Evaluating alternative targeting of social assistance in fragile settings

Pre-Analysis Plan

December 20, 2023

Abstract

This research project aims to address three research objectives. First, we aim to evaluate the performance of alternative targeting approaches including Proxy Means test (PMT), Community-Based Targeting (CBT) and Peer Targeting (PT) in fragile and conflict settings. We measure their relative performance in terms of identifying the poor. Second, we test the performance of alternative variants of community-based targeting approaches under different selection conditionalities and budget constraints. We evaluate the impact of alternative designs to CBT on the breadth and depth of social assistance transfers distributed to members of the community. Third, we evaluate decision-making and negotiation processes in CBT when real incentives are involved and in the absence of actual cash transfers.

Introduction

Effective targeting of social assistance is critically important to reach vulnerable communities rapidly and cost-effectively. This is particularly important in the context of increasing fragilities and conflicts in Africa where donors and development partners continue to face additional competing priorities making the availability of resources for overseas humanitarian assistance increasingly difficult. Recurring global and regional shocks, including those arising due to conflicts, food price and climate related shocks, are increasing the urgency and need for shock-responsive social assistance and targeting mechanisms. Similarly, reduced availability of resources in the face of shocks means that humanitarian organizations such as World Food Programme (WFP) are forced to follow more stringent targeting approaches and rethink existing ones, including in humanitarian contexts. This adds additional pressure to the growing challenge of targeting humanitarian and social assistance, particularly in conflict-affected settings (Sabates-Wheeler and Szyp, 2022; Maxwell et al., 2010). This challenge arises partly due to the uncertainty surrounding targeting caused by the shocks (as in the case of conflicts changing the status quo) and the urgency to move fast to save lives, which renders targeting either costly, as in universal coverage (Devereux et al., 2017) or it becomes "crude" and hence socially inappropriate (Ellis, 2012).

While a growing literature offers some guidance on alternative targeting approaches in relatively stable contexts, empirical evidence on the relative performance of alternative targeting approaches, especially in conflict-affected and fragile settings, remains scarce. For example, Premand and Schnitzer (2021) find that Proxy Means Targeting (PMT) generated slightly higher impacts of national cash transfer program in Niger than community-based targeting (CBT) while the former also appearing to be more legitimate among non-beneficiaries than CBT, despite the usual lack of transparency PMT based targeting face (Hanna and Olken, 2018). However, in fragile and conflict-affected settings, relevant data are usually not readily available and hence CBT may be preferred over data-driven approaches such as PMT (Alatas et al., 2012; Karlan and Thuysbaert, 2019). These community-driven approaches work best in close-knit communities that know each other well and in slow-onset emergencies with no conflicts or displacements (Maxwell, 2011; Karlan and Thuysbaert, 2019). However, in situations of frequent and prolonged conflicts, community members may lack up-to-date information about each other's livelihood conditions as conflict induced losses may alter pre-conflict status. Trachtman et al. (2022) finds that while community-based targeting can successfully identify long-term poverty, it falls short in identifying short-term and dynamic distress. It is, thus, unclear which alternative targeting approaches to follow in these contexts. Humanitarian agencies, including the WFP, constantly struggle with the complexities and trade-offs involved in the targeting of humanitarian programming, including who to target, how to target, and whether to target at all (Sabates-Wheeler and Szyp, 2022).

Even within the alternative Community-Based Targeting (CBT) approaches, there is a notable lack of empirical evidence on how alternative designs perform and compare in effectiveness. This gap becomes particularly significant in light of the growing demand and diminishing resources for humanitarian and social assistance because of global and regional crises. For example, understanding how alternative CBT approaches perform in the presence of resource constraints can inform targeting operations in various settings. Furthermore, the ideal balance of supervision, monitoring and discretion needed in community-based targeting is unknown. A key inquiry remains: what objectives do community leaders maximize in targeting social assistance programs and when they face resource constraints and other limitations?

In this study, we aim to rigorously evaluate the performance of alternative targeting methods under conflict-affected and fragile setting in Ethiopia. We particularly aim to address the following questions. (1) which targeting approaches can serve the poor and/or most affected in conflict and post-conflict settings? (2) How does community-based targeting (CBT) fare in conflict settings and in response to changes in resource constraints? (3) Does conflict change the performance of the alternative targeting mechanisms? (4) How do community leaders in contexts where protracted crises change the nature and extent of poverty allocate resources availed at their discretion vis-a-vis standard targeting criteria (often set by implementers)? (5) Do changes in resource constraints in community targeting affect incentives for elite capture and induce resource misappropriation? (6) How does real and/or perceived resource misappropriation in community targeting affect trust in local institutions? (7) Do community leaders in conflict-affected and fragile settings behave differently when transfers are hypothetical and real?

To answer these questions, we work with communities and community leaders in 180 villages or Enumeration Areas (EAs) in Ethiopia, leveraging large household and community surveys conducted prior to the outbreak of the recent conflict in Ethiopia. These surveys serve as a baseline for the upcoming community-level experiments. We then conduct follow-up household- and community-level surveys, revisiting the same households and communities with additional enriched instruments that are designed to gather insights on targeting using data-driven methods such as PMT as well as through alternative variants of community-based targeting.

Ethiopia is a particularly suitable setting to study targeting in the context of conflict and fragility. On one hand, the country has unfortunately experienced a surge in conflict in recent years, while on the other hand, in the two relatively stable decades preceding the conflict, it has implemented one of Africa's largest social assistance programs, the Productive Social Safety Net Programme (PSNP). This context provides a unique opportunity to understand targeting dynamics after conflicts set in following a relatively longer-term stability. To do so, we randomly assign communities into treatment and control arms and introduce real cash transfers, in which community leaders are allowed to allocate to beneficiary households, mimicking different forms of CBT for social assistance.

