provides a more comprehensive analysis of targeting performance through an experiment embedded in a national cash transfer program in Niger. Eligible villages were randomly assigned to have beneficiary households selected by community-based targeting (CBT), proxy-means testing (PMT), or a formula to identify the food-insecure (FCS). The study considers targeting legitimacy and the impact of targeting choice on program effectiveness based on data collected after program roll-out. PMT is more efficient in identifying households with lower consumption per capita. Nonbeneficiaries find formula-based methods (PMT and FCS) more legitimate than CBT. Manipulation and information imperfections affect CBT, which can explain why it is not the most legitimate. Program impacts on some welfare dimensions are larger among households selected by PMT than CBT.

JEL classification: H53, I32, I38, O12, O15, O18, R28

Keywords: targeting, safety nets, cash transfers, poverty, field experiment

Patrick Premand (corresponding author) is a senior economist in the Development Impact Evaluation Department (DIME) at the World Bank, Washington, DC, USA; his email address is ppremand@worldbank.org. Pascale Schnitzer is an economist in the Social Protection & Jobs Global Practice at the World Bank, Washington, DC, USA; her email address is pschnitzer@worldbank.org. The experiment for this study was undertaken in partnership with the Cellule Filets Sociaux (CFS) in the prime minister's office of the Government of Niger. CFS implemented the targeting protocols and the cash transfer program. The authors are thankful to Ali Mory Maidoka, Bassirou Karimou, and Kadi Aboubacar at CFS, and to Carlo del Ninno and Mahamane Maliki Amadou at the World Bank for countless contributions. Sahel Consulting, led by Adamou Hamadou, collected the study data. Julius Gunnemann, Ibrahim Ouattara, Karim Pare, Teslem Zeln, and Al-Mouksit Akim provided excellent support to data collection, management, and analysis. The authors are grateful for inputs and comments from Arthur Alik-Lagrange, Robin Audy, Phillippe Leite, Brad Mills, Jamele Rigolini, Marc Smitz, Quentin Stoeffler, Dominique Van de Walle, Kelsey Wright, Ruslan Yemtsov, the editor, and anonymous referees, as well as seminar participants at ETH Zürich and the World Bank. The computational reproducibility of the study results was verified by DIME analytics. The research for this article was supported by the World Bank Sahel Adaptive Social Protection Program, funded by the United Kingdom; and the Niger Adaptive Safety Nets Project funded by the World Bank and managed by CFS. Both authors worked for the World Bank Social Protection & Jobs Global Practice at the time the study was designed and implemented. Premand made additional contributions while a guest researcher at ETH Zürich and at DIME. The findings, interpretations, and conclusions of the paper are those of the authors and do not necessarily represent the views of the World Bank or the government of Niger. A supplementary online appendix for this article can be found at *The World Bank Economic Review* website.

1. Introduction

The targeting of social programs is a topic of active policy debates and research. Recent work has discussed when targeted programs or universal approaches may be preferable (Brown, Ravallion, and van de Walle 2018; Hanna and Olken 2018; Banerjee, Niehaus, and Suri 2019; Gentilini et al. 2020).[1] The performance of methods to select beneficiaries is a key consideration. Although achieving full coverage of eligible populations is a widespread aspiration, many countries have limited budgets for social spending. Even after applying geographical targeting criteria, methods to select beneficiary households are commonly required.

Analyses of targeting performance often focus on efficiency, that is, whether the pre-program profile of the selected population matches intended target groups (e.g., Coady and Skoufias 2004; Coady, Grosh, and Hoddinott 2004a; Coady, Grosh, and Hoddinott 2004b; Del Ninno and Mills 2015; Klasen and Lange 2016; McBride and Nichols 2016; Stoeffler, Mills, and del Ninno 2016; Bah et al. 2018; Brown, Ravallion, and van de Walle 2018). For instance, "inclusion errors" quantify the likelihood that selected households have welfare above a given threshold before program roll-out.

Few studies directly compare alternative targeting methods based on their implementation in real-life settings. Experiments testing different targeting methods are particularly rare: Alatas et al. (2012) compare proxy-means testing (PMT) and community-based targeting (CBT), while Alatas et al. (2016b) compare PMT and self-selection in Indonesia. Some comparative studies are based on the implementation of one targeting method and the simulation of another, such as PMT and CBT in Honduras and Peru (Karlan and Thuysbaert 2019), or PMT and selection by community leaders in Malawi (Basurto, Dupas, and Robinson 2019).

This study presents the results of a large-scale targeting experiment embedded in a national cash transfer program. It provides a comprehensive assessment of targeting performance by analyzing targeting efficiency, targeting legitimacy and the impact of targeting choice on program effectiveness.

The experiment compares two of the most popular methods, PMT and CBT, as well as another method to identify temporarily food-insecure households (FCS). Specifically, 318 villages eligible for a cash transfer program in Niger were randomly assigned to have beneficiary households selected by CBT (104 villages), PMT (105 villages), or FCS (109 villages).

For each method, the study measures targeting efficiency by comparing pre-program welfare indicators between selected and nonselected households. It then assesses legitimacy and estimates how program welfare impacts vary across targeting methods using data collected after program roll-out.[2] It also considers a range of mechanisms that can explain differences in the relative efficiency and legitimacy of targeting methods, in particular how imperfect information, manipulation and local inequality affect CBT.

PMT, FCS, and CBT can theoretically lead to the selection of beneficiaries with different profiles. When assessing targeting efficiency, a relevant welfare benchmark needs to be identified. Through CBT, local populations select households based on their own conception of poverty, which may not match standard welfare benchmarks such as consumption or food security. PMT seeks to target households with low average consumption per capita, which captures a relatively stable measure of welfare. In contrast, FCS aims to target temporarily food-insecure households by proxying the food consumption score[3] measured during the lean season. This approach is akin to the Household Economy Analysis

1 Earlier discussions can be found in Ravallion and Datt (1995).

2 This study is most closely related to Alatas et al. (2012), who compare the efficiency of CBT relative to PMT in an experiment in Indonesia. They consider differences in satisfaction after program roll-out, pointing to higher legitimacy of CBT relative to PMT. The present study sheds light on potential trade-offs across the three dimensions of targeting performance: efficiency, legitimacy, and impacts on program effectiveness.

3 The food consumption score is the food security indicator most widely used by humanitarian agencies. It captures dietary diversity through the quantity and quality consumed (World Food Programme 2008; Wiesmann et al. 2009).

(HEA), which is widely used by humanitarian actors.[4] Food security fluctuates and is more volatile than consumption. As such, the FCS method attempts to account for transient factors related to shocks or seasonality, which PMT may not capture (Del Ninno and Mills 2015). Differences in the profile of households selected by each method can affect targeting efficiency, targeting legitimacy, and program effectiveness.

Considering the legitimacy of targeting methods is relevant for several reasons. Measures of legitimacy obtained after program roll-out can complement subjective well-being indicators to assess which targeting methods best match definitions of poverty held by local populations. Yet targeting legitimacy may depend on the targeting process, not solely on the targeting results. Populations and policy makers may particularly value selection processes that are transparent, participatory, and free of manipulation. CBT has the distinct feature of relying on a participatory process. Formula-based targeting methods, such as PMT and FCS are sometimes considered as lacking transparency (Hanna and Olken 2018). Ultimately, targeting legitimacy (or illegitimacy) can affect social cohesion at the local level and influence broader political support for government programs.

The choice of targeting methods may also determine program effectiveness, that is, the magnitude of program impacts on outcomes related to program objectives. Targeting is only a means to effectively achieve program objectives, such as reducing poverty or improving food security. Ravallion (2009) highlights the limits of targeting efficiency measures, which have little correlation with the poverty impacts of a large social program in China. Basurto, Dupas, and Robinson (2019) and Banerjee, Niehaus, and Suri (2019) even suggest trade-offs between targeting efficiency and program effectiveness. Basurto, Dupas, and Robinson (2019) find that local leaders are less efficient at allocating subsidies to the poor but are able to identify individuals for whom returns are higher. The present study tests whether the impacts of a cash transfer program vary across villages assigned to different targeting methods.[5] It uses a difference-in-differences strategy that compares changes in baseline and post-program outcomes related to (chronic) food security, consumption, and poverty between beneficiaries selected by each method.

Results show that PMT is more efficient than other methods in identifying households with higher poverty rates (by 5–8 percentages points) and lower consumption per capita (by 10–13 percent). The targeting methods have similar efficiency with respect to other welfare measures. CBT and PMT are closely related to subjective well-being indicators that proxy perceptions of poverty, but FCS less so. Findings on legitimacy show that nonbeneficiaries consider the use of formula-based methods (PMT and FCS) more legitimate. Nonbeneficiaries are 9–14 percentage points more likely to want the PMT and FCS approaches to be repeated in the future, relative to CBT. An analysis of mechanisms explains why CBT is not necessarily considered the most legitimate by local populations. Some committee members attempt to manipulate the CBT process, and individual committees generate substantial exclusion errors due to imperfect local knowledge. Households are aware of manipulation risks, with almost half of respondents reporting that committee members try to benefit themselves. Finally, program impacts tend to be larger among households selected by PMT than CBT. For instance, impacts on food stock are larger by 20 percent and impacts on poverty reduction larger by 9 percent for cash transfer beneficiaries in PMT villages compared to CBT villages.

The findings highlight challenges with targeting in low-income settings where safety net coverage is constrained by limited funding. Variations in performance across targeting methods are not large in the study setting, but the results point to some gains when using PMT. Trade-offs between targeting efficiency, legitimacy, and program effectiveness are not found. The results call for particular caution when

4 For more information on HEA and its use in the Sahel region, see Schnitzer (2019).

5 Studies on the effect of targeting choice on poverty usually rely on simulations. For example, Brown, Ravallion, and van de Walle (2018) and Alatas et al. (2012) find limited variations in impact of cash transfer programs on poverty reduction across targeting methods.

considering CBT due to information imperfection and risks of manipulation. One caveat of the study is that baseline measures of food security were collected slightly after the end of the lean season. This limits the assessment of the performance of FCS, including its efficiency in reaching temporarily food-insecure households. As such, conclusions on the FCS approach are less clear-cut.

2. Study Design and Implementation

Context and Niger Safety Nets Program

The challenge of selecting beneficiaries for social programs is particularly salient in low-income settings, including in the Sahel. Niger is among the poorest countries in the world, with a poverty incidence of 51.4 percent in rural areas (World Bank 2017) and widespread food insecurity (Schnitzer 2019). Social protection programs in Niger are thin and fragmented. A large share of resources for social programs is allocated to emergency responses implemented with relatively weak government coordination. The total amount of funding for social programs can cover only a small share of the population in need of support. There are recurring questions on how best to select safety net beneficiaries. In this context, a better understanding of the performance of alternative targeting methods can facilitate coordination as part of a national social protection system.

In 2011, the Government of Niger, with support from the World Bank, set up a national safety nets project. It included a cash transfer program providing small, regular transfers of 10,000 FCFA per month (or approximately $17.2),[6] for 24 months, to poor households. The duration of the program contrasts with temporary humanitarian interventions providing cash transfers for three to four months during the lean season. By offering small regular cash transfers over a longer period of time, the program seeks to improve food security in the short term, while providing consumption support, reducing poverty, and facilitating investments in human capital and income-generating activities over the medium term. The cash transfer program was designed based on a pilot initiated in 2010 (Stoeffler, Mills, and Premand 2020), and has expanded over time to reach 100,000 households by 2018.

The present study is a large-scale randomized experiment that tests alternative targeting methods. The experiment was embedded in the second phase of the cash transfer program implemented from 2016 to 2018 through a collaboration with the government of Niger. This phase took place in five regions that host 95 percent of the country's poor population (World Bank 2011).[7] Geographical targeting identified 38 rural communes with the largest poor population in these five regions.[8] Within these communes, the number of villages to be served by the program was determined to cover on average 40 percent of households in participating villages. All villages were considered eligible, and, to ensure transparency, beneficiary villages were selected through public lotteries (Premand and Barry 2020).[9] The geographical targeting process to select communes and villages was applied everywhere and is not the focus of this paper.

