How to Target the Poor: Evidence from a Field Experiment in Indonesia[1]

Vivi Alatas, World Bank
Abhijit Banerjee, MIT
Rema Hanna, Harvard Kennedy School
Benjamin A. Olken, MIT
Julia Tobias, Stanford University

First Draft: September 2009

PRELIMINARY DRAFT:
PLEASE DO NOT QUOTE WITHOUT PERMISSION

ABSTRACT

In developing countries, identifying the poor for the purposes of redistribution or social insurance programs is challenging because the government lacks reliable information about people's incomes. This paper reports on a field experiment conducted in 640 Indonesian villages designed to investigate two main approaches to solving this problem: proxy-means tests, in which the government conducts a census of hard-to-hide assets and uses these data to predict consumption, and community-based targeting, in which villagers are asked to rank everyone from richest to poorest. When poverty is defined using per-capita expenditure and the common PPP$2 per day threshold, we find that community-based targeting performs worse in identifying the poor than proxy-means tests, particularly near the threshold. This worse performance does not appear to be due to elite capture. Instead, communities appear to be systematically using a different concept of poverty: the results of community-based methods are more correlated with how individual community members rank each other and with villagers' self-assessments of their own status. Consistent with this, community-based methods result in higher satisfaction with beneficiary lists and the targeting process.

[1] We thank Ritwik Sarkar, Prani Sastiono, Ririn Purnamasari, Hendratno Tuhiman, Matthew Wai-Poi, and Chaeruddin Kodir for outstanding research assistance and thank Mitra Samya and the Indonesian Central Bureau of Statistics for their cooperation implementing the project. Most of all we thank Lina Marliani for her exceptional work leading the field implementation teams and data analysis. Funding for this project came from a World Bank – Royal Netherlands Embassy trust fund. All views expressed are those of the authors, and do not necessarily reflect the views of the World Bank, the Royal Netherlands Embassy, Mitra Samya, or the Indonesian Central Bureau of Statistics.

I. Introduction

Targeted social safety net programs have become an increasingly common tool to reduce poverty. In developed countries, the selection of the beneficiaries of these programs ("targeting") is frequently accomplished through means-testing: only those with incomes below a certain threshold are eligible. In developing countries, most potential recipients work in the informal sector and lack verifiable records of their earnings. To collect income data, governments can conduct a census of every household in the country, gathering information on their consumption and earnings. However, this is both slow and expensive, besides being eminently manipulable: if people know that the data determine eligibility, their incentive to report their information correctly may be quite weak.

As a result, there is increased emphasis on targeting strategies that do not rely on directly observing incomes: either what are called proxy means tests (PMTs) or community-based targeting. In PMT, the government conducts a census of assets and demographic characteristics to create a "proxy" for household consumption or income, and this proxy is in turn used for targeting. In community-based methods, the government allows the community or some part of it (e.g. local leaders) to decide who should receive government assistance. Both methods aim to address the problem of unobservable incomes: in PMT, the presumption is that the assets and demographic characteristics used in the PMT are harder to conceal from government surveyors than income; in community-based targeting, the observability problem is addressed if one's wealth is harder to hide from neighbors than from the government.

The tradeoff between the two approaches is generally framed as a race between the better information that communities might have versus the risk of elite capture in the community process. By focusing on assets, PMTs aim to capture the permanent component of consumption, but they may miss out on transitory or recent shocks. For example, a family might have fallen into poverty because one of its members has fallen ill, but it may still own a large house that classifies it as non-poor under the PMT. Neighbors, on the other hand, might know the family's true economic situation, either from spending time with them or merely by observing the way they live (e.g. the way they dress, what they buy).[2] The perception that the PMT is getting it wrong can undermine the legitimacy of the PMT process and create political problems.[3]

However, while community targeting utilizes local information, it also allows for the possibility that targeting decisions may be based on a wide range of factors beyond poverty as defined by the government. For example, the community may not care as much about the poor or certain groups among the poor. It may also prefer to use the money to achieve broader social goals, such as paying off the potentially disgruntled to buy social peace. Or, elites may capture the process and select only their friends and relatives.[4]

Given the tradeoffs involved, it is ultimately an empirical question as to whether PMT or community methods work best. Coady, Grosh and Hoddinott (2004) conduct a meta-analysis of 111 targeted anti-poverty programs in 47 countries, including 7 PMTs and 14 cases of community-based targeting. They find no difference in the performance of these two models, as measured by the fraction of total targeted resources that went to the bottom 40 percent. This may be due, in part, to the small samples sizes in these studies. Moreover, at least two sources of bias, which work in opposite directions, may be present. First, the authors suggest that

---

[2] Seabright (1996) makes the theoretical argument that greater local information is one of the advantages of the community methods. Alderman (2002) and Galassso and Ravallion (2005) provide empirical evidence that communities may have additional information beyond the PMTs.
[3] See, for example, "Data Penerima BLT di Semarang Membingungkan" (BLT Beneficiary List in Semarang Confuses)" *Kompas* (5/15/08), "Old data disrupts cash aid delivery," *Jakarta Post* (9/6/08); "Poorest still waiting for cash aid," *Jakarta Post* (6/24/08); "Thousands protest fuel plan, cash assistance," *Jakarta Post* (5/22/08).
[4] Corruption is possible in PMTs as well, as households could use money or influence to induce the surveyor to misreport its assets. Camacho and Conover (2008) document widespread manipulation of PMT scores in Columbia.

community targeting is often chosen when state capacity is limited and the community functions well together. In such places, the PMT would have fared worse had it been tried. Second, the authors suggest that many relatively small projects have used the community model, but fail to report data in a way that would facilitate a meta-analysis. As such, the examples of community-based targeting that are included in that study tend to be bigger and, potentially, better run than the average. Thus, the relative efficacy of these methods remains a question.