Specifically, we work with community leaders to conduct an incentivized CBT experiment in which leaders conduct ranking exercises of their community members from the most to the least needy for social assistance, and then accordingly make a onetime lump sum budget transfer. This exercise partly mimics the traditional CBT where community leaders' knowledge and information is used to target social assistance (Dupas et al., 2022; Trachtman et al., 2022). To understand how CBT responds to resource scarcity, we exogenously vary budgets available to community leaders to allocate among community members. This allows us to identify (i) whether resource scarcity makes CBT more inclusive or exclusive (both in terms of number of beneficiaries and the size of transfers); (ii) whether the composition of beneficiaries varies by the level of budgets made available to community leaders, and (iii) whether resource leakages increase (decrease) under relaxed (constrained) budget constraints. We also exogenously vary the targeting criteria and level of discretion granted to community leaders. In one treatment arm, community leaders are given a full discretion of targeting beneficiaries using their own targeting criteria, while another arm was instructed to target based on a set of pre-determined targeting criteria. This variation allows us to study how different levels of autonomy given to community leaders affect targeting in CBT. To mimic real world settings and understand associated behaviors, our experimental design allows that a limited portion of the budget may be channeled to cover community leaders' "administration costs" if the community leaders wish to do so.

We also implement peer-ranking at the beneficiary level where respondents are asked to rank approximately 20 survey households within their village (enumeration area) from the “most needy” to the “least needy” for the purpose of targeting for a social assistance, irrespective of households’ current participation status. This follows Dupas et al. (2022), but its application in conflict affected settings offers valuable insights for targeting in those settings. The rankings from the CBT and peer-ranking by households are then compared with the consumption and poverty-based, asset-based and a parametrized PMT-based targeting using pre- and post-conflict data. In addition, we will compare our rankings with actual targeting and selection practices in social protection programs, including the PSNP, to assess how each of these targeting approaches perform in identifying poor households. We also examine the impact of conflict exposure on the performance of the alternative targeting approaches using data from the Armed Conflict Location and Event Data (ACLED) project. By linking these methods’ effectiveness with the ACLED data, we can establish whether conflict moderates the performance of one or more of the targeting methods.

The objective of this research project is to understand alternative targeting approaches for social assistance programs in conflict-affected settings. Over the last decade and half, with support from development partners, Ethiopia has implemented one of largest social assistance programs in Africa - the Productive Safety Net Programme (PSNP). The program reaches about 8 million rural people living in food insecure communities in the country (e.g., Gilligan et al., 2009; Hoddinott et al., 2012; Berhane et al., 2014; Abay et al., 2022). During much of this period, Ethiopia was characterized by relative stability and positive socioeconomic changes. In recent years, the country has been marred by political unrest and recurrent conflicts, leaving millions in dire emergency and social assistance needs. This surge in the numbers of people in need of assistance, coupled with resource constraints faced by international aid agencies and development partners, has necessitated the need to rethink existing targeting approaches, including community-based and other data-driven approaches such as PMT. Our study is motivated by these challenges and aims to contribute to the broader discourse on alternative targeting strategies in various settings, particularly in conflict-affected and fragile contexts.

We work with communities and community leaders in 180 Enumeration Areas (EAs) or villages across Ethiopia. An EA typically comprises 150 to 200 households within a *Kebele*, the lowest administration unit in Ethiopia. In the 2019 survey, a random sample of about 20 households from each of the 180 EAs/villages were interviewed, along with a community survey where information was collected on a range of *kebele* characteristics. The community leaders' experiment and survey brought together six individuals composed of key *Kebele* leaders, including the *Kebele* chairman and other individuals knowledgeable about the village. Again, to mimic actual targeting practices in Ethiopia and other countries, we define the six committee members to include the following members: (i) *Kebele* leader or a member of the kebele leadership, (ii) Elder man/woman, (iii) religious leader, (iv) women representative, (v) teacher or development agent or extension worker, and (vi) youth representative. We note that these members are commonly involved in targeting of social protection programs in Ethiopia, including in the targeting of the PSNP.

Interventions

The interventions in this study occur on two levels: the community leaders and the community members. We describe each of these as follows.

Intervention 1: Ranking of households and allocation of hypothetical and actual cash transfers by community leaders

We ask community leaders to rank households in their respective EAs *from the most to the least needy* based on their needs assessment for social assistance. We provide community leaders with a lump-sum of cash transfers that they will allocate among their ranked households based on community level random assignment into different budget categories and discretion levels built into the experimental design (see Experimental Design section). We exogenously vary the nature of the transfer (hypothetical versus real), the amount of money available for transfers, and the level of discretion granted to community leaders. This would allow us to study, for example, whether discretion in the selection of beneficiaries and allocation of funds makes community-based targeting more Rawlsian - maximizing the welfare and wellbeing of the worse-off households (Rawls, 1999; 2001) or Utilitarian and hence-maximize the sum of the individuals' utilities within the community regardless of differences in individual wellbeing (Arrow, 1973; Yaari, 1981; Mill, 1993).

Intervention 2: Peer- and self-ranking of households

Similar to the community-based ranking, we ask all survey households in each EA (approximately 20) to rank themselves and other fellow households in our sample in the same EA from the most to the least needy. Respondents provide these ranks with a view of targeting for social assistance. We build on the 2019 survey, which contained rich information on pre-conflict status of households including consumption, assets and other important modules that allow us to estimate alternative targeting measures including PMT.

These two interventions allow us to compare the rankings from community leaders and peers against the constructed consumption-, poverty-, asset-and a parametrized PMT-based targeting using pre- and post-conflict data.

Targeting social assistance in conflict-affected settings: Some theoretical foundations

Data-driven approaches for targeting social assistance are impractical in settings of prolonged conflicts and widespread poverty, as up-to-date data necessary to apply such methods is often unavailable. In addition, conflicts disrupt the capacity and workings of otherwise functioning social assistance delivery systems (Sabates-Wheeler et al., 2022). Thus, the burden of targeting under such circumstances relies on the knowledge and judgement of community leaders and others involved in the targeting process. However, conflicts can erode local knowledge as well as alter local caseloads, making community-based targeting challenging. Protracted crises and shocks can significantly alter the poverty landscape – pushing the already poor into further destitution and drifting new ones into transitory poverty. With a level of information asymmetry introduced by conflicts, and resource constraints in the face of increased caseloads, it is unclear how and whom community leaders choose to allocate social assistance that is left at their discretion.