6 The conversion is based on an exchange rate of 582 FCFA/U.S dollar.

7 The five regions are: Dosso, Maradi, Tahoua, Tillabéri, and Zinder.

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.

9 The public lotteries were organized by program staff in communal headquarters. Local authorities and representatives from all villages in the commune were invited to participate. In order to give the same chance to each village to be selected, all the village names were placed in a bucket, and a random draw took place to select villages up to a pre-established quota in each commune. Data from control villages were not collected as part of this study. Premand and Barry (2020) present results on the effectiveness of an earlier phase of the cash transfer program, including a control group. They also provide additional details on the lotteries.

Targeting Experiment Design and Implementation

The targeting experiment was implemented in 18 communes,[10] where 356 villages were selected to participate in the program; 318 villages were drawn to be part of the targeting experiment.[11] In all villages, a registry census was collected door to door among all households. Private firms were hired to collect the census based on a brief two-page questionnaire. The questionnaire contained information to compute PMT and FCS scores for each household, as further described below.

The registry census was used to determine a quota of beneficiary households per village. This was done based on (1) a quota of beneficiaries per commune depending on the estimated poor population in the commune, and (2) the application of a PMT formula to the registry census to allocate beneficiaries in the commune between villages.[12] As a result of this process, program implementers knew how many beneficiaries needed to be selected in each village.

The 318 study villages were then randomly assigned to have beneficiary households selected through one of three methods: community-based targeting (CBT, 104 villages), a proxy-means test formula designed to identify the poor (PMT, 105 villages), or an alternative formula designed to identify temporary food-insecure households (FCS, 109 villages). The randomization was performed in Stata and stratified by commune. To ensure balance, within each commune, randomization was further stratified using a PCA index based on village population, village average PMT score, and Gini coefficient.[13] These stratification variables were obtained from the registry census.[14] As a result of this process, program implementers knew which method had to be used to select beneficiaries in each village. Each village was informed one by one about the targeting steps that would be used in their specific case. This was done when it was time for the next step in the targeting protocol to be implemented.[15]

Implementation of Community-Based Targeting (CBT)

The CBT method relies on village members to select beneficiaries (Coady, Grosh, and Hoddinott 2004a; McCord 2013). In the context of the study, it was implemented through a participatory wealth ranking process. Villagers were asked to rank households according to their poverty level based on well-defined protocols. In each village randomized to have beneficiary households selected by CBT, a village-wide assembly was organized by trained facilitators to provide information about the process. To ensure that information was consistently explained throughout, scripts with key messages were developed in local languages. They aimed to (1) recall the objectives and key features of the cash transfer program; (2) explain the objectives and process to identify program beneficiaries; and (3) guide the community in

10 These communes were chosen by the implementing agency to reflect the diversity of the five targeted regions. The study is representative of the selected communes, which should be considered among the poorest in Niger, although they are not formally representative of all project areas.

11 Power calculations showed that 318 villages were needed in the study sample. These 318 villages were randomly drawn from the 356 beneficiary villages. The number of sample villages per commune was determined proportionally to the number of villages in each commune. Among the 318 sampled villages, 5 had to be replaced. Two villages were nomad villages where everyone had migrated at the time of data collection. In the other three villages, the registry census had issues that required systematic checks by program implementers, which could not be done in time for the survey. Replacement villages were randomly drawn.

12 A minimum share of beneficiaries was set at 10 percent and a maximum share at 70 percent per village. These were rarely binding in practice.

13 Thirty-six strata were created. Within each commune, two cells were created based on the top and bottom half of the PCA distribution. The median number of villages per stratum was 7, ranging from 3 to 22.

14 The Gini coefficient was computed based on the distribution of the PMT score, using all households in each village covered by the registry census.

15 Specifically, until the registry census and baseline survey, the steps were identical in all villages. The only visible difference was that a team visited villages assigned to CBT, while such a visit did not take place in PMT and FCS villages (where the next step was the registration phase, as detailed below). This reduces the risk of contamination across villages.

selecting three committees to perform CBT. Protocols were put in place to maximize participation during the assembly. In the assembly, and with guidance from the facilitator, all villagers were asked to jointly form three committees of five members each. Committee members were chosen on the basis that they were well known and trusted. The three committees include a group of leaders (with at least two women), a group of nonleader women, and a mixed group of nonleaders (with at least two women).[16]

Once the three committees were formed, each committee conducted its ranking independently with the support of a trained facilitator.[17] 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.[18] Next, the committees were given a set of pre-printed cards. Each card included the name of the household head (and adult household members) of one household visited during the registry census.[19] The order of the cards was randomized: each card included a random number, and the cards were sorted from the smallest to the largest number. The name on each card was read aloud one by one in the pre-defined random sequence. After each card was read, committee members were asked to place it on a long board containing three different colors (red, yellow, and green), capturing the perceived capacity of each household to satisfy needs during the entire year. When disagreements were observed, the facilitator would recall the characteristics of "poor households" defined during the discussion to help members find a consensus.

Once each committee completed its ranking, the information was compiled by the national safety nets unit. An average ranking was calculated for each household.[20] Based on this process, within each village, households were ranked according to their average score. This ranking led to the identification of a list of beneficiary households to be registered in the program, that is, households with average ranks lower or equal to the number of beneficiaries to be selected based on the quota for each village.

Implementation of Targeting by Proxy Means Testing (PMT)

The Proxy Means Testing 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 household demographics, dwelling characteristics, land ownership, and assets.[21] The PMT formula was developed based on a national household survey (LSMS) collected in 2011.[22] The survey included two rounds of data collection, the first during the lean season, and the second post-harvest. The

16 Committee members were selected by village members during the assembly. The protocols were detailed and closely monitored. The criteria to be a committee member were clearly stated to all village members. In addition to the gender and leadership criteria specific to each committee, and to the extent possible, diversity was encouraged within each group along dimensions such as household location, education levels, wealth, or ethnicity.

17 Each field team had three facilitators (one to work with each committee), as well as a supervisor responsible for checking that protocols were enforced.

18 CBT aimed to identify households based on the communities' own perceptions of poverty. As such, the meaning of "basic needs" was defined by communities as part of the process.

19 In practice, based on the door-to-door registry census data, around 7 percent of households in the top part of the PMT distribution were filtered out from the ranking process. This step aimed to reduce the time needed to rank households, while at the same time improving targeting efficiency and minimizing manipulation. PMT can be effective at excluding households in the top part of the consumption distribution (Schnitzer 2019). See section S1 in the supplementary online appendix, which outlines that this step does not affect the results on targeting performance.

20 If the rank given by a committee for any given household deviated by more than two standard deviations from the ranks from the other two committees, it was not used.

21 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 nonfood expenditures (such as clothes or household durables).

22 Section S1 of the supplementary online appendix shows that the time lag between the national survey (used to estimate the targeting formula) and the registry census (the data used to apply the formula) does not lead to large errors in the context of the study.

formula was derived from a stepwise regression starting from a broad set of variables including household demographics, dwelling characteristics, land ownership, livestock, and assets. It identified the variables that best predict average consumption[23] between the two rounds.[24] The PMT formula was then applied to the registry census data to calculate a PMT score for each household. Within each village, households were ranked according to their PMT score, which together with the village quotas determined a list of households to be registered in the program.

Implementation of Targeting by Food Insecurity Proxy Formula (FCS)

Many humanitarian actors use a method called Household Economy Analysis (HEA), which aims to identify households that are temporarily food insecure. Schnitzer (2019) shows that the formula on which the HEA method is based can be further improved to identify households at risk of temporary food insecurity.[25] This formula proxies the food consumption score indicator during the lean season and constitutes the third targeting method tested in the experiment (FCS). The food consumption score captures dietary diversity based on a set of questions asking about the consumption frequency of different food groups during the past seven days (World Food Programme 2008).[26] This indicator is widely used by agencies responding to food crises and aims to reflect both the quantity and quality of food consumed.

The FCS proxy formula aims to predict temporary food security in the lean season based on a limited number of easily verifiable variables. As for the PMT method, the formula was obtained from the nationally representative 2011 LSMS survey. The formula was derived from a step-wise regression model starting from a broad set of variables including household demographics, dwelling characteristics, land ownership, and assets. The formula contains the variables that best predict households at risk of temporary food insecurity based on the food consumption score measured in the lean season. The formula was then applied to data from the registry census to obtain a proxy FCS score for each household.[27] Together with the village quotas, the ranking of households based on this score produced a list of households to be registered in the program.

Registration of Selected Beneficiaries

After the safety nets unit produced a list of beneficiaries to enroll in the program based on the targeting method assigned to each village, implementation teams went back to each village to communicate the results and register beneficiaries. The registration phase started with a village assembly, when the list of selected households was read to the community. During this assembly, community members could publicly

23 Total consumption was considered. On average, food consumption represents 75 percent of total consumption.

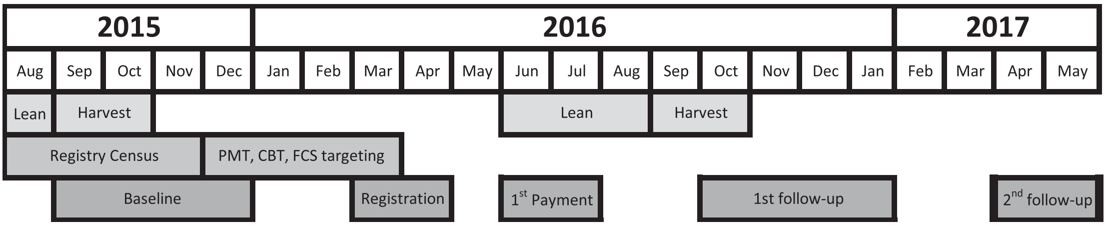
24 Section S2 in the supplementary online appendix provides more details on the construction of the formulas and the variables they include.

25 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 (2019) highlights the relatively weak correlation between consumption and temporary food insecurity during the lean season.

26 The food consumption score is the primary dietary diversity measure used by humanitarian actors working on seasonal insecurity in the Sahel (see broader discussion in Schnitzer 2019). Validation studies have shown that the food consumption score captures household dietary diversity and correlates with alternative measures such as the household dietary diversity score (Wiesmann et al. 2009). Several new approaches have recently been introduced, but there is no consensus in the literature on which measure is preferable (Cafiero et al. 2014; Bertelli 2020). A coping strategy index is also considered at times, but it captures responses conditional on the occurrence of shocks, which makes it harder to interpret.

27 As for CBT, around 7 percent of households in the top of the PMT distribution were filtered out from the selection process based on the registry data. This step does not affect targeting performance (see section S1 in the supplementary online appendix).

Figure 1. Study Timeline



Source: Authors' presentation of study timeline. Rows contain: year, month, season, targeting steps, data collection and program implementation steps.

object to the inclusion of households, though in practice very few did.[28] Based on the list of beneficiaries validated by the community, selected households were registered.

Timeline and Data

The study took place between August 2015 and May 2017. Figure 1 illustrates the timeline. The first source of data for the experiment comes from the registry census collected in all households in villages that were selected to participate in the cash transfer program. It included data on a limited set of variables needed to apply both the PMT and FCS formulas. In total, 54,051 households were surveyed between August and November 2015. The registry census was collected between the last month of the lean season and the first month of the post-harvest season.

A baseline survey was collected shortly after the registry census and before any beneficiary selection process was implemented. It included a sample of 12 households per village randomly drawn from the registry census: 3,816 households were sampled and 3,496 interviewed.[29] The baseline survey contains detailed information on several welfare dimensions (discussed further below in section 3), including consumption, food insecurity, anthropometrics (for children under the age of five), assets, household demographics, education and health status of household members. Note that the baseline survey took place between September and December 2015. The survey was collected during the harvest and post-harvest season, which is a limitation of the study since temporary food insecurity during the lean season is not observed. The registry census and baseline surveys were collected before the randomization of targeting methods, which prevents differential response bias across villages assigned to each method.