In this paper, we report on a field experiment designed to understand how community methods fare relative to the PMT. In 640 villages in Indonesia, the government implemented a cash transfer program that distributed 30,000 Rupiah (about $3) to households that fell below location-specific poverty lines. In one-third of the villages (randomly selected), the government conducted PMT to target the poor. In another third of these villages, once again chosen at random, it employed community-based targeting. Specifically, the community members were asked to rank everyone from richest to poorest during a meeting, and ranking determined eligibility. In the remaining villages, a hybrid of the two methods was used: communities engaged in the ranking exercise, and then the ranks were used to limit the universe of individuals who should be interviewed by the government. Eligibility was then determined by conducting PMT on this limited list. This hybrid method was designed to make use of the communities' greater knowledge, while at the same time using the PMT as a check on potential elite capture.

To evaluate the relative performance of these different methods, we focus on two measures: the ability to identify the poor and the local satisfaction with the beneficiary list. To measure targeting ability, we collected per capita expenditure data from a set of households in selected villages prior to the experiment. We then determined which targeting method was most successful in identifying the poor in terms of per capita expenditure. When poverty is defined using the common PPP$2 per day poverty cutoff, we find that both the community and hybrid methods perform worse than the PMT. Specifically, in both methods, there was a 3 percentage point (10 percent) increase in mis-targeting rates relative to the PMT. However, the community-based strategies actually do as well (if not better) at finding the very poor, those with consumption below PPP$1 per day.

Despite the worse targeting outcomes, the community methods resulted in higher satisfaction levels and greater legitimacy of the process along all dimensions we considered. For example, communities lodged 60 percent fewer complaints in the comment box, and there were many fewer reports of difficulties in distributing the actual funds. When asked ex-post about the targeting process, the community treatment villages suggested fewer modifications to the beneficiary list, and reported being much more satisfied with the process.

To understand why community methods differ from the PMT, we examined several alternative explanations: elite capture of the community process, the role of community effort in conducting the rankings, local concepts of poverty beyond per-capita information, and local information about poverty.

To test for elite capture of the ranking exercise, we randomly divided the community and hybrid villages so that, in half of these villages, everyone in the community was invited to participate in the ranking meeting, whereas in the other half, only a few local leaders (the elites) were invited. In addition, we gathered data in the baseline survey on which households were related to the local elites. We find no evidence of elite capture. The overall mis-targeting rates were the same, regardless of who comprised the meeting. Moreover, we find no evidence that elite connected households are more likely to receive funds in the community treatments relative to the PMT. In fact, we find the opposite: in the community treatments, elites are much less likely to be put on the beneficiary list, regardless of their actual income levels.

To examine the role of effort, we randomized the order in which households were considered at the meetings. This allows us to test whether the effectiveness of community targeting differs between those households who were ranked first (when the community members were still full of energy) with those ranked at the end (when fatigue may have set in). We find suggestive evidence that effort matters: for the first few households that the community considers, targeting performance is better than in the PMT, but by the time the community gets to the last household, targeting performance is worse.

To examine the role of preferences and information, we created and studied alternative metrics of which households were perceived to be poor in our baseline survey. First, we asked every survey respondent to rank a set of randomly chosen villagers from rich to poor ("survey ranks"). Second, we asked the head of the sub-village to conduct the same exercise. Finally, we asked each household we interviewed to subjectively assess their own welfare level. We find that the community treatment produces a ranking of villagers that is much more correlated with these three alternate metrics than the ranking produced by PMT. In other words, the community treatments moved the targeting outcomes away from a ranking based purely on per-capita consumption and towards the rankings you would get by polling different classes of villagers.

Two stories could explain these findings: the community has less information about household's per-capita consumption than the PMT or the community's conception of poverty is different from that based solely on per-capita consumption. The evidence suggests that the latter theory is closer to the truth. First, the correlation of self-assessments with the survey ranks is higher than the correlation of self-assessments with per-capita consumption. Thus, in assessing their own poverty, villagers (who presumably have complete information about their own poverty) use a welfare metric that looks more like what community members use to assess each other than their own per-capita consumption. Second, the survey ranks contain information about those villagers' per-capita consumption, even controlling for all variables in the PMT, i.e. the community has residual information about consumption beyond the PMT. Third, when we regress the survey rankings on the PMT variables, the coefficients on almost all of the variables are of plausible signs and magnitudes, suggesting that the community survey ranks incorporate the PMT information as well. Finally, when we investigate how the survey ranks differ from consumption, we find that communities place greater weight on factors that predict earnings capacity than would be implied simply by per-capita consumption. For examples, conditional on actual per capita consumption levels, communities consider widowed households poorer than the typical household. The fact that communities employ a different concept of poverty explains why community targeting performance might differ from the PMT, as well as why the community targeting might result in greater satisfaction levels.

The paper proceeds as follows. We discuss the empirical design in Section II, and we describe the data in Section III. In Section IV, we compare how each of the main targeting methods fared in identifying the poor. Section V tests for evidence of elite capture, while Section VI aims to understand the role of effort. In Section VII, we test whether the community and the government have different maximands. Section VIII concludes.

II. Experimental Design and Data

II.A. Setting

This project occurred in Indonesia, which is home to one of the largest targeted transfer cash programs in the developing world, the Direct Cash Assistance (*Bantuan Langsung Tunai*, or *BLT*) program. Launched in 2005, the BLT program provides transfers of about US $10 per month to about 20 million households.

Targeting in the BLT program was accomplished with a combination of community-based methods and proxy-means tests. In 2005, Central Statistics Bureau (*Biro Pusat Statistik*, or *BPS*) enumerators met with village leaders to determine a list of households who could potentially qualify for the program. The BPS enumerators then collected asset data only for the households that were suggested by the community leaders. These data were used to create a score that determined eligibility. In 2008, the list was updated: the enumerators met with village leaders to ask if anyone else should potentially be added to the list, and then collected asset data on this new combined list. PMT regressions were used to determine final eligibility.