In this study, we ask two important high-level questions (later divided into seven research questions), mainly (1) which targeting methods are most effective in identifying the poor, and how do alternative variants of CBT perform in such information and resource constrained environments; (ii) what objectives do community leaders maximize when targeting social assistance in various contexts and in the presence of resource constraints. For example, do community leaders maximize the number of beneficiaries or the amount of transfer that goes to specific types of (e.g., the poorest of the poor) households, and does this change with conflict? We also assess whether the resulting targeting outcomes broadly conform to choices reflected by the conventional measures of welfare and hence promote social cohesion. While most of these are empirical questions whose answers likely differ across contexts, they can be linked to existing theoretical foundations that are commonly used to explain the choice of targeting of social assistance. In particular, there are three important theories widely used to justify targeting and redistribution of social assistance in various systems, namely the Rawlsian theory of social welfare (Rawls, 1999; Metz et al., 2002), the Utilitarianism (Arrow, 1973; Yaari, 1981; Mill, 1993), and the universal basic income (UBI) (e.g., Stark et al., 2014;Van Parijs and Vanderborght, 2017; Banerjee et al., 2019; Ghatak and Maniquet, 2019); all of which merit some discussion to facilitate our understanding.

Rawls' social justice theory prioritizes fairness and equality in the distribution of social assistance (Rawls, 1999; Metz et al., 2002). According to Rawls, justice requires that resources are distributed in a way that benefits the least advantaged members of society (Metz et al., 2002). Rawls' theory thus provides a framework for considering the needs of these individuals in the allocation of social assistance (Brewer et al., 2014). In the context of conflict-affected settings, Rawls' theory of justice implies the allocation of social assistance should emphasize the need to address the inequities and disadvantages faced by individuals in these settings (Barrientos, 2016). Conflict often exacerbates existing inequalities and creates new vulnerabilities, making it crucial to prioritize the well-being of the most disadvantaged individuals (Campbell, 2021).

Rawls' theory stands in contrast to utilitarianism, which focuses on maximizing overall happiness or utility, without necessarily considering the distribution of resources. The utilitarian approach may lead to the neglect of those who are most in need of social assistance in conflict-affected settings (Sari, 2020). By prioritizing the needs of the least advantaged members of society, Rawls' theory, on the other hand, helps to address the underlying causes of conflict and promote social cohesion (Sen, 2006). Stark et al. (2014) identify the conditions under which the utilitarian approach and the Rawlsian approach may generate similar outcomes. For example, Stark et al. (2014) demonstrate the congruence of the utilitarian and Rawlsian approaches when utility of individuals depends not only on an individual's own income, but also on others' income.

The Universal Basic Income (UBI) approach to targeting social welfare in contrast relies on alternative normative notions and values to justify redistribution policies (Van Parijs and Vanderborght, 2017; Ghatak and Maniquet, 2019). Furthermore, Rawls' theory supports the idea of a universal basic income (UBI) as a means of addressing social and economic inequalities (Metz et al., 2002). In the context of conflict-affected settings, UBI would be a form of social assistance that provides a regular income to all individuals affected. UBI aims to ensure a basic level of economic security and reduce poverty. By guaranteeing a minimum income for all individuals, UBI aligns with Rawls' principle of justice as fairness, as it seeks to address the disadvantages faced by the least advantaged members of society (Metz et al., 2002).

However, it is important to note that the implementation of Rawls' theory and UBI in conflict-affected settings may face challenges. For example, community leaders may give more weight to exclusion than inclusion errors partly because targeting becomes practically infeasible making exclusion errors costly with the resulting outcome that everyone in the community gets social assistance, as in the case of damages that affect everyone in the community where excluding anyone does not make sense (Sabates-Wheeler and Szyp, 2022). On the contrary, providing social assistance to certain individuals may bring about significant societal benefits than thinly spreading resources based on need. The allocation of social assistance in these settings is often complex and depends on context, including society's views on social justice and community leaders' objective function, resource constraints and other competing priorities (Wahab & Khairi, 2020, Campbell, 2021). Additionally, the effectiveness of UBI in addressing the specific needs and vulnerabilities of conflict-affected populations requires careful consideration and adaptation to the unique contexts (Tyler, 2000).

Depending on the objective functions pursued by community leaders, each of these theories provides a framework for understanding the allocation of social assistance in conflict settings. Rawls' theory would emphasize the importance of fairness, equality, and addressing the needs of the least advantaged individuals. The concept of UBI aligns with Rawls' principles and can be seen as a means of addressing inequalities. Utilitarianism would imply a distribution that maximizes overall utility regardless of inclusion or exclusion errors based on neediness. Clearly, which of these approaches is pursued in practice remains an important hypothesis worth testing in our study.

Research Objectives and Research Questions

This research project aims to address three broad research objectives. The first objective of the project is to evaluate the performance of alternative targeting approaches including proxy means test (PMT), community-based and peer targeting in fragile and conflict settings. We measure their relative performance in terms of identifying the poor. The second objective revolves around testing the performance of alternative variants of community-based targeting approaches under different settings and budget constraints. We evaluate the impact of alternative designs to CBT on the breadth and depth of social assistance transfers distributed to members of the community. Third, we evaluate decision-making and negotiation processes in CBT when real incentives are involved and in the absence of actual cash transfers. Below, is the list of specific research questions associated with these broad research objectives:

1) How do alternative targeting approaches perform (in terms of identifying the poor) under conflict settings?
    a. Community-based targeting versus Proxy Means Test (PMT)
    b. Community-based targeting versus peer targeting
    c. Peer ranking versus PMT
2) Do community leaders behave differently (in terms of allocating social assistance transfers) when transfers are real versus hypothetical?
3) How do community leaders target and allocate resources when faced with constrained versus relaxed budgets? Do they maximize the number of beneficiaries or the size of transfer to beneficiaries?
4) How do community leaders target and allocate resources under rule-based versus discretionary settings?
5) How does conflict mediate targeting under resource constraint?
6) Do alternative community-based targeting approaches affect negotiation and pro-social behavior of community leaders?

The alternative targeting methods we implement along with the exogenous variations we introduce in the different variants of CBT enable us to address the above research questions. For example, as we are implementing three types of targeting: PMT, CBT and PT, we can compare their performance in terms of identifying the poor and the type of beneficiary households each targeting method target. The exogenous variation in exposure to hypothetical versus real cash transfer allows us to assess and evaluate community-leaders' decision-making and negotiation behavior in hypothetical and incentivized experiments. These behavioral decision-making processes can be captured through the amounts (or share) of resources they allocate to themselves and to community members. Similarly, the exogenous variations in the amount of transfer and level of discretion given to community leaders enables us to quantify differences in resource allocations across these settings.