Following the baseline survey, targeting protocols were implemented. The CBT process was implemented between December 2015 and January 2016. The PMT and FCS formulas were applied in parallel. The communication of results took place during registration assemblies (March to April 2016). Out of the 3,496 baseline households, 1,439 were selected to receive cash transfers for 24 months. The cash transfer payments started in June and July 2016.

Three to five months into the cash transfer program (between October 2016 and January 2017), a first follow-up survey was collected to analyze the legitimacy of the various targeting methods, as well as to assess short-term impacts of the choice of targeting methods on a limited set of outcomes related to food security. The survey took place in a random subsample including 162 villages, or approximately half the

28 Complaint committees were set up prior to the start of the targeting process, and these committees were present and could also voice concern at the registration stage. At that stage, 99 percent of households identified by each of method were "validated" by the community. This does not mean that communities are in full agreement with the results of the targeting process and that the targeting methods are necessarily effective, as the analysis below will make clear. Anecdotal evidence from the field suggests that during the validation assembly, community members tend to refrain from disagreeing in public with the presented lists of beneficiaries.

29 Among the 8 percent of households not interviewed, 70 percent were households not recognized. This was due to errors in the registry census, including errors in village identifiers. The remaining were absent.

villages in the baseline sample.[30] Eight households were randomly drawn from the baseline sample in each village, including four beneficiary and four nonbeneficiary households.[31] The first follow-up survey sample included 1,296 households: 96.5 percent were successfully interviewed, including 611 beneficiary households and 639 nonbeneficiary households.

Lastly, a second follow-up survey was collected approximately one year after the cash transfers started (in April–May 2017). While the second follow-up survey was not collected specifically for this study,[32] it contains additional information on household welfare that is helpful to assess whether program impacts on consumption per capita and poverty vary across targeting methods. The sample for the second follow-up survey included 1,280 beneficiary households randomly drawn from the baseline.

Administrative data on costs, and household-level information on registration and payment status, as well as monitoring data from CBT implementation complement the survey data.

Baseline Balance

To document balance in baseline household characteristics ($Y_{0,i}$) across villages, the following equation is estimated by OLS using CBT as the reference category:

$$Y_{0,i} = \alpha + \beta_1 PMT_i + \beta_2 FCS_i + \pi_S + \varepsilon_i \tag{1}$$

CBT, PMT, and FCS are dummy variables taking the value of 1 for households in villages assigned to each targeting method. π_S capture dummies for the randomization strata. Robust (White-Huber) standard errors are clustered at the village level. The reported results show the estimated difference in characteristics across households in villages assigned to each targeting method (β_1, β_2, $\beta_1 - \beta_2$) and related p-values.

Results for the full baseline sample are presented in table S3.1 in the supplementary online appendix, available with this article at *The World Bank Economic Review* website. Beneficiary households represent 40 percent of the sample. The balance tests show that the randomization worked as intended. There is only a small number of statistically significant differences across villages assigned to different targeting methods.[33] The baseline data also highlight that the sample is overwhelmingly poor, with 85 percent of households below the national poverty line. Food insecurity affected 12 percent of households during the harvest or post-harvest season. Households reported that their own agricultural production covered food needs for an average of 3.4 months during the year. While the overall level of inequality is low (with an average Gini of 0.24), there is variation across villages, with village-level Gini coefficients ranging from 0.1 to 0.4.[34]

30 The sample covered all 18 communes in the 5 regions. The sample size was determined based on power calculations to detect differences in legitimacy outcomes.

31 Due to sampling variation, the number of beneficiary or nonbeneficiary households sampled at baseline was less than 4 in some villages. In these cases, additional nonbaseline households were added to the sample; 86 percent of sample households were from baseline. Estimations based on the follow-up survey use survey weights, as discussed below.

32 The second follow-up survey was the baseline for complementary interventions that were later introduced during the final year of implementation of the cash transfer program (Bossuroy et al. 2019). For this reason, the survey focused on cash transfer beneficiaries only. In addition, the budget for the complementary interventions only allowed covering 17 of the 18 communes, so that one commune from the targeting study was not included in the sample. All program villages in the 17 communes were included in the sample. The second follow-up sample also included cash transfer beneficiary households not sampled at baseline, but these are not considered here.

33 The main difference is about subjective well-being indicators between CBT and FCS villages.

34 The low levels of human development in the sample are also noteworthy: only 20 percent of household heads have ever been to school, usually only for a few years at most. Households are large with a mean size of seven. Twenty-two percent of all households are polygamous. The population is young, with 44 percent of individuals less than 12 years old.

3. Efficiency of Targeting Methods

This section documents the relative efficiency of the three targeting methods. The objective is to understand if there are differences in the pre-program profile of beneficiaries selected by each method, and if so, along which welfare dimension. The analysis of efficiency can also provide information about the implicit welfare benchmarks used by local populations when they select beneficiaries through CBT, and whether they differ from other targeting methods.

First, the study compares the mean characteristics of beneficiary households selected by each of the three methods. This involves measuring differences in baseline welfare ($Y_{0,i}$) among program beneficiaries ($B = 1$) selected by each method. The following equation is estimated by OLS, using CBT villages as the reference category:

$$\left(Y_{0,i} \mid B = 1\right) = \alpha + \beta_1 PMT_i + \beta_2 FCS_i + \pi_S + \varepsilon_i \tag{2}$$

π_S are dummies for the randomization strata. Robust (White-Huber) standard errors are clustered at the village level. Table 1 presents the results. For each welfare indicator, columns 1–3 present the mean characteristics for households selected by CBT, PMT, respectively FCS. Columns 4–6 show the estimated differences in the characteristics of beneficiaries selected between methods (β_1, β_2, $\beta_1 - \beta_2$) and their statistical significance.[35]

Equation (2) is estimated for a range of welfare measures collected at baseline, including household-level indicators (panel A), children indicators (panel B), and subjective well-being indicators (panel C). The annex contains a detailed definition of each measure. Household-level welfare measures include consumption per capita, consumption per adult equivalent and consumption with adjustment for economies of scale, as well as a poverty rate calculated based on consumption per capita. Consumption includes annualized expenditures and self-consumption for food, nonfood (including health and schooling) and durables.[36] Other household-level welfare benchmarks include the food consumption score, asset and livestock indices,[37] income (monetary earnings) per capita and food stock coverage.[38] Children indicators in panel B include the incidence of severe stunting, wasting, and underweight for children under the age of five. Subjective well-being indicators include dummies capturing if a household (1) is perceived as very poor by other community members[39], and (2) has a low self-reported ability to satisfy basic needs.[40]

Results show that PMT is more efficient than FCS and CBT in identifying households with higher poverty rates and lower consumption per capita.[41] Households selected by PMT have a poverty rate

35 Note that the correlation between the PMT and FCS scores is 0.2. In CBT villages, the correlation between the PMT score and the CBT ranking is 0.12, and the correlation between the FCS score and CBT ranking is 0.13.

36 Food consumption is captured for the 30 days before the survey and multiplied by 12. Note that this is slightly different from what is done in the national household survey, which collects food consumption at two points in time (during the lean season and post-harvest) and averages it out before annualizing. Still, the correlation of food consumption between the two waves is quite high (0.7, see Schnitzer 2019).

37 The asset and livestock indices are constructed using Principal Component Analysis (PCA).

38 Food stock constitutes a measure of food security as it captures the duration of food coverage using own agricultural production over the past year.

39 Baseline households were asked to rank other sample households by poverty levels. For each household, a dummy variable is constructed and equals 1 if the average of rankings provided by other baseline respondents falls below the program cut-off line.

40 This measure is based on the average response across six questions that ask respondents how satisfied they feel (on a scale from 1 to 3), with respect to six basic needs, including food, housing conditions, health, education, clothing, and revenues. The dummy variable equals one if the ranking based on self-assessed basic needs falls below the program cut-off line.

41 The correlation between the PMT score and consumption per capita is 0.235. The correlation between the FCS score and consumption per capita is 0.022. In CBT villages, the correlation between the CBT ranking and consumption per capita is 0.

Table 1. Average Baseline Characteristics of Beneficiaries Selected by Each Method

	(1) CBT	(2) PMT	(3) FCS	(4) PMT-CBT	(5) FCS-CBT	(6) PMT-FCS
Panel A: Household indicators						
Poor	0.85	0.93	0.88	0.08***	0.03	0.05**
	(0.36)	(0.26)	(0.33)	(0.02)	(0.02)	(0.02)
Consumption PC (thousands FCFA)	119	104	114	−15***	−4	−11**
	(65)	(56)	(65)	(4)	(4)	(4)
Consumption AE (thousands FCFA)	166	147	159	−19***	−7	−12**
	(92)	(82)	(88)	(6)	(6)	(6)
Consumption EOS (thousands FCFA)	142	126	136	−16***	−6	−10**
	(76)	(67)	(75)	(5)	(5)	(5)
Food insecure (chronic)	0.14	0.15	0.11	0.01	−0.03	0.04*
	(0.35)	(0.36)	(0.32)	(0.02)	(0.02)	(0.02)
Asset index	−0.15	−0.11	−0.16	0.05	0	0.05
	(1.35)	(1.49)	(1.46)	(0.10)	(0.08)	(0.10)
Livestock index	−0.02	−0.08	−0.07	−0.06	−0.05	−0.01
	(1.44)	(1.22)	(1.71)	(0.10)	(0.12)	(0.10)
Income AE (thousands FCFA)	71	66	80	−4	9	−13
	(68)	(82)	(136)	(5)	(7)	(8)
Food stock coverage	3.12	2.9	3.34	−0.22	0.23	−0.44*
	(2.10)	(2.03)	(2.51)	(0.14)	(0.17)	(0.15)
Observations	518	487	483			
Panel B: Children indicators						
Severe stunting	0.21	0.27	0.23	0.05*	0.02	0.04
	(0.35)	(0.37)	(0.36)	(0.03)	(0.03)	(0.03)
Severe wasting	0.07	0.04	0.03	−0.03	−0.04**	0.01
	(0.23)	(0.17)	(0.14)	(0.02)	(0.02)	(0.01)
Severe underweight	0.16	0.2	0.18	0.05*	0.02	0.02
	(0.31)	(0.34)	(0.34)	(0.02)	(0.03)	(0.03)
Observations	326	310	288			
Panel C: Subjective well-being indicators						
Perceived very poor by others	0.62	0.59	0.52	−0.03	−0.09***	0.06**
	(0.49)	(0.49)	(0.50)	(0.03)	(0.03)	(0.03)
Low ability to satisfy basic needs	0.64	0.59	0.57	−0.04	−0.07**	0.02
	(0.48)	(0.49)	(0.50)	(0.03)	(0.03)	(0.03)
Observations	502	469	468			

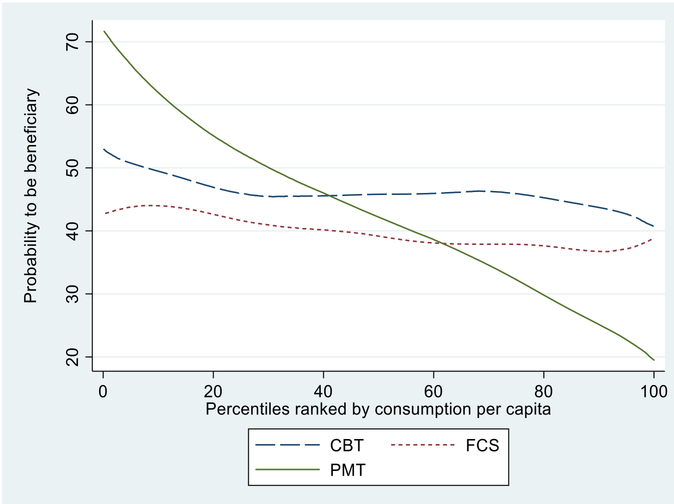
Source: Baseline Survey (beneficiary subsample).