Targeting is the key problem that Indonesia faces in the BLT program. Using the common $2 PPP per day poverty threshold, the World Bank estimates that 45 percent of the funds were mis-targeted to non-poor households and 47 percent of the poor were excluded from the program in 2005-2006 (World Bank, 2006).[5] Perhaps more worrisome from the government's perspective is the fact that citizens voiced substantial dissatisfaction with the beneficiary lists. Protests about mis-targeting in the program led some village leaders to resign rather than defend the beneficiary lists to their constituents.[6] In fact, over 2,000 village officials

[5] Targeting inaccuracy has been documented in many government anti-poverty programs that offer subsidized rice (*Raskin*), basic commodities *(Operasi Pasar Khusus*), health insurance (*Askeskin/Jamkesmas*), and scholarships for poor households. See, for example, Olken (2006); Daly and Fane (2002); and Cameron (2002).
[6] See for example: "BLT Bisa Munculkan Konflik Baru" (BLT May Create New Conflicts), *Kompas* (5/17/08), and "Kepala Desa Trauma BLT" (A Village Head's Trauma with BLT) *Kompas* (5/24/08).

refused to participate in the program for this reason.[7] The experiment reported in this paper was designed and conducted in collaboration with the Central Statics Bureau of Indonesia to investigate the two primary targeting issues: improving targeting performance and increasing popular acceptance of targeting results.

II.B. Sample

The sample for the experiment consists of 640 sub-villages spread across three Indonesian provinces: North Sumatra, South Sulawesi, and Central Java. The provinces were chosen to be broadly representative of Indonesia's diverse geography and ethnic makeup, with one province located on each of the three most populous islands (Sumatra, Sulawesi, and Java). Within these three provinces, we randomly selected a total of 640 villages, stratifying the sample to consist of approximately 30 percent urban and 70 percent rural locations.[8] For each village, we obtained a list of the smallest administrative unit within it (a *dusun* in North Sumatra and *Rukun Tetangga (RT)* in South Sulawesi and Central Java), and randomly selected one of these sub-villages for the experiment. These sub-village units are best thought of as neighborhoods. Each sub-village has an elected or appointed administrative head, whom we refer to as the sub-village head, and contains an average of 54 households.

---

[7] See for example: "Ribuan Perangkat Desa Tolak Salurkan BLT" (Thousands of Village Officials Refuse to Distribute BLT), *Kompas* 5/22/08 and "DPRD Indramayu Tolak BLT," (Village Parliament of Indramayu Refuses BLT), *Kompas*, 5/24/08.

[8] An additional constraint was applied to the district of Serdang Bedagai because it had particularly large-sized sub-villages and more populous areas entail lengthier data collection. All villages in this district with average populations above 100 households per sub-village were excluded. In addition, five of the originally-selected villages were replaced prior to the randomization due to an inability to reach household respondents during the baseline survey, the village head's refusal to participate, or the presence of local conflict.

II.C. Experimental Design

In this section, we describe the basic experimental setup, the main treatments, and the sub-treatments. We then discuss the randomization design and the timing of the experiment.

Basic setup

In each sub-village, the Central Statistics Bureau and Mitra Samya, an Indonesian NGO contracted to run the experiment, implemented an unconditional cash transfer program, where beneficiary households would receive a one-time, Rp. 30,000 (about $3) cash transfer. The amount of the transfer is equal to about 10 percent of the median beneficiary's monthly per-capita consumption, or a little more than one day's wage for an average laborer.[9]

Each sub-village was randomly allocated to one of the various targeting treatments (described in more detail below). The fraction of households in the sub-village that would receive the transfer was set in advance through a geographical targeting approach, which was applied identically in all villages such that the fraction of households receiving the subsidy was held constant across all treatments in the experiment. We then observed how each treatment selected *the set of* beneficiaries. After the targeting process occurred and the beneficiary lists were finalized, the funds were distributed to the designated households. To publicize the results, the facilitators posted two copies of the beneficiary list in highly visible community locations such as roadside food stalls, mosques/churches, or the sub-village head's house. The facilitators also placed a suggestion box and a stack of complaint cards next to the list, along with a reminder about the program details and complaint process. Depending on the sub-village head's preference, the cash distribution could occur either through door-to-door handouts or by

[9] While the amount of the transfer is substantially smaller than in the national BLT program (which distributed Rp. 100,000 per month), the amount is substantial enough that poor households would want to receive it. In fact, in September 2008, more than twenty people were killed during a stampede involving thousands when a local wealthy person offered to give out charity of Rp. 30,000 per person (Kompas, 9/15/08).

gathering the recipients at a central location.[10] After at least three days, the facilitators collected the suggestion box.

Main Treatment 1: PMT

In the PMT treatment, the government created formulas that mapped easily observable household characteristics into a single index using regression techniques. Specifically, it created a list of 49 indicators, encompassing the household's home attributes (wall type, roof type, etc), assets (TV, motorbike, etc), household composition, and household head's education and occupation.[11] Using pre-existing survey data, they estimated the relationship between these variables and household per-capita consumption.[12] While it collected the same set of indicators in all regions, the government estimated district-specific formulas due to the high variance in the best predictors of poverty across regions. On average, these PMT regressions had an R^2 value of 0.48.

Government enumerators collected these indicators from all households in the PMT villages by conducting a door-to-door survey.[13] These data were then used to calculate a computer-generated poverty score for each household using the district-specific PMT formula. A list of beneficiaries was generated by selecting the pre-determined number of households with the lowest PMT scores in each sub-village.[14]

---

[10] If no adult household members were present in the household, facilitators were required to revisit the household. If the revisit was unsuccessful, the facilitator could entrust either the sub-village head or a neighbor to pass on the funds to the designated beneficiary.
[11] The questions were similar to those in the survey used during Indonesia's recent BLT targeting survey.
[12] Data used to determine the weights on the PMT equations are from Indonesia's SUSENAS (2007) and World Bank's Urban Poverty Project (2007) datasets.
[13] Enumerators had been provided with each sub-village census (from the baseline survey) prior to their visits. The Indonesian government's usual definition of "household" was used: "a person or group of persons who share part or all of fiscal expenses and usually eat together from the same kitchen"
[14] The fraction of households that would be beneficiaries in a given village was kept secret by the computer program until the PMT data had been fully collected and entered into the computer.

Main Treatment 2: Community Targeting

In the community treatment, the sub-village residents determine the list of beneficiaries through a poverty-ranking exercise. At a community meeting, residents ranked all households in the neighborhood from richest to poorest, with the poorest households becoming the beneficiaries of the program.