Primary outcomes

1. Households' access to CBT transfers (whether the household receives a transfer or not)
2. Amount of transfer received by households.
3. Amount of transfer devoted to running costs or administrative costs by community leaders
4. Distribution of transfers or inequality in the distribution of transfers

Primary outcomes explanation

1. Households' access to CBT transfers (whether the household receives a transfer or not). The first outcome we consider is a binary outcome capturing whether a household receives transfer from community leaders from the pool of funds community leaders are given to disburse to households. The number of households or the share of households receiving transfer from the community leaders is one of the key measures we employ to capture how community-based targeting functions under different scenarios and settings. Financial constraints and discretion may encourage community leaders to distribute social assistance differently. Exposure to conflict may also affect how community leaders target beneficiaries.
2. The amount of transfer households receive from CBT transfers. This slightly relates to the first outcome and hence measures the amount of transfer households receive from community leaders. This considers the intensive margin of participation in social assistance. The amount of transfer each household receives is related to the number of households selected to receive cash transfers. Community leaders may maximize the total

number of households covered by a social assistance program or the amount of transfer each household receives. Thus, the first and second outcomes will be analyzed jointly.

3. The share of resources used to cover administrative costs of community leaders. Community leaders were offered the option of keeping up to 10 percent of the budget allocated for the community as remuneration for their time. First, each community leader will be asked to decide independently how much he/she would like to keep as a "running cost" for the group's operations. The choice set contains the following options: 0 percent, 2 percent, 4 percent, 6 percent, 8 percent, and 10 percent. Then the full group of community leaders will decide on a negotiated amount to keep for themselves. This allows us to study how alternative designs of CBT affect negotiation behavior and whether negotiated decisions generate better pro-social outcome for the community compared to individual decisions.
4. Distribution of transfers or inequality in the distribution of transfers, measured by Gini coefficient. This outcome uses Gini coefficient based on the amounts of transfers received among members of the community. This measure allows us to identify whether some features and designs associated with CBT make community leaders inequality averse or not.

Secondary outcomes

1. The distribution of transfers across beneficiaries
2. The number of beneficiaries in each village
3. Average transfer received by households

Experimental design

In this section, we outline the experimental design of the intervention at the community leaders' level. As discussed above, community leaders of the 180 communities are the key actors in this intervention. The intervention follows community level clustered randomization in which the 180 communities are randomly assigned into one of four groups/ treatment arms. The treatment assignment is based on (i) whether communities receive actual transfer or hypothetical (control), (ii) the nature of discretion given to community leaders when allocating transfers (full discretion versus rule-based) and (iii) the size of the transfer pool available to community leaders to distributed among households within the community (constrained budget involving 10,000 Birr versus relaxed budget of 20,000 Birr). The treatment arms generated by combinations of these treatments are outlined below (see also Figure 1).

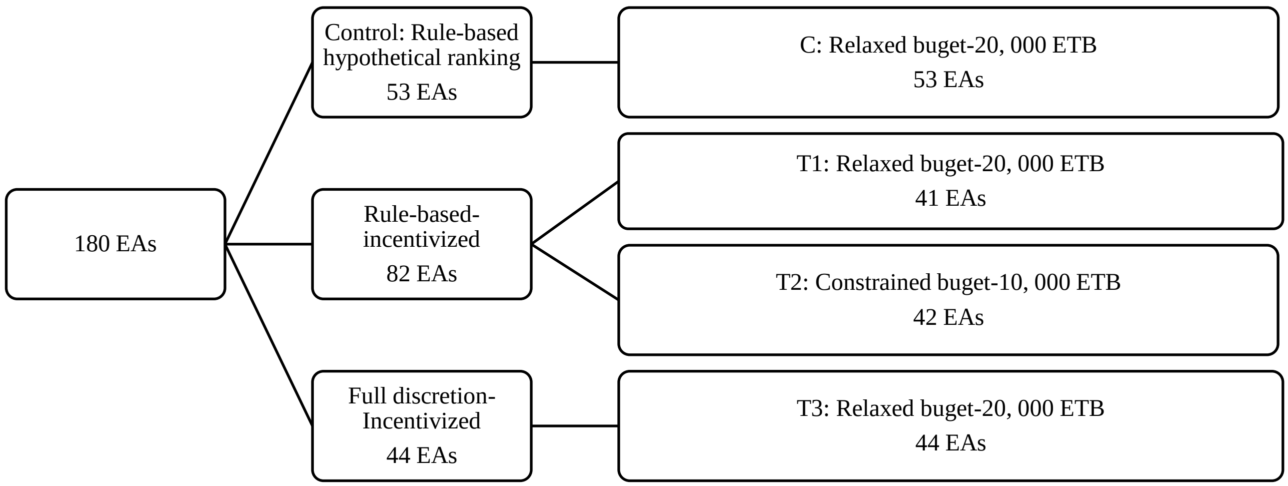

Figure 1. Random assignment of communities across treatment and control arms

(1) ***Control: Rule-based targeting using hypothetical transfer of 20,000 Birr (C)***: This group serves as a control cluster where community leaders are not given any actual funds but are instructed to act as if they have a hypothetical budget of 20,000 Birr to distribute among households in their community. Community leaders are first asked to rank households based on their need for social assistance. They are then asked to allocate this notional budget among the 20 households included in our sample. During this ranking process, leaders are required to strictly adhere to pre-defined rules provided by the research team. These rules are carefully selected to mimic the targeting criteria used in actual social assistance programs in Ethiopia. More specifically, community leaders are asked to prioritize those households who: (i) had difficulty satisfying their food needs; (ii) own no or little asset (e.g., livestock, land); (iii) have limited income-generating activities or capacity; (iv) have lost productive assets due to shocks (e.g., conflict, drought); and (v) have lost family members recently.

(2) ***Rule-based incentivized targeting with relaxed budget (T1)***: Another group of communities are randomly assigned to a cluster that receives real transfer funds with a budget of 20,000 Birr (about 360 USD). In this cluster (i) Community leaders are required to rank households based on five pre-determined targeting criteria and allocate the transfers. These criteria are similar to those in the control group and mimic the targeting criteria used by the national safety net program in Ethiopia, the PSNP (e.g., Gilligan et al., 2009; Hoddinott et al., 2012; Abay et al., 2022).