Note: Consumption PC, AE, and EOS correspond to consumption per capita, per adult equivalent, and per capita adjusted for economies of scales, respectively. Consumption and income values are in thousands of FCFA (KFCFA). See Annex for a detailed definition of variables. Regressions include strata fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

higher by 8 percentage points than households selected by CBT, and higher by 5 percentage points than households selected by FCS. Similarly, households selected by PMT consume 15 KFCFA (or 13 percent) less than households selected by CBT, and 11 KFCFA (or 10 percent) less than households selected by FCS.[42] It shows that, even if the population is largely poor, there are meaningful welfare differences

42 The differences remain significant, though somewhat smaller in magnitude, when consumption is adjusted by economies of scale, or adult equivalency scales.

Figure 2. Targeting Efficiency along the Consumption Distribution



Source: Baseline Survey.
Note: The graph displays the probability of households to be selected as beneficiary (y axis) depending on their ranking along the distribution of baseline consumption per capita (percentiles in x axis). This is plotted separately for each targeting method: community-based targeting (CBT), proxy means testing (PMT) or a formula to proxy temporary food insecurity (FCS). The graph is based on a semiparametric regression model (Yatchew 1998) where commune and village fixed effects enter linearly in the functional form.

between households selected by alternative methods. The FCS and CBT methods perform similarly relative to poverty and consumption welfare benchmarks.

Figure 2 plots the likelihood of a household being selected as a beneficiary by each method along the baseline consumption distribution. The figure illustrates that PMT is relatively more efficient at including the poorest and excluding the better-off, based on consumption per capita. In contrast, the performance of CBT and FCS is rather flat along the consumption distribution.

The efficiency of targeting methods is often measured by calculating inclusion (or type 1) errors and exclusion (or type 2) errors. For each targeting method, inclusion errors (IE) are measured as the share of individuals with baseline welfare ($Y_{0,i}$) above the eligibility threshold (Z) that are selected for the program (B = 1).

$$IE_m = N\left(B = 1, Y_{0,i} > Z, T = m\right)/N\left(Y_{0,i} \geq Z, T = m\right), \quad \textit{with } m = CBT, FCS, \textit{respectively } PMT \tag{3}$$

Inclusion errors are calculated using consumption per capita.[43] Compared to a scenario of universal coverage,[44] the PMT method reduces inclusion errors by 31 percent (from 60 percent to 39 percent, see table S3.2). This illustrates some of the potential gains from targeting in the population. The CBT and

43 An alternative measure of inclusion error is also presented using a threshold equivalent to the national poverty line, as opposed to a threshold based on program eligibility (see table S3.2 in the supplementary online appendix). In addition, exclusion errors are measured (in table S3.2 in the supplementary online appendix) as the share of individuals with baseline welfare ($Y_{0,i}$) below the poverty line (Z) that are not selected for the program (B = 0):

$$EE_m = N(B = 0, Y_{0,i} < Z, T = m)/N(Y_{0,i} < Z, T = m), \quad \textit{with } m = CBT, FCS, \textit{respectively } PMT \tag{4}$$

44 This scenario assumes that every household in the sample benefits from the program.

FCS methods lead to slightly smaller reductions in inclusion errors, respectively to 42 percent and 45 percent. The difference is only weakly statistically significant between PMT and FCS, however.[45]

When using welfare measures other than consumption, no large difference is found in the efficiency of the three targeting methods. Table 1 shows that there are fewer differences in average characteristics of beneficiary households selected by each method across other welfare benchmarks such as (chronic) food insecurity, assets, livestock, income per capita, food stock coverage, as well as stunting, wasting and underweight of children under the age of five.[46]

The finding that methods perform similarly based on (chronic) food insecurity may be explained by the fact that the FCS method was designed to predict the food consumption score during the lean season, while the baseline survey measures the food consumption score during the harvest season. (This was due to delays in the baseline data collection resulting from implementation delays in the registry census). This finding is consistent with Schnitzer (2019), who finds that HEA performs much better than PMT in proxying (temporary) food insecurity during the lean season, but does less well in proxying (chronic) food insecurity in the harvest season. This is related to the limited correlation between households experiencing (temporary) food insecurity in the lean season and (chronic) food insecurity in the harvest season (which is less prevalent).

It is noteworthy that no significant difference in subjective well-being indicators is found between PMT and CBT, although FCS is less aligned with subjective well-being. For instance, among beneficiary households selected by CBT, 62 percent are perceived as very poor by other community members, and 64 percent have a low self-reported ability to satisfy basic needs. These indicators are not significantly different for households selected by PMT but are lower among households selected by FCS. These results contrast with those from Indonesia, where CBT was found to perform considerably better than PMT in matching community perceptions of poverty (Alatas et al. 2012). Determinants of targeting efficiency are further analyzed in section 5 below.

Targeting efficiency can also be assessed through differences in baseline welfare measures between beneficiaries (B = 1) and nonbeneficiaries (B = 0) selected by each method.

$$\left(Y_{0,i}|T=m\right) = \alpha + \gamma_m B_i + \pi_S + \varepsilon_i \quad with\, m = CBT, FCS, respectively\, PMT \tag{5}$$

This equation is estimated by OLS separately for villages assigned to each targeting method, including dummies for the randomization strata (π_S) and calculating robust (White-Huber) standard errors clustered at the village level. Table 2 presents the mean characteristics for beneficiaries and nonbeneficiaries, as well as the estimated difference in means in columns 3, 6, respectively 9.

The PMT method leads to the selection of households with higher poverty and lower consumption, as well as lower income per adult equivalent and food stock. Overall, the mean consumption per person of PMT-selected beneficiaries is 49 cents a day, compared to 63 cents for those not selected. While in absolute

45 The relative performance of methods varies depending on which eligibility threshold is used: for instance, whether it is the program eligibility threshold or the poverty line. This is because the performance of targeting methods differs along the consumption distribution (fig. 2). Table S3.2 in the supplementary online appendix provides results when using the poverty line as the eligibility threshold instead. Compared to a benchmark scenario of universal coverage, the PMT method reduces inclusion errors from 15 percent to 6 percent when using the national poverty line. Exclusion errors based on the national poverty line are large. This is explained by relatively low program coverage due to limited funding. While 85 percent of households in the project areas are poor, the program can only cover 40 percent of households in the project areas. Given the limited program coverage relative to needs, exclusion errors remain high with all methods.

46 The only statistical difference between households selected by CBT and FCS is a slightly higher share of severely wasted children in households selected by CBT. There are few statistical differences in the profile of households selected by PMT and CBT, aside from a slightly higher share of severely underweight and severely stunted children selected by PMT (significant at 10 percent). There are also only a few differences between households selected by FCS and PMT, with FCS households slightly less (chronically) food insecure, and having slightly more food stock. The few observed differences are small in magnitude, and with no clear or robust patterns.

Table 2. Differences in Characteristics between Beneficiaries and Nonbeneficiaries Selected by Each Method

	CBT			PMT			FCS		
	(1) NB	(2) B	(3) Diff. 1–2	(4) NB	(5) B	(6) Diff. 7–8	(7) NB	(8) B	(9) Diff. 4–5
Panel A: Household indicators									
Poor	0.85	0.85	0	0.8	0.93	−0.13***	0.85	0.88	−0.03
	(0.36)	(0.36)	(0.02)	(0.4)	(0.26)	(0.02)	(0.36)	(0.33)	(0.02)
Consumption PC	123	119	4	134	104	30***	122	114	8
(thousands FCFA)	(66)	(65)	(4)	(69)	(56)	(3)	(67)	(65)	(4)
Consumption AE	173	166	8*	182	147	35***	172	159	12*
(thousands FCFA)	(94)	(92)	(6)	(94)	(82)	(5)	(94)	(88)	(6)
Consumption EOS	146	142	4	157	126	31***	146	136	10
(thousands FCFA)	(77)	(76)	(5)	(80)	(67)	(4)	(78)	(75)	(5)
Food insecure	0.11	0.14	−0.03	0.11	0.15	−0.04	0.12	0.11	0.01
(chronic)	(0.31)	(0.35)	(0.02)	(0.31)	(0.36)	(0.02)	(0.32)	(0.32)	(0.02)
Asset index	0.05	−0.15	0.2***	0.04	−0.11	0.14	0.21	−0.16	0.37***
	(1.45)	(1.35)	(0.08)	(1.66)	(1.49)	(0.09)	(1.84)	(1.46)	(0.1)
Livestock index	0.15	−0.02	0.17***	0	−0.08	0.08	−0.02	−0.07	0.04
	(1.61)	(1.44)	(0.1)	(1.51)	(1.22)	(0.07)	(1.24)	(1.71)	(0.1)
Income AE	86	71	15**	89	66	23***	79	80	−1
(thousands FCFA)	(83)	(68)	(5)	(75)	(82)	(5)	(122)	(136)	(8)
Food stock coverage	3.53	3.12	0.41**	3.67	2.9	0.77***	3.54	3.34	0.19
	(3.02)	(2.1)	(0.16)	(2.35)	(2.03)	(0.13)	(2.38)	(2.51)	(0.16)
Observations	671	571	1,242	727	543	1,246	783	530	1,302
Panel B: Children indicators									
Severe stunting	0.23	0.21	0.02	0.21	0.27	−0.06	0.22	0.23	−0.02
	(0.37)	(0.35)	(0.03)	(0.36)	(0.37)	(0.04)	(0.34)	(0.36)	(0.03)
Severe wasting	0.04	0.07	−0.03	0.05	0.04	0	0.03	0.03	0
	(0.18)	(0.23)	(0.02)	(0.19)	(0.17)	(0.01)	(0.15)	(0.14)	(0.01)
Severe underweight	0.16	0.16	0.01	0.14	0.2	−0.07*	0.14	0.18	−0.03
	(0.32)	(0.31)	(0.03)	(0.3)	(0.34)	(0.02)	(0.3)	(0.34)	(0.03)
Observations	359	326	685	363	310	673	458	288	742
Panel C: Subjective well–being indicators									
Perceived very poor	0.34	0.62	−0.28***	0.33	0.59	−0.25***	0.33	0.52	−0.19***
by others	(0.47)	(0.49)	(0.04)	(0.47)	(0.49)	(0.04)	(0.47)	(0.5)	(0.03)
Low ability to satisfy	0.38	0.64	−0.26***	0.38	0.59	−0.22***	0.35	0.57	−0.22***
basic needs	(0.48)	(0.48)	(0.03)	(0.49)	(0.49)	(0.04)	(0.48)	(0.5)	(0.03)
Observations	602	502	1,104	639	469	1,104	723	468	1,181

Source: Baseline Survey.

Note: NB and B stand for nonbeneficiaries and beneficiaries, respectively. Consumption PC, AE, and EOS correspond to consumption per capita, per adult equivalent, and per capita adjusted for economies of scale, respectively. Consumption and income values are in thousands of FCFA (KFCFA). See Annex for a detailed definition of variables. Regressions include strata fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

terms the difference may seem small, it represents almost 30 percent of beneficiaries' consumption. PMT also selects households with a slightly higher share of underweight children than nonbeneficiaries.[47] The CBT process leads to the selection of households with fewer assets, fewer livestock, lower income per adult equivalent, and lower food stock. The FCS method leads to the selection of households with fewer household assets.[48] CBT, PMT and to a lesser extent FCS select households with lower subjective well-being indicators. For instance, beneficiary households selected by CBT, PMT, respectively FCS are 28, 25, respectively 19 percentage points more likely to be considered very poor by other community members than nonbeneficiary households.