To start, a local facilitator visited each sub-village and informed the sub-village head about the program. Meeting dates were set several days in advance to allow the facilitator and sub-village head time to spread awareness. Facilitators encouraged attendance by making door-to-door household visits. On average, 45 percent of households attended the meeting.

At the meetings, the facilitator first explained the program. Next, they displayed the list of all households in the sub-village (which came from the baseline survey), and asked the meeting attendees to make corrections to the list if necessary. To assist the community in ranking households, the facilitator then spent about 15 minutes having the community brainstorm a list of characteristics that differentiate the poor from the wealthy households in their community.[15]

The facilitator then proceeded with the ranking exercise using a set of randomly-ordered index cards that displayed the names of each household in the neighborhood. He hung a string from wall to wall, with one end labeled as "most well-off" (*paling mampu*) and the other side labeled as "poorest" (*paling miskin*). Then, he started by holding up the first two name cards from the randomly-ordered stack and asking the community, "Which of these two households is better off?" Based on the community's response, he ranked the two cards by paper-clipping them along the string in order with the poorer household closer to the "poorest" end of the string.

[15] These characteristics or criteria were not used explicitly in the ranking; they were instead used more to help communities think about how they would differentiate households in making ranking decisions.

Next, he displayed the third card from the stack and asked how this household ranked relative to the first two households. The activity continued with each card being positioned relative to the already-ranked households one-by-one until all were ranked.[16] By and large, the community reached a consensus on how to rank each card.[17] Before the final ranking was recorded, all household names were read aloud to the community so adjustments could be made if necessary.

After all meetings were complete, the facilitators were provided with "beneficiary quotas" for each sub-village based on the geographic targeting procedure described above. Households ranked below the quota were eligible to receive the transfer. Note that prior to the ranking exercise, facilitators told the meeting attendees that outsiders would determine the quotas, but that once the quota was determined, all households who were ranked below the quota would receive the transfer. Facilitators also emphasized that the government would not interfere in any way with the community's ranking.

Main Treatment 3: Hybrid

The hybrid method combines the community ranking procedure with a subsequent PMT verification. In this method, the entire community-ranking procedure, described above, was implemented first. However, there was one key difference: at the start of these meetings, the community was told that lowest-ranked households would be independently checked by the government enumerators before the beneficiary list was finalized.

---

[16] Once at least 10 households had been ranked, to speed up the process the facilitator began by comparing each card to the middle card, and if it was higher than middle, to the 75th percentile card, and so on to most quickly arrive at the exact position of the new card.

[17] If the community did not know a household or the household's status could not be decided upon, the facilitator and several community members visited the household after the meeting and then added them to the rank list based on the information gained from the visit. In practice, only 2 of the 431 community or hybrid villages had any households that could not be ranked at the meetings (19 out of 67 households at one meeting, all of whom were boarders at a boarding house, and 5 out of 36 households at the second meeting.).

After the community meetings were complete, government enumerators visited the lowest-ranked households to collect the data needed to calculate the PMT score. The number of households to be visited ("visiting quotas") was computed using the same method described above for the community treatment's "beneficiary quotas," but multiplied by 150 percent. Beneficiary lists were then determined using the PMT formulas. Thus, it was possible that some households could become beneficiaries according to the PMT criteria even if they were ranked as slightly wealthier than the beneficiary quota cutoff line on the community list. Conversely, some relatively poor-ranked households on the community list might become ineligible.

The hybrid treatment aims to take advantage of the relative benefits of both methods. First, as compared to the community method, the hybrid method's additional PMT verification phase serves as a check against elite capture. Second, in the hybrid method, the community is incentivized to accurately put the poorest households at the bottom of the ranking, as richer households would later be eliminated by the PMT. Third, as compared to the PMT treatment, the hybrid method's use of the community rankings to narrow the set of households that need to be visited may be potentially more cost-effective, in light of the fewer household visits required.

Community Sub-Treatments

We designed several sub-treatments in order to test three hypotheses about why the results from the community process might differ from those that resulted from the PMT treatment: elite capture, community effort, and heterogeneity in preferences within the community.

First, to test for elite capture, we randomly assigned the community and hybrid sub-villages to two groups: a "whole community" sub-treatment and an "elite" sub-treatment. In "whole community" areas, the facilitators actively recruited all community members to participate in the ranking. In the "elite" areas, meeting attendance was restricted to no more than seven invitees chosen by the sub-village head. Sub-village heads were encouraged to invite community members that were usually involved in village decision-making, such as religious leaders or school teachers. At least one woman was required to attend. The elite meetings have the advantage that it is easier and less costly to organize the meetings with just a few individuals relative to the whole community approach. Moreover, the elites may have greater legitimacy (and possibly better information) to make difficult allocation decisions on behalf of their citizens. However, the elite meetings also carry the risk that the elites may choose to funnel aid to friends and family in a setting where the broader community is absent (Bardhan and Mookherjee 2005).

Second, we introduced several treatments to investigate whether effort levels on the part of the community played an important part in the accuracy of the community rank lists. The rank list procedure takes time and considerable effort. On average, it took 1.68 hours for the community to complete the rankings. For a community with the mean number of households (54), by the time the community places the last card, even an optimal sorting algorithm would require the community to make 6 pair-wise comparisons in order to accurately place the card in the rank list. It is therefore plausible to assume that, in longer meetings, the community members might become fatigued and expend less effort.

We introduced two randomizations to investigate the role of effort. First, we randomized the order in which households were ranked. Specifically, we ordered the stack of index cards used in the ranking based on a randomized number.[18] Second, in half of the meetings, the facilitator led an exercise to identify the ten poorest households in the sub-village prior to the ranking exercise. If effort is important, this treatment should increase targeting accuracy by

[18] Any new household cards that were added to the stack during this process were ranked last due to the logistical complexity of re-randomizing the order of the entire stack.

allowing the community to focus first (when it presumably has the most energy and attention) on correctly identifying the very poor. More generally, this treatment could also clarify the community's conception of poverty.

The third set of hypotheses concerns the role of preferences. If the community results differ from the PMT results because of preferences, it is important to understand whether these preferences are broadly shared or are simply a function of who attends the meeting. Thus, we varied the meeting times in order to attract different subsets of the community. Half of the meetings were randomly assigned to occur at night (after 7:30 pm), when attendance would be dominated by men who work during the day. The remaining meetings were assigned to occur during the afternoon, when we expected higher female attendance.