(3) ***Rule-based incentivized targeting with constrained budget (T3)***: This group of communities follows similar rules as those in control group, but they receive a constrained budget of 10,000 Birr (about 180 USD). Community leaders are required to rank households based on the five criteria outlined above and allocate the 10, 000 Birr to the community members in our sample. These criteria are designed to mimic the targeting criteria used by the national safety net program in Ethiopia, the PSNP. This treatment arm serves to test the implication of budget constraint.

(4) ***Incentivized-discretionary targeting (T4)***: The fourth group of communities are provided a budget of 20,000 Birr to distribute as social assistance to households identified as in need. Here, community leaders rank households based on their own criteria they collectively agree upon. The establishment of these ranking criteria is entirely left to the discretion of the community leaders. It is up to the leaders to determine who among the ranked households gets how much of the 20,000 Birr transfer assigned to the community.

Randomization Method

The randomization is done at the village (EA) level using the baseline list of villages. The initial selection of villages into our sample considers their accessibility for a survey. The selected and accessible villages are then randomly assigned into four groups. A reserve list has been prepared in case some of the villages become inaccessible due to conflict.

Randomization unit

Village or community level

Was the treatment clustered

Yes

Planned number of observations

180 communities and about 3,000 households

Number of clusters by treatment arm

T1 (Control): 53 villages

T2 (Rule-based, 20, 000 ETB): 41 villages

T3 (Rule-based, 10, 000 ETB): 42 villages

T4 (Discretionary, 20, 000 ETB) : 44 villages

Statistical power: minimum detectable effect size for primary outcomes

As this project is designed to test multiple hypotheses, the power calculation considers these hypotheses and respective outcomes. We compute the number of clusters needed for the primary outcomes described above, assuming that there are a known and fixed number of households in each cluster (village). In the baseline sample there were an average of 20 households in each village, and we anticipate being able to trace about 85 percent of them, 17 households per village. Our power calculations aim to achieve the standard and widely adopted 80 percent power at a significance level of 5 percent. We note that the power calculations are computed and reported only for the primary outcomes. Given that we have several hypotheses and primary outcomes, we computed the number of clusters and associated sample size needed for each primary outcome separately, and then selected the maximum sample needed to detect impacts across these outcomes.

To compare the performance of alternative variants of community-based targeting, we compute statistical power and sample size needed to detect a reasonable impact on households' access to CBT transfers (whether the household receives a transfer), the amount of transfer received by households as well as the share of budget devoted as "running cost". The rule-based community-based targeting we adopt in this study mimics the targeting approach followed by Ethiopia's flagship national social safety program, the PSNP. This allows us to exploit important information on the distribution of PSNP participation in our baseline data.

Our design introduces three important variations that generate three hypotheses related to these changes to the usual rule-based CBT: (i) what happens when real stakes and payments are introduced; (ii) how do community leaders prioritize allocation when they face resource constraints, and (iii) does granting community leaders more discretion on the criteria to be used for ranking and distributing the transfers matter. To test these hypotheses, we employ alternative indicators that capture the decision-making processes and hence distribution of the cash transfers offered by community leaders.

About 30 percent of households in our sample are PSNP beneficiaries at baseline, and we assume that in the absence of any additional intervention community leaders could choose 30 percent of households to receive for transfer distribution. Furthermore, we hypothesize that reducing the available resources by 100 percent to 10,000 Birr (about 180 USD) can reduce the number of beneficiary households by 17 percentage points, and also reduce the transfer amount that goes to beneficiary households by 35 percent. To detect these impacts, we need 41 clusters in each treatment arm.

Community leaders are expected to behave differently in hypothetical versus real transfer scenarios, both in terms of the share of the total budget they take out for "running costs" as well as the way they allocate resources to potential beneficiaries. For instance, community leaders may behave more pro-socially and hence devote smaller resources for themselves in hypothetical scenarios. We assume that in the hypothetical setting, about half of the community leaders might request the maximum amount, while this is likely to increase by 23 percentage points when there are real incentives and transfers. Detecting this impact requires about 50 communities in the control group and roughly three times that in the treatment group. Based on these calculations, we allocate about 30 percent of the 180 communities into the hypothetical arm with no actual transfer and divide the remaining communities into three equal groups. We stratify the random assignment of communities across regions.

Empirical Estimation Strategy

Evaluating targeting performance under various settings

We start by implementing an empirical specification that allows us to evaluate and compare the performance of alternative targeting approaches in various settings. Before estimating parametric and more saturated specifications, we estimate the concordance between the different rankings using a non-parametric statistic, Spearman's rank correlation coefficient. This helps to establish whether alternative targeting approaches generate significantly correlated rankings. We then assess the performance of the different targeting approaches (CBT, peer ranking and PMT) in accurately identifying the poor and most needy. We do so by comparing the rankings from the three targeting approaches with a poverty-based targeting (e.g., Alatas et al., 2012; Trachtman et al., 2022; Dupas et al., 2022). For testing these hypotheses, we run the following simple empirical specifications:

$$P_{hc} = \alpha_{m1}R_{hcm} + \alpha_{m2}HH_{hc} + \alpha_{m3}CM_c + \alpha_c + \epsilon_{hc} \quad (1)$$

$$R_{hcm} = \beta_{m1}R_{hc(-m)} + \beta_{m2}HH_{hc} + \beta_{m3}CM_c + \beta_c + \varepsilon_{hcm} \quad (2)$$

where P_{hc} stands for standard and widely used welfare and consumption-based poverty measures for household h in community c. R_{hcm} stands for ranking of household h living in community c using targeting method *m*. These targeting methods include: PMT, CBT and peer ranking or targeting. HH_{hc} captures household level characteristics, which includes both demographic as well as socioeconomic characteristics of households, and CM_c represents community level characteristics. α_c and β_c represent either village or district level dummies while ϵ_{hc} and ε_{hcm} capture additional unobservable factors that may explain household welfare or the ranking of households.