4. Legitimacy of Targeting Methods

This section tests whether there are differences in the legitimacy of targeting methods among local populations after the program was rolled out. The perceived legitimacy of each method may depend either on the targeting process or on the targeting results. Four measures from the first follow-up survey ($L_{1,i}$) are considered. The first two questions capture overall legitimacy based on (i) whether respondents are very satisfied with the targeting method,[49] and (ii) whether respondents would like the same method to be used again in similar programs in the future.[50] The next two questions capture perceptions of accuracy in targeting results, namely (iii) whether respondents believe that *any* (nonpoor) household was wrongly included; and (iv) whether respondents believe that *any* (poor) household was wrongly excluded from the program. For each of these variables, an OLS regression is estimated. It includes indicators for the targeting method assigned to the village (using CBT as the reference category), as well as fixed effects for the randomization strata:

$$L_{1,i} = \alpha + \psi_1 PMT_i + \psi_2 FCS_i + \pi_S + \varepsilon_i \tag{6}$$

Robust (White-Huber) standard errors are clustered at the village level, and the estimation includes sampling weights. First, differences across methods are assessed by pooling all respondents. Equation (6) is then re-estimated separately for beneficiaries ($B = 1$) and nonbeneficiaries ($B = 0$).[51]

Overall, legitimacy appears relatively high across targeting methods (table 3):[52] 81 percent of sample households are very satisfied with the targeting process, with no statistical difference across methods; 80 percent of respondents state that they would like the same approach to be used again in the future, with few significant differences across methods, except for a small preference for FCS over CBT. Inclusion and exclusion issues are reported, but again with little differences across targeting methods. On average,

47 Related to assets, note that the difference in assets between PMT and other methods (CBT and FCS) is not significant in table 1. The finding that PMT beneficiaries do not have statistically significant lower assets than nonbeneficiaries in table 2 is not inconsistent with previous literature from Niger which highlights that the performance of PMT can vary depending on which asset is being considered (Schnitzer 2019). Results presented in the study are for an overall asset index.

48 FCS and CBT also lead to the selection of households with slightly lower consumption, though only based on the adult-equivalent measure (significant at 10 percent).

49 This measure is based on the following question, "*to what extent are you satisfied with the process used to identify program beneficiaries?*" Response options are given on a scale from 1 to 4 (1, unsatisfied; 2 a little satisfied, 3 satisfied, and 4 very satisfied). For the analysis, a dummy was created equaling 1 if very satisfied and 0 otherwise. Results are consistent using the continuous variable instead.

50 This is based on the following question: "*If there was a similar program, would you like to use the same approach to identify program beneficiaries?.*"

51 Section 5 also estimates equation (6) separately for villages with low and high inequality (by splitting the sample at the median of the Gini coefficient).

52 While there is no benchmark to compare levels of legitimacy with other studies, this assessment is based on the fact that little dissatisfaction is noted.

Table 3. Legitimacy and Perceptions of Accuracy of Targeting Methods – All Households

	(1) Very satisfied with process?	(2) Repeat same approach?	(3) Any non poor included?	(4) Any poor excluded?
PMT	−0.00142	0.0424	0.0887***	−0.0120
	(0.0291)	(0.0258)	(0.0301)	(0.0307)
FCS	0.0374	0.0610**	0.0394	0.0388
	(0.0270)	(0.0277)	(0.0265)	(0.0314)
Observations	1,250	1,250	1,250	1,250
R-squared	0.040	0.048	0.096	0.133
PMT-FCS	−0.0389	−0.0186	0.0493	−0.0508
	(0.0321)	(0.0246)	(0.0314)	(0.0294)
CBT mean	0.812	0.796	0.282	0.779
	(0.0239)	(0.0217)	(0.0244)	(0.0233)

Source: First Follow-up Survey.
Note: Regressions include strata fixed effects and sampling weights. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4. Legitimacy and Perceptions of Accuracy of Targeting Methods – Beneficiaries Only

	(1) Very satisfied with process?	(2) Repeat same approach?	(3) Any nonpoor included?	(4) Any poor excluded?
PMT	0.00371	−0.00746	0.0179	0.00730
	(0.0146)	(0.0202)	(0.0476)	(0.0399)
FCS	−0.0198	−0.0210	−0.0113	0.0347
	(0.0197)	(0.0217)	(0.0372)	(0.0424)
Observations	611	611	611	611
R-squared	0.025	0.071	0.109	0.119
PMT-FCS	0.0235	0.0136	0.0292	−0.0274
	(0.0191)	(0.0223)	(0.0428)	(0.0348)
CBT mean	0.986	0.952	0.250	0.761
	(0.0102)	(0.0161)	(0.0331)	(0.0339)

Source: First Follow-up Survey.
Note: Regressions include strata fixed effects and sampling weights. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

28 percent of respondents report that at least one nonpoor household has been selected. Exclusion issues are more prevalent, with 78 percent of respondents reporting that at least one poor household has been excluded.[53] The only significant difference is that more inclusion issues are reported with PMT than CBT.

However, legitimacy is much higher among beneficiaries (table 4) than among nonbeneficiaries (table 5): 99 percent of beneficiaries are "very satisfied," compared to only 65 percent of nonbeneficiaries; 95 percent of beneficiaries would like the same targeting approach to be used in the future, but only 65 percent of nonbeneficiaries. Perceptions of accuracy differ less, but nonbeneficiaries are also slightly more likely to report that there are poor, respectively nonpoor households wrongly targeted.

Importantly, there are sharper differences in legitimacy and perceived accuracy of targeting methods among nonbeneficiaries (table 5). Nonbeneficiaries are particularly unlikely to want the CBT approach to be repeated: they are 9.5 and 13.9 percentage points more likely to want to repeat the PMT, respectively

53 As discussed in section 3, it is expected that exclusion is more frequent, since the program operates in high-poverty areas, and does not have resources to cover all poor households in selected villages.

Table 5. Legitimacy and Perceptions of Accuracy of Targeting Methods – Nonbeneficiaries Only

	(1) Very satisfied with process?	(2) Repeat same approach?	(3) Any nonpoor included?	(4) Any poor excluded?
PMT	0.00513	0.0946**	0.159***	−0.0326
	(0.0565)	(0.0474)	(0.0454)	(0.0386)
FCS	0.0870*	0.139***	0.0927**	0.0438
	(0.0493)	(0.0493)	(0.0422)	(0.0395)
Observations	639	639	639	639
R-squared	0.086	0.095	0.131	0.171
PMT-FCS	−0.0819	−0.0448	0.0665	−0.0764
	(0.0629)	(0.0440)	(0.0458)	(0.0409)
CBT mean	0.650	0.651	0.313	0.796
	(0.0404)	(0.0354)	(0.0354)	(0.0322)

Source: First Follow-up Survey.

Note: Regressions include strata fixed effects and sampling weights. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the FCS approach.[54] This holds even though perceptions of accuracy favor CBT over PMT: in PMT (respectively FCS) villages, nonbeneficiaries are 16 percentage points (respectively 9 percentage points) more likely to say that there was a nonpoor household selected than nonbeneficiaries in villages where the CBT approach was used. This suggests that higher legitimacy is related to the targeting process used with formula-based methods, rather than the precision of targeting results. No significant difference is found across methods in terms of legitimacy and perceived accuracy among beneficiaries (table 4).[55]

5. Determinants of Efficiency and Legitimacy

This section considers mechanisms that can explain differences in the relative efficiency and legitimacy of targeting methods. First, it tests whether imperfect information and manipulation affect CBT performance. The results of CBT are sometimes considered as revealing local populations' own welfare rankings. The ability of CBT to consider local knowledge is widely considered a key advantage. However, *information failures* may arise when local populations do not have full information on other households' welfare levels.[56] *Manipulation* (of which *elite capture* is a form) is a common concern with CBT (Conning and Kevane 2002), though evidence is mixed (McCord 2013).[57] This section shows that these factors can explain why CBT is not necessarily perceived as the most legitimate process in the study context. The section then explores whether targeting legitimacy varies by local levels of inequality. On the one hand, Bardhan and Mookherjee (2000) show that elite capture is more likely in high-inequality settings. On the

54 This is in line with pretesting that took place prior to the study and qualitative work during the study. This is also consistent with expecting manipulation to happen, and manipulation being observed in the data (see "Manipulation" in section 5).

55 In addition to sampled households, in each village four leaders were selected to answer legitimacy questions. Leaders included the head of village, the president of the women's group, the religious leader, and the leader of the complaint committee. Table S3.6 in the supplementary online appendix replicates the legitimacy analysis for leaders. Results suggest patterns similar to the results at the household level, though estimates are more imprecise with few statistically significant results. The point estimates suggest that leaders may prefer to repeat formula-based approaches (though results are not significant), while perceptions of targeting accuracy are slightly better for CBT.

56 Alatas et al. (2016a) highlight how network efficiency affects information transmission and CBT results.

57 For instance, Pan and Christiaensen (2012) find evidence of elite capture in a voucher program in Tanzania. In contrast, Alatas et al. (2019) find little elite capture in the targeting process in Indonesia.

Table 6. Effect of Local Knowledge on Exclusion Errors

	(1) Women committee	(2) Leader committee	(3) Mixed committee	(4) Average ranking
Local knowledge	−0.359***	−0.376***	−0.315***	0.00957
	(0.0947)	(0.109)	(0.112)	(0.0959)
Observations	491	491	491	491
R-squared	0.119	0.113	0.105	0.105
Mean exclusion error	0.536	0.526	0.549	0.423
	(0.022)	(0.022)	(0.0219)	(0.0219)

Source: Baseline Survey (CBT subsample, households below eligibility cut-off of consumption distribution).
Note: The baseline survey asked each household to provide a welfare ranking of the other sampled households in the same village. For any given household, the "local knowledge" proxy captures the share of other sampled households in the village who report knowing that household (i.e., for whom the information about welfare ranking is not missing). Regressions include strata fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

other hand, CBT might not be able to identify households with different welfare levels in low-inequality settings where there are few observable welfare differences across households. Lastly, a range of other potential determinants of targeting performance are considered, including fatigue in CBT, the ability of methods to take into account the effects of shocks, as well as implementation issues in formula-based methods.

Imperfect Information

Information failures at the local level may hinder CBT targeting efficiency. Households who are not known by other village members may be more likely to be wrongly excluded. This hypothesis is tested by creating a proxy indicator of local knowledge. The baseline survey asked each household to provide a welfare ranking of other households sampled in the village. In 27 percent of cases, respondents report not knowing another household. This information is used to create a proxy variable for "local knowledge." For any given household, the proxy captures the share of other sampled households in the village who report knowing that household (that is, for whom the information about welfare ranking is not missing). An OLS regression of exclusion errors ($E_{0,i}$) is estimated on the local knowledge variable (K_i), including randomization strata fixed effects.

$$E_{0,i} = \alpha + \varphi K_i + \pi_S + \varepsilon_i \tag{7}$$

This is done separately for the rankings produced by each of the three committees and the combined ranking. Table 6 shows that imperfect local knowledge is associated with higher exclusion errors for all committees. A household that is known by a higher share of households in the village is less likely to be excluded by error by each committee. The effects are substantial, with exclusion probabilities increasing between 32 and 38 percentage points.[58] However, the triangulation of results across the three committees fully corrects for imperfect information. Based on the *combined* rankings from the three committees, households less well-known are not more likely to be excluded. While each committee has imperfect information, the creation of multiple committees thus effectively addresses information asymmetries. These results are consistent with findings in Alatas et al. (2016a) who show that more efficient networks improve the performance of CBT. However, results also highlight that a CBT process that would not triangulate results across multiple committees would suffer from information failures in the study setting.

58 Figure S3.3 in the supplementary online appendix illustrates that exclusion errors linearly decrease with local knowledge.

Table 7. Likelihood of Being a Beneficiary Based on CBT Committee Membership

	(1) Women committee	(2) Leader committee	(3) Mixed committee	(4) Average ranking
Committee member	0.0803	0.164**	0.241***	0.117***
	(0.0883)	(0.0732)	(0.0573)	(0.0379)
Observations	998	999	998	1,006
R-squared	0.133	0.152	0.172	0.162

Source: Baseline Survey (CBT subsample), and project monitoring data.
Note: Regressions control for consumption per capita, food consumption score, assets, livestock, food stock, and subjective well-being indicators. Regressions include strata fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Manipulation

Another concern with CBT is that committee members may manipulate the selection process and try to benefit themselves. The ranking provided by each of the three independent committees can be analyzed to assess whether manipulation occurs. This would arise if committee members' households have a higher chance to be selected as beneficiaries. Estimation controls for baseline household welfare measures.[59]

Table 7 shows that some committee members attempt to manipulate the process. After controlling for baseline welfare measures, the probability that committee members' households are selected is higher than for other households. It is statistically significant for two of the three committees: households of the mixed nonleader and leader group have a 24.1 and 16.4 percentage point higher chance to be selected as beneficiaries, respectively. No significant manipulation is observed in the women committee. Triangulating rankings across several committees only partly mitigates manipulation. Based on the average ranking, committee members' households have an 11.7 percentage point higher chance to be selected as beneficiaries (significant at the 1 percent level).[60]

These results can explain why CBT was not considered the most legitimate process by nonbeneficiaries. Results are also consistent with respondents' perceptions: since 48 percent declare that committee members try to benefit themselves. Households displaying low trust (by believing committee members try to benefit themselves) are also 12 percentage points less likely to want to repeat the CBT approach.