Randomization Design

Each of the 640 sub-villages was randomly assigned to treatments as follows (see Table 1). First, in order to ensure experimental balance across geographic regions, we created 51 geographic strata, where each stratum consists of all villages from one or more sub-districts (*kecamatan*) and is entirely located in a single district (*kabupaten*).[19] Then, we randomly allocated sub-villages to one of the three main treatments (PMT, community, or hybrid), stratifying such that the proportion of sub-villages allocated to each main treatment was identical (up to integer constraints) within each stratum. We then randomly and independently allocated each community or hybrid sub-village to the community sub-treatments, with each of these three sub-treatment randomizations stratified by stratum and main treatment.

[19] Specifically, we first assigned each of the 68 subdistricts (*kecamatan*) in the sample to a unique stratum. We then took all subdistricts with 5 or fewer sampled subdistricts and merged them with other kecamatans in the same district, so that each of the resulting 51 strata had at least 6 sampled villages. Note that a subdistrict is the next-highest administrative unit above a village, and consist of 5-20 villages and 15,000 – 50,000 inhabitants.

Timing

In November to December 2008, an independent survey company conducted a census in each sub-village and then collected the baseline data. The targeting treatments and the creation of the beneficiary lists started immediately after the baseline survey was completed during the months of December 2008 and January 2009. Fund distribution, the collection of the complaint form boxes, and interviews with the sub-village heads occurred during February 2009. Finally, the survey company conducted the endline survey in late February and early March 2009.

III. Data

We collected four main sources of data: a baseline household survey; household rankings generated by the treatments; data on the community meeting process (in community/hybrid treatments only); and data on community satisfaction. In this section, we describe the data collection effort, and then provide summary statistics and a test of the randomization.

III.A. Baseline Data

We conducted a baseline survey in November and December 2008. The survey was administered by SurveyMeter, an independent survey organization. No mention was made of the targeting experiment until after the baseline survey had been concluded.[20] We began by constructing a complete list of all households in the sub-village. From this census, we randomly sampled eight households from each sub-village and the head of the sub-village, for a total sample size of 5,756 households.[21] To ensure gender balance among survey respondents, in each sub-village, households were randomized as to whether the household head or spouse of the

[20] SurveyMeter enumerators were not told about the targeting experiment.
[21] In four of the sub-villages, there were seven respondents rather than eight due to respondent refusals.

household head would be targeted as the primary respondent. The survey included questions about demographic variables, family networks in the sub-village, participation in community activities, relationships with local leaders, access to existing social transfer programs, and detailed data on the households' per capita consumption.

The baseline survey also included (as part of a module on social capital) a variety of measures of the household's subjective poverty assessments. In particular, we asked each household to rank the other eight households surveyed in their sub-village from poorest to richest. We also asked each respondent to list whom he or she considered to be the five poorest and five richest households in the sub-village, as well as any households whom they considered formal or informal leaders in the sub-village. To measure elite connectedness, we asked respondents to identify any household in the sub-village who was related by marriage or blood to the households that they identified as poor, rich, or leaders. Finally, we asked respondents several subjective questions to determine how they assessed their own poverty levels.

III.B. Data on treatment results

Each of our treatments – PMT, community, and hybrid – produces a rank ordering of all households in the sub-village ("targeting rank list"). For the PMT treatment, this is the rank ordering of the PMT score, i.e., the regression-predicted per-capita consumption based on asset data collected in the PMT survey. For the community treatment, it is the ranking of households that was constructed during the community meetings. For the hybrid treatment, it is the final ranked list, where all households below the visiting quota are ordered based on their PMT score and all households above the visiting quota are ordered based on their rank at the community meeting. For all treatments, we additionally collected data on which households actually received the program (i.e. which households fell below each sub-village quota).

III.C. Data on community meetings

For the community and hybrid sub-villages, we collected data on the meetings' functioning. We collected attendance lists (cross-referenced to the unique household identifiers assigned when we constructed the census), and also the number of female participants.[22] Timesheets recorded how many minutes were spent on each part of the community meeting (e.g. updating the household list, discussing poverty criteria, and conducting the ranking). Facilitators also recorded the household identifiers of the ten poorest households listed in the "poor household priming" treatment. After each meeting, the facilitators filled out a questionnaire on their perceptions of the community's interest and satisfaction with the ranking exercise. They were also asked to evaluate community participation levels, including participation by women and the poor, and to list the five most actively participating households.

III.D. Data on community satisfaction

After the cash disbursement was complete, we collected data on the community's satisfaction level using four different tools: suggestion boxes, sub-village head interviews, facilitator feedback, and household interviews. First, facilitators placed suggestion boxes in each sub-village along with a stack of complaint cards. Each anonymous complaint card asked three yes/no questions in a simple format: (1) Are you satisfied with the beneficiary list resulting from this program? (2) Are there any poor households not included on the list? (3) Are there any non-poor households included on the list? Second, on the day when suggestion boxes were collected, the facilitators interviewed the sub-village heads, and asked them about complaints submitted to

[22] The attendance form used was somewhat different in the elite sub-treatment, so we verified the attendance results by treatment using data from the endline household survey. All other forms were identical in all sub-treatments.

them verbally.[23] Sub-village heads were also asked if they were personally satisfied with the targeting outcome. Third, each facilitator filled out feedback forms on the ease of distributing the transfer payments. Finally, in Central Java province, an independent survey company conducted an endline survey using questions that were similar to those asked to the sub-village head. Three randomly-selected households were interviewed out of the eight households surveyed in the baseline survey.[24]

III.E. Summary statistics

Table 2 provides sample statistics of the key variables. Panel A shows that average monthly per capita expenditures are approximately Rp. 558,000 (about $50 dollars).