The empirical specification in equation (1) allows us to address the following important questions: (i) which one of the alternative targeting approaches predict poverty and hence identify the poor better; (ii) how correlated are alternative targeting methods to social assistance; and (iii) what explains the variation in ranking/targeting across alternative approaches as well as across households: household or community characteristics? These are important research questions that can inform targeting operations in humanitarian and social protection programs. Comparing the sizes of α_{m1} across alternative targeting approaches can inform the relative performance of each targeting method in identifying the poor. We note that one can also compute targeting errors associated with each method by applying poverty lines and dividing the sample into poor and non-poor as in Alatas et al. (2012). In our case, however, we use continuous measures of per capita consumption, mainly because in a population with a large share of poor households, a continuous measure of deprivation is more informative than a binary indicator.

The vector of parameters contained in β_{m1} provides important insights on the strength of the correlation (in household ranking) across the alternative targeting methods. Strong correlations across these methods suggest that governments and donors may apply either of these methods to identify the needy without compromising targeting accuracy. The vector of parameters captured by α_{m3} and α_{m4} as well as β_{m3} and β_{m4} helps us to identify the role of household and community-level attributes in the ranking of households. For example, if household characteristics appear to be negligible, governments and donors may save resources by optimizing the efficiency of spatial targeting instead. In situations where poverty is highly concentrated and hence within community targeting of households is likely to generate an inaccurate list, spatial targeting may suffice and reduce operational costs (e.g., Dupas et al., 2022). To quantify the role of household and community characteristics, we compute and report the Shapley decomposition of the explained variation (measured by R^2) in ranking across households and alternative targeting methods (Huettner and Sunder, 2012).

Besides assessing the performance and strength of the correlations across alternative targeting and ranking approaches, we also explore how these targeting approaches perform across the following settings and situations: (i) in the presence of violent conflicts that may disrupt information flows while also increasing the vulnerability of communities; (ii) when communities are characterized by high (low) level of inequality; (iii) when there is information asymmetry and lack of interaction across peers; (iv) in dynamic settings that involve significant temporal and spatial mobility across wealth quintiles; and (v) in densely populated urban areas with limited social network and sparsely populated rural areas.

To test these hypotheses, we extend the empirical specifications in equations (1) and (2) by incorporating interaction terms. For example, if conflict impairs the informational advantage that community leaders or peers may have, particularly because prolonged conflicts could distort community member's knowledge of each other's livelihood conditions, community-based and peer-ranking may perform worse in conflict-affected settings. To test this, we compile data on household and community-level exposure to conflict from the Armed Conflict Location and Event Data (ACLED) project and interact these variables with the specific type of targeting approach. Similarly, we will compute community-level inequality indicator using the Gini coefficient and interact it with the rankings generated by the alternative targeting methods. As we are collecting pre-conflict and post-conflict data on consumption and assets, we also aim to identify which targeting methods can sufficiently capture dynamic evolution of household welfare and wealth profiles. Using detailed information about the acquaintances and relationships among peers and community members, we identify whether information asymmetries and the nature of social networks in a community affect the performance of peer and community-based targeting. Finally, we have baseline and follow-up data on various demographic characteristics of households and communities, which enables us to explore additional heterogeneities in the performance of the alternative targeting methods.

Comparing the performance of alternative variants of community-based targeting

To compare the performance of alternative variants of CBT, we will exploit the random variation generated by the intervention that assigned villages and hence community-leaders into a control and three treatment arms. The random assignment of communities is done based on (i) whether community-leaders receive hypothetical or actual transfer funds to distribute among households (i.e., hypothetical ranking/distribution exercise or incentivized ranking exercise); (ii) the nature of discretion granted to community leaders when ranking households and allocating transfers (i.e., rule-based versus full discretion); and (iii) the level of transfers community leaders are given to distribute to households (i.e., constrained, or birr 10, 000 ETB versus 20,000 ETB).

The randomly generated treatment variations produce the following treatment arms: (i) *Control: Rule-based hypothetical ranking/distribution (C)*: a group that serves as control and where community leaders receive no real transfer but are asked to hypothetically imagine a 20,000 Birr transfer fund while ranking households in these communities. However, leaders are required to strictly follow pre-defined rules in ranking households. (ii) *Rule-based incentivized targeting with relaxed budget (T1)*: a group where community leaders receive a real 20,000 Birr transfer budget but are asked to rank households and distribute transfers based on pre-defined criteria. (iii) *Rule-based incentivized targeting with constrained budget (T2)*: a group where community leaders receive a 10,000 Birr transfer fund and are asked to rank households and distribute transfers based on pre-defined criteria. This group is similar to T1, with the only difference being the smaller transfer budget. (iii) *Discretionary Incentivized targeting (T3)*: this group receives a one-off 20,000 Birr transfer as social assistance to help needy households based on some criteria that is left to the discretion of the community leaders. To check the validity of the randomization we will conduct balance test of observable baseline characteristics of communities and households across these arms. We also use these observable demographic and socioeconomic characteristics to explore potential heterogeneities in the impact of the interventions.

As described earlier, we employ two types of outcomes and methods to compare and evaluate how alternative variants of community-based targeting perform. The first group of outcomes are measured at the household level and relate to the extent and intensity of the community-based transfers. These outcomes include the number of beneficiary households, whether a household receives a transfer, and the amount of transfer going to each household. The second group of outcomes are measured at the community-level and relate to the distribution or concentration of the transfers across the entire community as well as across the poor members of the community. We examine the first set of outcomes parametrically while we assess the distribution of transfers using non-parametric approaches using Kolmogorov-Smirnov tests.

While the random assignment of villages into treatment and control groups generates unbiased average treatment effects using simple mean differences, the availability of observable characteristics facilitates more structured and powered estimations. Hence, we will estimate the following empirical specification to identify the impact of the various treatment arms:

$$Y_{hc} = \beta_0 + \beta_1 T1_{hc} + \beta_2 T2_{hc} + \beta_3 T3_{hc} + \beta_4 HH_{hc} + \epsilon_{hc} \tag{3}$$

Where Y_{hc} measures access to community-based cash transfers associated with each household h living in community *c*. We measure households' access to community-based cash transfers at the extensive margin using a binary indicator for access to transfers and at the intensive margin using the amount of transfers. $T1_{hc}$, $T2_{hc}$, and $T3_{hc}$ stand for indicator variables for those households assigned to the incentivized rule-based with relaxed budget, incentivized rule-based with constrained budget and discretionary community-based targeting, respectively. Note that those communities and households assigned to the hypothetical ranking/ distribution serve as control group and are the base outcome. β_1, β_2 and β_3 capture the impact of incentivizing the community-based targeting under rule-based and discretionary approaches, respectively. Successful randomization ensures that these parameters are unbiased and hence capture average treatment effect of the incentives under different settings and arrangements. If incentives make community-leaders apply stricter inclusion standards (e.g., devote additional time to carefully identify the neediest households), we expect β_1, β_2 and β_3 to be negative and statistically significant. We hypothesize that incentivizing the targeting exercise could encourage community leaders allocate the maximum amount of the budget for "administrative" purposes, which is at 10 percent of the budget. This would leave less funds to distribute to households.