Variation in Targeting Performance by Local Inequality Levels

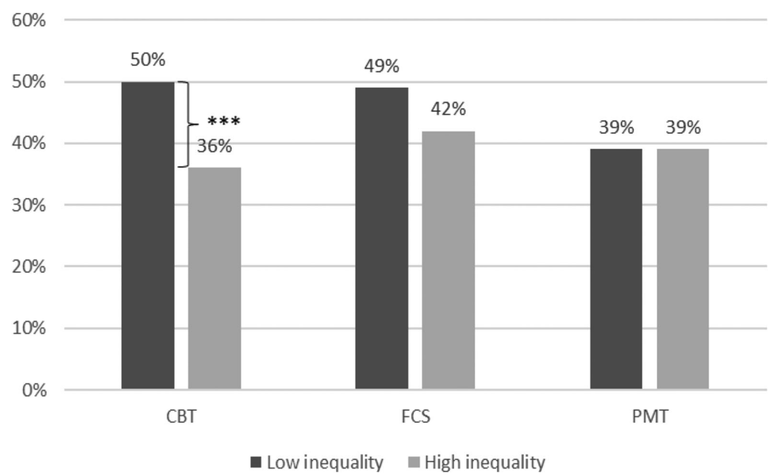
Empirical studies are rarely able to analyze whether local levels of inequality affect targeting performance. This section assesses variations in targeting efficiency and legitimacy between low and high-inequality villages. This is done by splitting the sample at the median of the village-level Gini coefficient from the registry census.

Figure 3 illustrates results for inclusion errors. The efficiency of formula-based methods (particularly PMT) remains similar at low and high levels of inequality. In contrast, the results of CBT vary strongly by levels of inequality: inclusion errors are 14 percentage points higher in low-inequality than in

59 Controls include consumption per capita, the food consumption score, assets, livestock, food stock, and poverty perceptions by other community members.

60 A similar analysis was performed to assess whether committee members' relatives also had a higher chance to be selected as beneficiaries. Table S3.3 in the supplementary online appendix shows these results. No significant difference is found, suggesting that risks of manipulation primarily stem from committee members trying to favor their own household. Another consideration is whether inclusion errors change when restricting the sample to non–committee members. No significant difference is found (table S3.4 in the supplementary online appendix). This is in part be due to committee members representing only a small fraction of households (a median of 12 percent of village households). Table S3.5 in the supplementary online appendix provides descriptive statistics about committee members. Results show almost no welfare differences between groups. The only difference is that the mixed-gender committee seems moderately worse in terms of food security outcomes than the leader committee.

Figure 3. Inclusion Errors (by Village Inequality Levels)



Source: Baseline Survey.
Note: This figure displays inclusion errors for each targeting method (CBT: community-based targeting, PMT: proxy means testing or FCS: formula to proxy temporary food insecurity) between low- and high-inequality villages. Low-inequality villages have Gini coefficient below the median. Gini coefficients are estimated from the registry census.

Table 8. Legitimacy and Perceptions of Accuracy of Targeting Methods – Low-Inequality Villages

	(1) Very satisfied with process?	(2) Repeat same approach?	(3) Any nonpoor included?	(4) Any poor excluded?
PMT	−0.0515	0.0503	0.0739*	−0.00232
	(0.0386)	(0.0303)	(0.0430)	(0.0377)
FCS	0.0257	0.0781**	0.0229	−0.0170
	(0.0429)	(0.0357)	(0.0406)	(0.0362)
Observations	619	619	619	619
R-squared	0.065	0.078	0.124	0.179
PMT-FCS	−0.0772	−0.0278	0.0510	0.0146
	(0.0481)	(0.0298)	(0.0510)	(0.0389)
CBT mean	0.815	0.773	0.278	0.750
	(0.0321)	(0.0315)	(0.032)	(0.0352)

Source: First Follow-up Survey.
Note: Regressions include strata fixed effects and sampling weights. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

high-inequality villages. CBT is as efficient as formula-based methods in high-inequality villages, but more inefficient in low-inequality villages.

Tables 8 and 9 show how legitimacy varies between low- and high-inequality villages. The higher legitimacy of formula-based methods may be partly driven by low-inequality villages, where households assigned to formula-based methods are 5 (marginally insignificant) and 7.8 percentage points more likely to prefer repeating PMT, respectively FCS compared to households assigned to the CBT approach. In contrast, legitimacy is found to be similar across methods in high-inequality villages. Relative to CBT, perceived inclusion and exclusion errors also tend to be more consistently significant for PMT and FCS in high-inequality villages.

Results can be interpreted as consistent with CBT committee members being better able to identify worse-off households in higher-inequality settings. They are not consistent with higher elite capture in high inequality settings (as in Bardhan and Mookherjee 2000).

Table 9. Legitimacy and Perceptions of Accuracy of Targeting Methods – High-Inequality Villages

	(1) Very satisfied with process?	(2) Repeat same approach?	(3) Any nonpoor included?	(4) Any poor excluded?
PMT	0.0400	0.0341	0.107***	−0.00468
	(0.0373)	(0.0389)	(0.0303)	(0.0471)
FCS	0.0548*	0.0602	0.0802**	0.0773**
	(0.0326)	(0.0384)	(0.0360)	(0.0384)
Observations	631	631	631	631
R-squared	0.049	0.065	0.101	0.111
PMT-FCS	−0.0148	−0.0260	0.0272	−0.0820
	(0.0460)	(0.0395)	(0.0386)	(0.0494)
CBT mean	0.809	0.817	0.287	0.806
	(0.0352)	(0.0299)	(0.0364)	(0.0305)

Source: First Follow-up Survey.
Note: Regressions include strata fixed effects and sampling weights. Robust standard errors are clustered at the village level and shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Other Factors: Fatigue, Shocks, and Implementation Issues in Formula-Based Methods

A range of other factors can affect targeting performance. Alternative explanations are explored in the supplementary online appendix. First, no statistically significant effects of fatigue are found in CBT. Second, some targeting methods may in principle better account for transient poverty factors. However, no differences are found across methods in the probability of excluding households exposed to shocks in the study setting. Finally, potential implementation issues with formula-based methods are found to have little effects on targeting efficiency.

6. Impacts of Targeting Methods on Program Effectiveness

This section complements the analysis of targeting efficiency and legitimacy by testing whether the choice of targeting method affects program effectiveness. Specifically, it analyzes whether the impacts of the cash transfer program vary across villages assigned to different targeting methods. The focus is on key welfare indicators related to core program objectives, including proxies for food security in the short term, as well as consumption and poverty. This section does not aim to estimate program impacts directly,[61] but rather to highlight if there are differences in program impacts across targeting methods.

The impacts of targeting choice on program effectiveness are measured for outcomes collected in the first and second follow-up surveys. The first follow-up survey took place a few months after the start of the cash transfer program. It captures four proxy measures of food security: a (chronic) food insecurity dummy, the food consumption score, the number of months household food stock could last (food coverage), and a hunger scale (see annex for detailed description of each indicator). It also includes an indicator measuring whether any conflict was reported between the household and another community member. This is to assess potential unintended effects of targeting choice. The second follow-up survey was collected one year after the start of the program and included outcomes such as consumption per capita and poverty.

61 There are two studies analyzing the effectiveness of earlier phases of the Niger cash transfer program. Stoeffler, Mills, and Premand (2020) analyze the impact of a cash transfers combined with savings promotion on assets 18 months after the termination of the transfers. Premand and Barry (2020) isolate the value-added of a behavioral change component delivered to cash transfer beneficiaries on parenting practices and child development outcomes.

Table 10. Impacts of Targeting Methods (First Follow-Up Survey, All Sample)

	(1) Food insecure	(2) Food consumption score	(3) Food stock coverage	(4) Hunger	(5) Conflict
PMT × POST	0.0246	1.935	0.556**	0.00241	0.0117
	(0.0369)	(2.067)	(0.270)	(0.0319)	(0.0591)
FCS × POST	0.0376	1.518	−0.191	−0.00252	−0.0452
	(0.0360)	(2.174)	(0.245)	(0.0363)	(0.0601)
PMT	0.0106	−1.694*	0.0574	0.00329	−0.0517
	(0.0151)	(0.930)	(0.130)	(0.0170)	(0.0348)
FCS	−0.00427	−0.291	0.105	−0.00911	−0.00398
	(0.0166)	(0.930)	(0.132)	(0.0184)	(0.0343)
POST	−0.0495**	0.545	0.400**	0.0400	0.360***
	(0.0224)	(1.356)	(0.195)	(0.0253)	(0.0404)
Observations	4,318	4,318	4,215	4,214	4,309
R-squared	0.046	0.053	0.078	0.021	0.077
Diff. effects PMT & FCS	−0.0129	0.417	0.747***	0.00492	0.0568
	(0.0406)	(2.329)	(0.242)	(0.0326)	(0.0622)
CBT mean	0.126	44.59	3.522	0.841	3.601
	(0.005)	(0.281)	(0.0374)	(0.0055)	(0.0109)

Source: Baseline and First Follow-up Survey.

Note: See annex for a detailed definition of variables. Difference-in-differences estimates are based on equation (8). Regressions include strata fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. Number of observations for data in long format (observations with baseline and follow-up information is counted twice). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Targeting methods were randomized at the village level, and characteristics were balanced at baseline between villages (see table S3.1). The analysis tests whether short-term outcomes vary across all households in the study villages. This indicates whether the cash transfer program had differential impacts in villages assigned to each targeting method. A difference-in-differences specification is estimated including all households in the village, independently of their status as beneficiary or nonbeneficiary:[62]

$$Y_{t,i} = \delta_1 PMT_i * \ POST_i + \delta_2 FCS_i * \ POST_i + \ \delta_3 POST_i + \ \delta_4 PMT_i + \ \delta_5 FCS_i + \pi_S + \varepsilon_{ti} \tag{8}$$

POST is a time dummy taking the value of 0 at baseline, and 1 at follow-up. As mentioned above, the focus is not on the impact of the program itself, but on whether the choice of targeting method leads to changes in program effectiveness. δ_1 and δ_2 are thus the parameters of interest.

The impacts of the cash transfer program may vary depending on the profile of selected households. If there is no trade-off between targeting efficiency and program effectiveness, larger program impacts on welfare would be expected for targeting methods identifying poorer populations.

Table 10 presents the results. No difference is found in the food consumption score, hunger and conflict indicators. However, larger cash transfer program impacts are found on food stock in villages where PMT targeting is used.[63] Food stocks are on average 16 percent higher across all households in villages randomized to PMT targeting compared to villages randomized to CBT targeting, and 21 percent higher compared to villages assigned to FCS.

Average differences across all households (independently of beneficiary status) can be estimated relying on the randomized design at the village level (such as in table 10). Yet these comparisons may not be sufficiently powered since only a subset of households is covered by the program. The analysis thus proceeds to test whether impacts of the cash transfer program vary between beneficiaries selected by alternative targeting methods. A difference-in-differences approach is used to account for differences in the profile of

62 Given that the randomization was performed at the village level, this analysis could be performed by calculating single-difference in short-term outcomes at follow-up, which provides consistent results.