Panel B provides statistics on mis-targeting rates. By construction, about 30 percent of the households received the cash transfer. We calculated how many individuals were mis-targeted, using per capita consumption in the baseline survey to determine eligibility. Specifically, we calculated the per-capita consumption level in each province of our data (separately by urban and rural areas) that corresponded to the percentage of households who were supposed to receive the transfer; this threshold level is approximately equal to the PPP$2 poverty line.[25] We defined "mis-target" to be equal to 1 if either the household's per capita

[23] We were concerned that sub-village heads might give positive-biased feedback on the program in community and hybrid areas, since they had previously met the facilitators in these areas during the community meeting process a few weeks earlier (while in the PMT treatment areas, the facilitators did not meet the sub-village heads until the fund distribution time). For this reason, the researchers intended to randomly re-assign facilitators' designated sub-villages after the fund distribution so that no facilitator would collect the sub-village head's feedback from an area he/she had already visited. While it proved to be logistically impossible to make this change in North Sumatra, the re-assignment was implemented in the other provinces.

[24] Time and budget constraints prevented the possibility of surveying in North Sumatra and South Sulawesi.

[25] To see this, note that adjusting the 2005 International Price Comparison Project's PPP-exchange rate for Indonesia for inflation through the end of 2008 yields a PPP exchange rate of PPP$1 = Rp. 5549 (author's calculations based on World Bank 2008 and the Indonesian CPI). The PPP$2 per day per person poverty line therefore corresponds to per-capita consumption of Rp. 338,000 per month. In our sample, the average threshold below which households should have received the transfer is is Rp. 320,000 per month, or almost exactly PPP$2 per day. The slight discrepancy is due to different regional price deflators used in the geographic targeting procedure.

consumption was below the threshold line and did not receive the transfer (exclusion errors) or if it was above the threshold line and did receive the transfer (inclusion errors). We also calculate mis-target for subsets of the population: those below the threshold, whom we call the "poor" (divided in half into the "very poor", with per-capita consumption below approximately PPP$1, and the "near poor", with per-capita consumption between approximately PPP$1 and PPP$2) and those above the threshold, whom we call the "non-poor" (divided in half into "middle income" and "rich").

As shown in Panel B, 32 percent of the households were mis-targeted. Twenty percent of households that should not have qualified for the transfer (the non-poor) received it, while 53 percent of poor households did not. Disaggregating the data further, the rich are the least likely to be mis-targeted (14 percent), while the near poor are the most likely to be (58 percent).

Panel C provides summary statistics for another metric of targeting: the rank correlation for each sub-village between alternative metrics of household well-being and results of the targeting experiment ("targeting rank list"). By using rank correlations, we can flexibly examine the relationship between the treatment outcomes and various measures of well-being on a comparable scale. First, we compute the rank correlation with per capita consumption, which tells us how closely the final outcome is to the government's metric of well-being. Second, we compute the rank correlation with the ranks provided by the eight individual households during the baseline survey. This allows us to understand how close the targeting rank list is to the community member's individual beliefs about their fellow community members' well-being. Third, we compute the rank correlation with the ranks provided by the sub-village head in the baseline survey, which does same things for the sub-village head's views. Finally, we compute the rank correlation with respondents' self-assessment of poverty, as reported in the baseline survey.[26] This final rank correlation allows us to understand how closely the treatment result matches individual's beliefs about their own well-being.

While targeting rank lists are associated with consumption rankings, they are more highly associated with the community's rankings of well-being. While the mean rank correlation between the consumption rankings and the targeting rank lists is 0.41, the mean correlation of the targeting rank list with the individual community members' ranks is 0.64, and the correlation with the sub-village head's ranks is 0.58. Finally, we see a 0.40 correlation between the ranks from the targeted lists with the individual' self assessment of their own poverty.

III.F. Randomization Balance Check

Before turning to the results, we first examine whether the randomization for the main treatments appears balanced across covariates. We chose ten variables for this check prior to obtaining the data from the experiment.[27] Specifically, we examined the following characteristics from the baseline survey: per capita expenditures, years of education of the household head, calculated PMT score, the share of households that are agricultural, and the years of education of the sub-village head. We also examined five village characteristics from the 2008 PODES, a census of villages conducted by BPS that contains about 400 village characteristics: log number of households, distance to district center in kilometers, log size of the village in hectares, the number of religious buildings per household, and the number of primary schools per household. We present the results from this analysis in Table 3. In Columns 1, 2, and 3, we present the mean of each variable for sub-villages assigned to the PMT, community,

[26] Specifically, each household was asked "Please imagine a six-step ladder where on the bottom (the first step) stand the poorest people and on the highest step (the sixth step) stand the richest people. On which step are you today?" Each respondent responded with a number from 1 to 6, and we use the rank of this response among respondents in a village in computing the rank correlation.
[27] In fact, we specified and documented all of the main regressions in the paper before examining the data (April 3, 2009). The document is available from the authors upon request.

and hybrid treatments, respectively. Standard deviations are listed below the mean in brackets. We present the difference in means between the community and PMT groups in Column 4, between the hybrid and PMT in Column 5, and between the hybrid and the community in Column 6. In Columns 7 – 9, we replicate the analysis shown in Columns 4-6, but additionally control for stratum fixed effects. Robust standard errors are shown in parentheses in Columns 4 – 9. All variables are aggregated to the sub-village level; thus each regression includes 640 observations. In the final row of Table 3, we provide the p-value of a test of joint significance of the difference across each of the outcome variables.

The sub-villages appear to be generally well-balanced across the ten characteristics. Out of the sixty individual differences presented, three are statistically significant at the 5 percent level, precisely what one would expect from random chance. All of these differences are in Column 9, which compares community to hybrid, controlling for stratum fixed effects. Specifically, controlling for stratum fixed effects, households in community locations have less education, there are fewer agricultural households in the villages assigned to the community treatment than the hybrid treatment, and hybrid villages have 8 percent fewer households than community villages. Looking at the joint significance tests across all ten variables considered, without stratum fixed effects, the only jointly significant difference is between hybrid and community (column 6, p-value 0.082); with stratum fixed effects (column 9), the p-value is 0.028. All results in this paper are robust to specifications that include these additional ten control variables.

IV. Main Results: Targeting Performance and Satisfaction

The government had two main goals for the experiment: 1) to understand how to increase the accuracy of the beneficiary lists, and 2) to assess the best method to increase satisfaction with and legitimacy of the results. This section presents the results as pertains to these goals.