Although the parameters we estimated thus far are important, the empirical specification in equation (3) does not directly answer how community leaders and community-based targeting function in the presence of binding resource constraints. To explicitly address these questions, we restrict the sample to those treatment arms receiving actual transfers and estimate a modified version of equation (3). That is, we modify the empirical specification in equation (3) by restricting the sample to households assigned to treatment arms (T1, T2 and T3), but with varying budgets and level of discretion to community leaders:

$$Y_{hc} = \alpha_0 + \alpha_1 T2_{hc} + \alpha_2 T3_{hc} + \alpha_3 HH_{hc} + \alpha_4 CM_c + \vartheta_{hc} \quad (4)$$

where all terms are as defined before. Equation (4) compares the targeting outcomes of households exposed to the rule-based targeting with relaxed budget ($T1_{hc}$) with that of households in the rule-based targeting with constrained budget ($T2_{hc}$) and discretionary targeting ($T3_{hc}$). The coefficient on $T2_{hc}$, α_1, captures the impact of budget constraints on household targeting outcomes. Depending on whether community leaders maximize access to the program or size of transfer, we expect either a negative or statistically insignificant α_1. The impact of variation in community leaders' discretion on targeting outcomes is captured by α_2.

We also estimate the impacts of the alternative designs associated with CBT on the community-level outcomes discussed earlier. For this purpose, we estimate the following empirical specification at the community level:

$$U_c = \gamma_0 + \gamma_1 T1_c + \gamma_2 T2_c + \gamma_3 T3_c + \gamma_4 CM_c + \epsilon_c \quad (5)$$

where U_c stands for community-level outcomes, including the share of resource community leaders allocate for themselves and the Gini coefficient associated with the transfers within communities. The impacts of incentivizing the community-based targeting using varying levels of budget and discretion to community leaders are captured by γ_1, γ_2, and γ_3. We hypothesize that incentivizing the targeting exercise could encourage community leaders reduce their pro-social behaviour and hence allocate a larger share of the budget for themselves, especially under a relaxed budget. Thus, we expect γ_1, γ_2, and γ_3 to assume positive values, and we anticipate that $\gamma_1 > \gamma_2$.

Another important policy question is which of the CBT targeting approaches are pro-poor? This is a crucial question directly related to the targeting error involved in each targeting method (e.g., . As we have consumption-based poverty measure, we can directly test this by interacting the treatment indicators specified in equation (3) as follows:

$$Y_{hc} = \delta_1 Poor_{hc} + \delta_2 T1_{hc} + \delta_2 T2_{hc} + \delta_3 T3_{hc} + \delta_4 T1_{hc} * Poor_{hc} + \delta_5 T2_{hc} * Poor_{hc} + \delta_6 T3_{hc} * Poor_{hc} + \delta_7 HH_{hc} + \mu_{hc} \tag{6}$$

where all terms except $Poor_{hc}$ are as defined before. $Poor_{hc}$ is an indicator variable assuming a value of 1 for those households whose per capita consumption falls below the national poverty line. The coefficients associated with the interaction terms between the alternative CBT and poverty status (δ_4, δ_5, and δ_6) allow us to test whether some of the targeting approaches are particularly more effective in serving the poor. The empirical specification in equation (6) can also provide important insights on whether and which of these targeting mechanisms are Rawlsian and which are utilitarian. For example, if either of the targeting approaches are pro-poor and transfers a large share of the budget to poor households, we expect δ_4, δ_5 and δ_6 to be positive and statistically significant.

Finally, our design allows us to probe what type of objective functions community-leaders or peers maximize. We will collect three sets of information that may inform this: (i) for all households they rank we ask respondents to identify the three most important reasons that prompted such ranking, (ii) in the rule-based ranking we ask community leaders whether our criteria are comprehensive enough, and (iii) in the discretionary arm we elicit the criteria agreed upon by community-leaders. Comparing these responses across treatment arms can help us to identify the objective functions community leaders maximize.

Households living in the same community are assigned to the same treatment group and they are likely to face similar shocks, market conditions and food security environment, which could generate spatial correlation of unobserved effects (error terms) across households within the same community. To account for this, standard errors will be clustered at the community level, which is the level of treatment in our case and, thus, the recommended level of clustering for standard errors (Abadie et al., 2023).

Theory of Change

A series of conflicts have hit Ethiopia in recent years, substantially increasing the number of people in dire need of social and humanitarian assistance. Increased caseloads in the face of severe resource constraints means that the Government of Ethiopia (GoE) and a consortium of donors including WFP are seeking for alternative targeting criteria that can fit the context. Several targeting methods are deployed in various contexts, including (1) Community-based targeting (CBT) that involves a group of community representatives who use their local knowledge to inform decision-making regarding who benefits from social assistance; (2) Proxy Means Testing (PMT) that uses observable characteristics to construct a score that proxies the poverty or wealth status of beneficiaries; and (3) Universal targeting, which is often associated with the difficulty to implement other targeting approaches; and instead, simply follows blanket distribution of benefits to all community members either to avoid risk of exclusion errors (of potential beneficiaries) or simply based on the principle of equality. This latter approach is often applied in humanitarian contexts in which applying a method of discrimination among community members is nearly impossible and instead universal coverage is preferred.

While the latter two methods are straight forward in terms of implementation and expected outcomes, it is not clear how CBT works under different scenarios of resource constraints as well as level of flexibility in the use of targeting criteria. CBT has been at the center of Ethiopia's flagship social protection program, the productive safety net program (PSNP) implemented in the last decade and half – a period in which Ethiopia enjoyed significant growth and relative stability. With increased caseloads and expanded geographic area to cover due to conflicts, it is unclear whether CBT still fit for purpose given the new circumstances in Ethiopia.