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

Table 11. Impacts of Targeting Methods (First Follow-Up Survey, Placebo Test on Nonbeneficiaries Only)

	(1) Food insecure	(2) Food consumption score	(3) Food stock coverage	(4) Hunger	(5) Conflict
PMT × POST	0.0235	0.279	0.503	0.01000	0.0256
	(0.0426)	(2.723)	(0.371)	(0.0493)	(0.0708)
FCS × POST	0.0115	1.101	0.0667	0.0184	−0.0602
	(0.0403)	(2.556)	(0.339)	(0.0527)	(0.0786)
PMT	0.00647	−1.324	0.132	0.0223	−0.0491
	(0.0184)	(1.199)	(0.163)	(0.0230)	(0.0425)
FCS	0.0124	−0.669	0.0251	0.0118	−0.0101
	(0.0188)	(1.091)	(0.164)	(0.0244)	(0.0437)
POST	−0.0308	−0.383	0.0954	0.0519	0.342***
	(0.0295)	(1.819)	(0.280)	(0.0395)	(0.0470)
Observations	2,537	2,537	2,464	2,462	2,533
R-squared	0.040	0.052	0.076	0.025	0.079
Diff. effects PMT & FCS	0.0120	−0.822	0.436	−0.00835	0.0859
	(0.0411)	(2.730)	(0.311)	(0.0460)	(0.0819)
CBT mean	0.120	44.98	3.614	0.826	3.593
	(0.0064)	(0.377)	(0.0514)	(0.0076)	(0.0144)

Source: Baseline and First Follow-up Survey (nonbeneficiary subsample).
Note: See annex for a detailed definition of variables. Difference-in-differences placebo estimates based on equation (9). Regressions include strata fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. Number of observations for data in long format (observations with baseline and follow-up information is counted twice). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

households selected by each method (as documented in section 3).[64] It focuses on the outcomes observed in the baseline and follow-up surveys. Restricting the sample to beneficiaries, a regression of the outcome variable is estimated on a dummy capturing the targeting method assigned to the village in which households resided at baseline, a dummy for each of the two survey rounds, and interactions between the round dummy and the targeting method dummies:

$$(Y_{t,i}|B = 1) = \delta_1 PMT_i * POST_i + \delta_2 FCS_i * POST_i + \delta_3 POST_i + \delta_4 PMT_i + \delta_5 FCS_i + \pi_S + \theta_t + \varepsilon_{ti} \tag{9}$$

δ_1 and δ_2 are the parameters of interest. They provide an estimate of the difference in program impacts across groups selected by different targeting methods.

The difference-in-differences approach relies on the assumption that, absent the intervention, outcomes for beneficiaries selected by each targeting method would have evolved in parallel over time. To test whether time-varying factors affect households differently, a placebo test is run by estimating the same specification in the sample of nonbeneficiaries with data in the short-term follow-up survey:

$$(Y_{t,i}|B = 0) = \delta_1 PMT_i * POST_i + \delta_2 FCS_i * POST_i + \delta_3 POST_i + \delta_4 PMT_i + \delta_5 FCS_i + \pi_S + \theta_t + \varepsilon_{ti} \tag{10}$$

This provides a test of the equal trend assumption on which the difference-in-differences estimator relies. The parameters δ_1 and δ_2 are expected to be 0.[65] Table 11 presents the results of the placebo test. All coefficients are statistically insignificant and no spurious difference in trend is found.[66]

Table 12 presents the results on the effect of targeting choice on program effectiveness based on the estimation of equation (10) for outcomes from the first (columns 1–5) and second (columns 6–7)

64 Table S3.13 in the supplementary online appendix provides descriptive statistics at the first and second follow-up survey, including differences between beneficiary and nonbeneficiary households for the main outcome indicators.

65 To complement the specifications in equations (9) and (10), an ANCOVA specification is also estimated (McKenzie 2012) and results are robust.

66 All coefficients in table 11 are very close to zero and far from being significant. The only exception is the food stock variable in PMT villages, which suggests a 13.9 percent difference in food stocks, but it is not close to being marginally significant.

Table 12. Impacts of Targeting Choice – Beneficiary Sample

	First Follow-up Survey					Second Follow-up Survey	
	(1) Food insecure	(2) Food consumption score	(3) Food stock coverage	(4) Hunger	(5) Conflict	(6) Poverty	(7) Consumption
PMT × POST	0.0190	3.872	0.696**	−0.00198	−0.00928	−0.0665*	8,815
	(0.0503)	(2.648)	(0.345)	(0.0457)	(0.0849)	(0.0383)	(7,130)
FCS × POST	0.0708	2.181	−0.401	−0.0171	−0.0282	−0.0544	9,694
	(0.0513)	(2.855)	(0.304)	(0.0470)	(0.0825)	(0.0395)	(7,610)
PMT	0.0217	−2.533**	−0.104	−0.0205	−0.0524	0.089***	−14,565***
	(0.0223)	(1.170)	(0.163)	(0.0222)	(0.0517)	(0.0255)	(5,019)
FCS	−0.0246	−0.408	0.190	−0.0394*	0.00482	0.0255	−2,335
	(0.0228)	(1.311)	(0.171)	(0.0230)	(0.0519)	(0.0274)	(5,427)
POST	−0.0717**	1.725	0.733***	0.0215	0.388***	−0.0163	15,514***
	(0.0315)	(1.799)	(0.232)	(0.0321)	(0.0604)	(0.0275)	(5,171)
Observations	1,769	1,769	1,739	1,740	1,764	2,470	2,470
R-squared	0.069	0.080	0.116	0.030	0.083	0.036	0.086
Diff. effects PMT & FCS.	−0.0518	1.691	1.097***	0.0152	0.0189	−0.0121	−879.4
	(0.0564)	(2.949)	(0.330)	(0.0476)	(0.0837)	(0.0390)	(74330)
CBT mean	0.133	44.02	3.397	0.860	3.614	0.767	125,603
	(0.0078)	(0.421)	(0.0541)	(0.0081)	(0.0165)	(0.0077)	(1276)

Source: Baseline, First Follow-Up Survey, Second Follow-Up Survey (beneficiary subsample).
Note: See annex for a detailed definition of variables. Difference-in-differences placebo estimates based on equation (10). Regressions include strata fixed effects. Robust standard errors are clustered at the village level and shown in parentheses. Number of observations for data in long format (observations with baseline and follow-up information is counted twice). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

follow-up surveys. No significant difference in impacts is found on (chronic) food security, the food consumption score, and the hunger scale. However, results show that short-term program impacts on food stocks are larger in the villages where the PMT method was implemented. Compared to villages assigned to CBT (respectively FCS), the impact on food stocks is higher by 0.7 months or 20 percent (respectively 1.1 months or 32 percent) in villages assigned to PMT, and the difference is significant at the 5 percent level (respectively 1 percent). No significant effect is found on the variable capturing the frequency of conflict between beneficiaries selected by alternative targeting methods.[67]

The results on food stocks are consistent with cash transfers having relatively larger impacts on households with higher unmet food needs. Various baseline indicators suggest that households selected by PMT had larger unmet food needs relative to those selected by CBT or FCS.[68] There are two potential explanations for these observed impacts.[69] First, the results are consistent with beneficiaries buying a relatively larger share of food from the market, reducing the share of consumption from own production and

67 Comparing PMT with CBT beneficiaries, the point estimates are positive for food insecurity (0.0190) and the food consumption score (3.872). These are respectively 14 percent and 8.8 percent changes from the mean in the CBT group, but the effects are far from being statistically significant. On the other hand, the coefficients for the effects on the hunger scale (−0.00198, or 0.2 percent) and conflict (−0.00928 or 0.26 percent) are very close to zero in magnitude.

68 For instance, households selected by PMT had lower consumption per capita, higher underweight and stunting rates, and, while differences were not statistically significant, they also presented a lower food security score and lower food stocks.

69 The supplementary online appendix describes whether households selected based on alternative targeting methods report using cash transfers differently. Table S3.14 in the supplementary online appendix documents (i) the share of households declaring spending cash transfers across various categories, (ii) average spending amounts, and (iii) average spending shares by categories. Households report spending between 65 and 69 percent of the transfers on food. Savings account

saving larger stocks (at a given level of consumption and production). Second, higher food stocks would be observed if cash transfer beneficiaries increase production but keep the same share of production from own consumption and the same consumption level. While it is not possible to formally disentangle these two mechanisms, the relative in food purchases seems the more likely channel, given the timing of the study.[70]

Program impacts on poverty reduction tend to be larger among households selected by formula-based methods, in particular PMT. Poverty reduction is larger by 6.7 pp (8.7 percent) in PMT villages compared to CBT villages (significant at 10 percent level). The effect on poverty is also negative but nonsignificant in FCS villages compared to CBT villages (5.4 pp or 7 percent). The point estimates on average consumption per capita are positive and of similar magnitudes for villages assigned to formula-based targeting methods (PMT and FCS) compared to CBT (respectively by 7 and 7.7 percent), but the differences are not statistically significant.[71] The effects of targeting choice on consumption and poverty impacts are not consistently statistically significant. They show that there is no trade-off between targeting efficiency and program effectiveness in the context of the study.

Ultimately, 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 improves program effectiveness based on some welfare dimensions (but not all). On the other hand, the costs of methods studied in this paper are very close, and only amount to 1.1–1.5 percent of program benefits. The cost of collecting registry data and applying PMT or FCS methods is roughly $6.8 USD per household. The cost of CBT is $5.4 USD per household, which represents a lower bound estimate as it doesn't account for the household listing that was taken from the registry census.[72]

7. Conclusion and Discussion

This study tests the relative performance of three commonly used targeting methods based on their implementation in a large-scale randomized experiment in Niger. Three dimensions of targeting performance are considered: efficiency, legitimacy, and the impact of targeting choice on program effectiveness.

Findings show that 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, although CBT and PMT are more closely aligned with subjective well-being than FCS. Importantly, nonbeneficiaries find formula-based methods (PMT and FCS) more legitimate than CBT. To explain why, a range of potential determinants of targeting performance are analyzed. CBT is affected by imperfect information at the local level and by manipulation since selection committee members attempt to prioritize their households. Local populations are aware of these issues, with approximately half the respondents reporting

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. There is no significant difference in the use of transfers across the groups selected by alternative targeting methods.

70 Cash transfer impacts on agricultural production through investments in inputs have been documented in the literature and would be consistent with the second channel. However, the study timeline may be too short to realistically expect such effects: the first cash transfer payments were made right at the beginning of the agricultural cycle, and the food security measures were obtained around harvest time. It is unlikely that cash transfers starting during the agricultural season were sufficient and delivered early enough for households to purchase inputs and immediately increase production in that same season.

71 It is not necessarily inconsistent to find effects on poverty reduction but not on average consumption. There could be distributional effects concentrated around the poverty threshold. Still, the results on consumption and poverty remain largely consistent with effects of similar magnitude, even though their statistical significance is marginally different.

72 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.

they expect committee members to benefit themselves. CBT appears to perform worse in low-inequality villages.

The results obtained in the context of Niger contrast with those from other settings. Findings on the relative legitimacy of PMT and CBT among local populations differ from those obtained by Alatas et al. (2012) in Indonesia. While both studies find limited differences in targeting efficiency between methods, communities prefer CBT over PMT in Indonesia, which is not the case in Niger. Out-of-sample comparisons are bound to speculative, but there are sharp differences across the two contexts. Indonesia is an emerging lower-middle income country with a relatively low poverty rate, while Niger is a low-income country with a higher poverty rate. Inequality is also slightly higher in Indonesia than in Niger.[73] The observed determinants of targeting performance in the study setting can further help interpret the differences in results. For instance, in the context of Niger, imperfect information at the local level and manipulation are shown to affect CBT, while these factors were not observed in Indonesia.

The study results show that concerns about manipulation and imperfect information in CBT are in part justified, but adjustments in CBT implementation can improve targeting performance. This is consistent with reviews that highlight the role of targeting implementation (Coady, Grosh, and Hoddinott 2004a; Devereux et al. 2017). The study findings have clear practical implications. The selection of committee members as well as the triangulation of results across multiple selection committees mitigate local information failures and (in part) the risk of manipulation when CBT is used.