IV.A. Targeting performance

We begin by comparing how the different targeting methods performed based on actual per-capita consumption levels, the metric of poverty used by the government. Specifically, we compute location-specific poverty lines based on the PPP$2 per day consumption threshold, and then classify a household as mis-targeted if its per capita consumption levels is below the poverty line and it was not chosen as a beneficiary, or if it was above the poverty line and it was identified as a recipient (MISTARGET_{ivk}). We then examine which method minimized mis-targeting by estimating the following equation using OLS:

$$\text{MISTARGET}_{ivk} = \alpha + \beta_1 \text{COMMUNITY}_{ivk} + \beta_2 \text{HYBRID}_{\text{ivk}} + \gamma_k + \varepsilon_{ivk} \tag{1}$$

where i represents a household, v represents a sub-village, and k represents a stratum. Note that the PMT treatment is the omitted category, so β_1 and β_2 are interpretable as the impact of the community and hybrid treatments relative to the PMT treatment. Stratum fixed effects (γ_k) are included. Since the targeting methods were assigned at the sub-village level, the standard errors are clustered to allow for arbitrary correlation within a sub-village.

The results, shown in Table 4, indicate that the PMT method outperforms both the community and hybrid treatment in terms of the overall mis-target rates. Under PMT, 30 percent of households are mis-targeted (Column 1). Both the community and hybrid methods increase the mis-targeting rate by about 3 percentage points —or about 10 percent— relative to the PMT method (significant at the ten percent level).

Adding a rich household to the list may have different welfare implications than adding a household that is just above the poverty line. Thus, we next determine whether mis-targeting varies by the income level of the households. Figure 1 graphs the log per capita consumption distribution of the beneficiaries (left panel) and non-beneficiaries (right panel) for each targeting treatment. The vertical lines in the graphs indicate the PPP$1 and PPP$2 per day poverty lines. Overall, the graphs confirm that all methods select relatively poorer households: for all methods, the mode per-capita consumption for beneficiaries is below PPP$2 per day, whereas it is above PPP$2 per day for non-beneficiaries.

Examining the impact of the treatments, the left panel shows that the consumption distribution of beneficiaries under PMT is centered to the left of the distribution under the community and hybrid methods, indicating that on average the PMT identifies poorer individuals. However, looking at the leftmost part of the distributions, the community methods select a greater percentage of beneficiaries whose log per-capita consumption is less than PPP$1 per day. Thus, the figures suggest that the community methods may include more of the *very* poor, even if they do worse on average. Moreover, the figures suggest that all three methods contain similar proportions of richer individuals (those with log income greater than about 6.5), and that the difference in mis-targeting across the three treatments is driven by differences in near poor (PPP$1 to PPP$2) and the middle income group (those above the PPP$2 poverty line but with a log income less than 6.5).

We more formally examine the findings from Figure 1 in the remaining columns of Table 4. In Columns 2 and 3, we examine mis-targeting separately for the poor and the non-poor. In Columns 4 and 5, we disaggregate the non-poor by splitting the sample approximately in half into rich and middle, and in Columns 6 and 7, we disaggregate the poor by splitting the sample approximately in half into near poor and very poor. The results show that in the PMT treatment, 18 percent of the non-poor were wrongly classified as poor (Column 2), and 52 percent of the poor were wrongly classified as non-poor (Column 3). Relative to the PMT treatment, the community treatment was 4.5 percentage points more likely to mis-classify the non-poor (Column 2, significant at the 5 percent level), while the hybrid was 3.7 percentage points more likely (Column 3, significant at the 5 percent level). Both the community and hybrid were a bit more likely to misclassify the poor, but these differences are not statistically significant.

Much of the difference in the error rate between the community methods and the PMT occur near the cutoff for inclusion. The community and hybrid methods are 6.6 and 5.2 percentage points, respectively, more likely to misclassify the middle non-poor (Column 5, both statistically significant at 5 percent). They are also more likely to misclassify the near poor by 5.2 and 3.6 percentage points, respectively, although these results are not individually statistically significant. In contrast, we observe much less difference between the PMT and community methods for the rich and the very poor, and in fact the point estimate suggests the community method may actually do better among the very poor.[28]

Finally, in Column 8, we examine the per capita consumption of beneficiaries across the three groups. As expected, given that the community treatment includes more of the very poor among beneficiaries but also includes more individuals who are just above the PPP$2 poverty line, the average per capita consumption of beneficiaries is not substantially different between the various treatments. This suggests that even though the community treatments do worse in mis-targeting the poor as defined by the PPP$2 cutoff, the welfare implications of the three methods appear similar based on the consumption metric.

---

[28] In results not reported in the table, we find that the difference in the community treatment's impact on mis-targeting rates between the very poor and the non-poor is statistically significant at the 10% level.

IV.B. Satisfaction

The second major goal was to understand how to increase satisfaction with and legitimacy of the targeting results. Table 5 presents data on how the treatments affected community satisfaction levels. Panel A presents data from the endline survey. Panel B presents data from the follow-up survey of sub-village heads. Panel C presents the results from the anonymous comment box, the community's complaints to the village head, and feedback from the facilitator on the ease of distributing the transfer payments.[29]

The results from the endline survey (Panel A) show that communities are overwhelmingly more satisfied with the community treatment than the PMT or hybrid treatments. For example, in community treatment, respondents wish to make fewer changes to the beneficiary list; they would prefer to add about one-third fewer households to the list of beneficiaries (Column 4) and subtract about one-half as many households (Column 5) than in the PMT or hybrid treatments. Villagers in community treatment sub-villages are more likely to report that the method used was appropriate (Column 1) and that they were satisfied with the program (Column 2). A joint test of the dependent variables in Panel A indicates that the community treatment differences are jointly statistically significant (p-value < 0.001).

Sub-village heads are also much more satisfied as well: the sub-village head was 38 percentage points more likely to say that the targeting method was appropriate in the villages where community-based targeting was used than in those which used PMT, and 17 percentage points less likely to name any households that should be added to the list (Panel B).