Our experiment aims to answer these questions by artificially introducing variations in resource constraints and level of discretion and evaluating how CBT performs under these conditions compared to other alternative methods. Specifically, to understand how CBT responds to resource scarcity and the degree of discretion in the targeting and allocation of social assistance resources, we exogenously introduce: (i) hypothetical and real transfer of resources; (ii) two levels of cash transfer funds (i.e., a low and high stake), and (iii) two levels of degree of discretion in the targeting and allocation of transfers – the no discretion context mimics the criteria-based community targeting approach often pursued in poor rural environments as in Ethiopia's PSNP.

Our experimental design allows us to explore whether resource constrained social assistance (low versus high stakes) combined with full discretion (or no discretion) of community leaders results in CBT targeting outcomes to be more Rawlsian, Utilitarian, or otherwise. Furthermore, we hypothesize that incentivizing the targeting exercise could encourage community leaders allocate the maximum allowed amount of the budget for "administrative" purposes, which stands at 10 percent in our design, which could leave less amount of the budget to distribute to beneficiaries. The degree of this leakage itself may be a function of the local context such as poverty, inequality, recent experiences of shocks. The alternative targeting methods are interacted with these measures of local context to prob their impact. Figure 2 illustrates the theory of change of use of alternative targeting methods in our context. All available targeting methods are weighed in view of the prevailing circumstances and the most effective and efficient method involves one that fits-the-context, and is broadly consistent with local choices, minimizes the overall implementation cost and delays.

Figure 2. The theory of change for alternative targeting methods

References

Abadie, A., Athey, S., Imbens, G. W., Wooldridge, J. M. (2023). When should you adjust standard errors for clustering?. *The Quarterly Journal of Economics*, 138(1), 1-35.

Abay, K. A., Abay, M. H., Berhane, G., & Chamberlin, J. (2022). Social protection and resilience: The case of the productive safety net program in Ethiopia. *Food Policy*, 112, 102367.

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., & Tobias, J. (2012). Targeting the poor: evidence from a field experiment in Indonesia. *American Economic Review*, *102*(4), 1206-1240. https://doi.org/10.1257/aer.102.4.1206

Banerjee AV, Niehaus P, Suri T. 2019. Universal basic income in the developing world. *Annual Review of Economics* 11:961–85.

Barrientos, A. (2016). Justice-based social assistance. *Global Social Policy*, 16(2), 151-165. https://doi.org/10.1177/1468018116633576

Berhane, G., Gilligan, D.O., Hoddinott, J., Kumar, N., Taffesse, A.S., 2014. Can Social Protection Work in Africa? The Impact of Ethiopia's Productive Safety Net Programme. *Economic Development and Cultural Change* 63 (1), 1–26.

Devereux, S., Masset, E., Sabates-Wheeler, R., Samson, M., Rivas, A. M., & te Lintelo, D. (2017). The targeting effectiveness of social transfers. *Journal of Development Effectiveness*, *9*(2), 162-211. https://doi.org/10.1080/19439342.2017.1305981

Dupas, P., Fafchamps, M., & Houeix, D. (2022). *Measuring relative poverty through peer rankings: Evidence from Côte d'Ivoire* (No. w29911). National Bureau of Economic Research.

Ghatak, M., & Maniquet, F. (2019). Universal basic income: some theoretical aspects. Annual Review of Economics, 11, 895-928.

Gilligan, D.O., Hoddinott, J., Taffesse, A.S., 2009. The Impact of Ethiopia's Productive Safety Net Program and its Linkages. *Journal of Development Studies* 45 (10), 1684–1706.

Hoddinott, J., Berhane, G., Gilligan, D.O., Kumar, N., Taffesse, A.S., 2012. The Impact of Ethiopia's Productive Safety Net Programme and Related Transfers on Agricultural Productivity. *Journal of African Economies* 21 (5), 761–786.

Karlan, D., Thuysbaert, B., 2019. Targeting ultra-poor households in Honduras and Peru. The *World Bank Economic Review*, 33(1), 63-94.

Maxwell, D., Young, H., Jaspars, S., Frize, J. and Burns, J., 2011. Targeting and distribution in complex emergencies: Participatory management of humanitarian food assistance. *Food Policy*, *36*(4), pp.535-543.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of development Economics*, *99*(2), 210-221.

Metz, T., Rawls, J., & Kelly, E. (2002). Justice as fairness: a restatement. The Philosophical Review, 111(4), 618. https://doi.org/10.2307/1556442

Mill, John S. and Geraint Williams. *Utilitarianism; On Liberty; Considerations on Representative Government; Remarks on Bentham's Philosophy*. London: Dent. 1993.

Rawls, J. (2001). *Justice as fairness: A restatement*. Harvard University Press. Cambridge, MA.

Rawls, J. (1999). *Collected papers*. Harvard University Press. Cambridge, MA.

Sabates-Wheeler, R., & Szyp, C. (2022). Key Considerations for Targeting Social Assistance in Situations of Protracted Crises. BASIC Research Working Paper 12, Brighton: Institute of Development Studies. https://doi.org/10.19088/BASIC.2022.012

Sari, C. (2020). Rawls's theory of justice and its relevance in analyzing injustice on ethnic phenomenon. Journal of Ethnic and Cultural Studies, 210-219. https://doi.org/10.29333/ejecs/497

Sen, A. (2006). What do we want from a theory of justice?. *The Journal of Philosophy*, 103(5), 215-238. https://doi.org/10.5840/jphil2006103517

Stark, O., Jakubek, M. and Falniowski, F., 2014. Reconciling the Rawlsian and the utilitarian approaches to the maximization of social welfare. *Economics Letters*, *122*(3), pp.439-444.

Trachtman, C., Permana, Y., & Sahadewo, G. (2022). How much do our neighbors really know? The limits of community-based targeting. *University of California, Berkeley Working Paper*.

Wahab, A. and Khairi, A. (2020). Right to justice and legal aid barriers to the vulnerable non-citizens in Malaysia. Malaysian Journal of Society and Space, 16(1). https://doi.org/10.17576/geo-2020-1601-02

Van Parijs P, Vanderborght Y. 2017. Basic Income: A Radical Proposal for a Free Society and a Sane Economy. Cambridge,MA: Harvard Univ. Press