The study results also contrast with those from Basurto, Dupas, and Robinson (2019), who find a trade-off between reaching the largest share of poor households and achieving the highest returns from a subsidy program in Malawi. In the context of cash transfers in Niger, no trade-off between targeting efficiency and program effectiveness is found. PMT appears to be relatively more efficient in reaching poor households, and program impacts on some welfare dimensions tend to be larger among households selected by PMT (compared to CBT). Potential trade-offs between targeting efficiency and program effectiveness may be less of a concern in the context of social protection programs seeking to improve food security and provide consumption support. These trade-offs may be stronger for other interventions, such as those aiming to improve livelihoods and raise earnings, though this topic would deserve additional research.

The complexity of selecting beneficiaries for social programs is particularly salient in low-income settings. Despite high poverty rates, social programs typically have limited budgets and coverage. The study highlights some of the challenges with targeting in these settings, while also documenting some meaningful differences in performance across methods. The clearest differences arise between CBT and PMT, with CBT appearing less efficient, legitimate, and effective than PMT. The study thus calls for caution when considering CBT due to risks of manipulation and information asymmetries. On the other hand, differences in performance between the formula-based methods (FCS and PMT) are less clear cut. The baseline survey took place slightly after the end of the lean season, which limits the possibility of providing firm conclusions on the performance of FCS in reaching temporarily food insecure households. Additional research would be welcome to assess the performance of targeting methods for seasonal or shock-response interventions.

Ultimately, variations in performance across targeting methods are not large, especially relative to the overall level of exclusion stemming from limited funding for social programs. In this sense, policy discussions at early stages of social protection systems may benefit from focusing less on the determination of a single optimal targeting method. Instead, they could focus more 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 based on harmonized data.

73 In 2014, Indonesia had a poverty rate of 7.9, and Niger had a poverty rate of 44.5. Further, the Gini coefficient was 39.4 in Indonesia and 34.3 in Niger.

Annex. Definition of Key Variables

Variable	Description
Poor	A household is considered poor if its aggregate annual consumption is below the national poverty line (using the latest available 2014 national poverty line at FCFA 189,233.2).
Consumption per capita (PC)	Consumption includes annual expenditures and self-consumption for food, nonfood (incl. health and schooling) and durables items. Consumption per capita is the total consumption divided by the household size. Consumption is displayed in thousands of FCFA.
Consumption adjusted for economies of scale (ECOS)	Consumption includes annual expenditures and self-consumption for food, nonfood (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, at a lower average unit price. Consumption is displayed in thousands of FCFA.
Consumption per adult equivalent (AE)	Consumption includes annual expenditures and self-consumption for food, nonfood (incl. health and schooling) and durables items. Consumption adjusted for adult equivalence takes account of the fact that children or the elderly do not consume as much food as adults. The standard FAO scale was used to make gender-age specific adjustments for adult equivalence. Consumption is displayed in thousands of FCFA.
Food consumption score	Standard measure of household dietary diversity (World Food Programme 2008). It is a composite score based on dietary diversity, food frequency, and relative nutritional importance of different food groups. Food groups include cereals and cereal products, tubers and plantains, legumes and seeds, vegetables, fish and meat, fruits, milk and milk products, oil and fat, sugar products, spices and condiments.
Food insecurity	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 a set of productive and nonproductive 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 owned of calves, cows, cattle, chickens, sheep, goats, donkeys; 0.2367 of the overall variation is explained by the first component.
Income per capita	Income is the total of all reported income (monetary earnings) from crop sales, self-employment, wage-employment, livestock, and received transfers. Income per capita is the total income divided by the household size. Income is displayed in thousands of FCFA.
Food stock coverage	Duration of food coverage using own agricultural production (in months) in past year.
Severely stunted	A child is severely stunted if its height-age z-score is 3 z-scores below the age-gender specific WHO world medians.
Severely wasted	A child is severely wasted if its height-weight z-score is 3 z-scores below the age-gender specific WHO world medians.
Severely underweight	A child is severely underweight if its weight-age z-score is 3 z-scores below the age-gender specific WHO world medians.
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, e.g., is blind, deaf, or mute; or has mental illness, leprosy.
Education	An individual that has *any* level of education.
Land ownership	A household that owns any cultivated land.

Variable	Description
Plots cultivated	Total number of plots cultivated by the household.
Perceived very poor by others	Baseline households were asked to rank other sample households by poverty levels. This dummy variable equals 1 if the average of rankings provided by other baseline respondents falls below the program cut-off line (bottom 4 quartiles of the distribution).
Low ability to satisfy basic needs	This measure is based on the average response across six questions that ask respondents how satisfied they feel (on a scale from 1 to 3), with respect to six basic needs, including food, housing conditions, health, education, clothing, and revenues. The dummy variable equals 1 if the ranking based on self-assessed basic needs falls below the program cut-off line (bottom 4 quartiles of the distribution).
Covariate shocks	Over the past 12 months, household was affected by any covariate shocks (including drought, flooding, food price increase).
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	Household reports that, in the past 12 months, it experienced a situation where there was not enough food for the entire household.
Conflict	Household reports that, in the past 12 months, there was 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.

References

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias. 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review* 102 (4): 1206–40.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2019. "Does Elite Capture Matter? Local Elites and Targeted Welfare Programs in Indonesia." *AEA Papers and Proceedings* 109: 334–39.

Alatas, Vivi, Abhijit Banerjee, Arun G. Chandrasekhar, Rema Hanna, and Benjamin A. Olken. 2016a. "Network Structure and the Aggregation of Information: Theory and Evidence from Indonesia." *American Economic Review* 106 (7): 1663–704.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2016b. "Self-Targeting: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 124 (2): 371–427.

Bah, Adama, Samuel Bazzi, Sudarno Sumarto, and Julia Tobias. 2018. "Finding the Poor vs. Measuring Their Poverty: Exploring the Drivers of Targeting Effectiveness in Indonesia." *World Bank Economic Review* 33 (3): 573–97.

Banerjee, Abhijit, Paul Niehaus, and Tavneet Suri. 2019. "Universal Basic Income in the Developing World." *Annual Review of Economics* 11 (1): 959–83.

Bardhan, Pranab K., and Dilip Mookherjee. 2000. "Capture and Governance at Local and National Levels." *American Economic Review* 90 (2): 135–39.

Basurto, Maria Pia, Pascaline Dupas, and Jonathan Robinson. 2019. "Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi." *Journal of Public Economics* 185: 1–25.

Bertelli, Olivia. 2020. "Food Security Measures in Sub-Saharan Africa. A Validation of the LSMS-ISA Scale." *Journal of African Economies* 29 (1): 90–120.

Bossuroy, Thomas, Markus Goldstein, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Christopher Udry, and Julia Vaillant. 2019. "Promoting Productive Inclusion and Resilience among the Poor: Multi-country RCT of the Sahel Adaptive Social Protection Program." AEA RCT Registry.

Brown, Caitlin, Martin Ravallion, and Dominique van de Walle. 2018. "A Poor Means Test? Econometric Targeting in Africa." *Journal of Development Economics* 134: 109–24.

Cafiero, Carlo, Hugo R. Melgar-Quiñonez, Terri J. Ballard, and Anne W. Kepple. 2014. "Validity and Reliability of Food Security Measures." *Annals of the New York Academy of Sciences* 1331 (1): 230–48.

Coady, David, Margaret E. Grosh, and John Hoddinott. 2004a. *Targeting of Transfers in Developing Countries: Review of Lessons and Experience*. Vol. 1. Washington, DC: World Bank Publications.

———. 2004b. "Targeting Outcomes Redux." *World Bank Research Observer* 19 (1): 61–85.

Coady, David, and Emmanuel Skoufias. 2004. "On the Targeting and Redistributive Efficiencies of Alternative Transfer Instruments." *Review of Income and Wealth* 50 (1): 11–27.

Conning, Jonathan, and Michael Kevane. 2002. "Community-Based Targeting Mechanisms for Social Safety Nets: A Critical Review." *World Development* 30 (3): 375–94.

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

Devereux, Stephen, Edoardo Masset, Rachel Sabates-Wheeler, Michael Samson, Althea-Maria Rivas, and Dolf te Lintelo. 2017. "The Targeting Effectiveness of Social Transfers." *Journal of Development Effectiveness* 9 (2): 162–211.

Gentilini, Ugo, Margaret Grosh, Jamele Rigolini, and Ruslan Yemtsov. 2020. *Exploring Universal Basic Income: A Guide to Navigating Concepts, Evidence, and Practices*. Washington, DC: World Bank.

Grosh, Margaret, and Judy L. Baker. 1995. "Proxy Means Tests for Targeting Social Programs." Working Paper 118, Living Standards Measurement Study, World Bank. Washington, DC, USA.

Hanna, Rema, and Benjamin A. Olken. 2018. "Universal Basic Incomes versus Targeted Transfers: Anti-Poverty Programs in Developing Countries." *Journal of Economic Perspectives* 32 (4): 201–26.

Karlan, Dean, and Bram Thuysbaert. 2019. "Targeting Ultra-poor Households in Honduras and Peru." *World Bank Economic Review* 33 (1): 63–94.

Klasen, Stephan, and Simon Lange. 2016. "How Narrowly Should Anti-poverty Programs Be Targeted? Simulation Evidence from Bolivia and Indonesia." Discussion Paper Series No. 213, Göttingen University. Göttingen, Germany.

McBride, Linden, and Austin Nichols. 2016. "Retooling Poverty Targeting Using Out-of-Sample Validation and Machine Learning." *World Bank Economic Review* 32 (3): 531–50.

McCord, A. 2013. "Community-Based Targeting in the Social Protection Sector." Working Paper No. 514, Overseas Development Institute, London, UK.

McKenzie, D. 2012. "Beyond Baseline and Follow-Up: The Case for More T in Experiments." *Journal of Development Economics* 99 (2): 210–21.

Pan, Lei, and Luc Christiaensen. 2012. "Who is Vouching for the Input Voucher? Decentralized Targeting and Elite Capture in Tanzania." *World Development* 40 (8): 1619–33.

Premand, P., and O. Barry. 2020. "Behavioral Change Promotion, Cash Transfers and Early Childhood Development, Experimental Evidence from a Government Program in a Low-income Setting." World Bank Policy Research Working Paper No. 9368, World Bank: Washington, DC, USA.

Ravallion, Martin, and Gaurav Datt. 1995. "Is Targeting Through a Work Requirement Efficient? Some Evidence for Rural India." In *Public Spending and the Poor: Theory and Evidence*. Edited by Dominique van de Walle and Kimberly Nead, 413–44. Baltimore: Johns Hopkins University Press.

Ravallion, Martin. 2009. "How Relevant is Targeting to the Success of an Antipoverty Program?" *World Bank Research Observer* 24 (2): 205–31.

Schnitzer, Pascale. 2019. "How to Target Households in Adaptive Social Protection Systems? Evidence from Humanitarian and Development Approaches in Niger." *Journal of Development Studies* 55 (sup1): 75–90.

Stoeffler, Quentin, Bradford Mills, and Carlo del Ninno. 2016. "Reaching the Poor: Cash Transfer Program Targeting in Cameroon." *World Development* 83: 244–63.

Stoeffler, Quentin, Bradford Mills, and Patrick Premand. 2020. "Poor Households' Productive Investments of Cash Transfers: Quasi-Experimental Evidence from Niger." *Journal of African Economies* 29 (1): 63–89.

Wiesmann, D., L. Bassett, T. Benson, and J. Hoddinott. 2009. "Validating the World Food Programme's Food Consumption Score and Alternative Indicators of Household Food Security." Discussion Paper 00870, International Food Policy Research Institute, Washington, DC, USA.

World Bank. 2011. "Niger Safety Net Project, Project Appraisal Document." World Bank, Washington, DC, USA.

———. 2017. "Republic of Niger. Priorities for Ending Poverty and Boosting Shared Prosperity. Systematic Country Diagnostic." World Bank, Washington, DC, USA.

World Food Programme. 2008. "Food Consumption Analysis: Calculation and Use of the Food Consumption Score in Food Consumption and Food Security Analysis." Technical Guidance Sheet, Rome.

Yatchew, Adonis. 1998. "Nonparametric Regression Techniques in Economics." *Journal of Economic Literature* 36 (2): 669–721.