[29]For simplicity of interpretation, we use OLS / linear probability models for all dependent variables in Table 5. Using ordered probit for categorical response variables and probit for binary dependent variables produces the same signs of the results, and the same levels of statistical significance, as the results show in Table 5.

The higher levels of satisfaction were manifested in fewer complaints (Panel C). There were on average 1.7 fewer complaints in comment box for the villages assigned to the community treatment relative to the PMT treatment, and 0.93 fewer complaints in the hybrid treatment relative to the PMT villages (Column 2). The sub-village head also reported receiving 2.69 and 2.02 fewer complaints directly to him in the community and hybrid treatment, respectively (Column 3).

Higher satisfaction levels in the community led to a smoother disbursal process of the actual transfer. First, the facilitators who distributed the cash payment were 4-6 percentage points less likely to experience difficulties while doing so in sub-villages assigned to the community or hybrid method (Panel C, Column 4). Second, sub-village heads had a choice of how the facilitator would conduct the cash distribution: they could do so in an open community meeting or, if the head felt that they would encounter problems in the village, the facilitator could distribute the transfer door-to-door. Facilitators were 8 percentage points more likely to distribute the cash in an open meeting in the sub-villages assigned to the community treatment (Panel D, Column 5). They were also 5 percentage points more likely to do so in villages assigned to the hybrid treatment, but this result is not significant at conventional levels.[30]

IV.C Understanding the differences between PMT and community targeting

The findings present an interesting puzzle. The results on mis-targeting suggest that the community-based methods actually do somewhat worse at identifying the poor. However, the community method results in much greater satisfaction among both citizens and the head of the

[30] An important question is whether these differences in satisfaction represent changes from the act of directly participating in the process (as in Olken 2008). Although this is hard to address definitively, one suggestive piece of evidence comes from the elite treatment. In results not reported in the table, we find no differences in our measures of satisfaction between the whole community treatments (when 48 percent of households attended the meeting) and elite community treatments (when only 17.6 percent of households attended a meeting). This finding suggests that it is either differences in the list, or knowing that some type of local process was followed, that drives the differences in satisfaction in Table 5.

sub-village.

To understand these differences, it is useful to unpack the community targeting process to isolate several different ways in which it differs from a PMT approach. In making targeting decisions, the community maximizes some function $\Sigma\, \lambda_i u_c(\hat{C}_{ci}(e), X_i)$, where λ_i is the social welfare weight of household i, u_c is the utility function which communities use to assess welfare, $\hat{C}_{ci}(e)$ is the community's estimate of consumption (which is more accurate as the community puts in more effort, e, to observe it), and X_i are other characteristics (e.g., widowhood) that might enter the community's utility function. For the PMT, the equivalent "government" maximand would have $\lambda_i = 1$, u_g a concave function of per-capita consumption, and $\hat{C}_{pmti}$ the estimate of per-capita consumption based on the PMT asset variables.[31]

This simple framework suggests investigating several avenues through which PMT might differ from community targeting. First, there might be *elite capture* – i.e., the elites might control the ranking process so that λ_i is greater for elite-connected households. Second, the community could put in insufficient effort e at calculating each household's consumption. Third, the communities' utility function u may differ from the government, specifically in that it may take into account factors other than household per-capita consumption. If this is the case, community-based targeting, while resulting in less pro-poor targeting as judged by consumption, would result in higher welfare as judged by the community's welfare metric. Fourth, communities could simply have different information $\hat{C}_i$ about per-capita consumption than that which is gathered by the PMT. Information could be better or worse: the community might not know about all the assets measured in the PMT, or, alternatively, the community might know about these assets (and more), and also know more about the local income elasticity of demand for

[31] The choice variable here is whom to give a discrete amount of money. Even though getting the money might cause ranking reversals, it is always optimal, given that maximand, to give it to the poorest people.

each asset than the central government. The subsequent sections use additional sub-interventions and features of the data collected to investigate each of these potential differences in turn.

V. Elite Capture

Community-based targeting may involve a tradeoff: it allows the government to make use of local knowledge in targeting, but it also potentially opens the door for capture of the process by local elites (Bardhan and Mookerjee, 2005). To the extent that elites have different social welfare weights from the community as a whole ($\lambda_e \neq \lambda_c$), more elite control over the process should lead to more resources being directed at those households with high λ_e and worse overall targeting performance. The increased latitude for elites to capture the targeting process is one potential explanation for why the community targeting fared worse than the PMT.

We test for elite capture by examining the community sub-treatments that allow more or less elite control. Specifically, we would expect less elite capture in the hybrid treatment, when there is ex-post verification of the community's ranking, and we would expect more elite capture in the elite sub-treatment, when only elite members, rather than the whole community, were invited to participate in the rankings. We start by re-estimating equation (1), including a dummy for the ELITE sub-treatment and, in some specifications, the interaction of ELITE and HYBRID. Table 6 provides these results.

We first verify that the treatment had an impact on the meeting attendance. Columns 1 and 2 calculate the attendance rate using data collected at the meeting, while Columns 3 and 4 calculate it using the data from the endline household survey.[32] Both measures confirm that the

[32]Since the data in Column 1 and 2 come from the actual meetings, they are only available for the community and hybrid treatments. Since the data in column 3 and 4 come from questions about generic targeting meetings (not meetings specifically related to our project), it is possible to report having attended a meeting even though our project held no ranking meeting in their villages. These meetings could include meetings during the socialization of the program or meetings about another targeting related activity.

whole community meetings were substantially better attended than the elite-only meetings. For example, the survey data (Column 3) show that 49 percent of households attended the targeting meetings in the whole community treatment, whereas only 19 percent of households did so in the elite sub-treatment.

Despite these differences in attendance, the mis-targeting rate for the elite treatment was not significantly different than for the whole community treatment (Column 5 of Table 6). In Column 6, we examine the interaction of elite and hybrid. We would expect less elite capture in hybrid treatment, where the government verifies the results. Surprisingly, we find more mis-targeting in the hybrid methods when only the elites are invited. Overall, while the whole community meetings were more inclusive than the "elite" meetings, it does not appear that the presence of the full community affected the degree of elite capture.

This evidence presented in Table 6 does not entirely settle this issue: the data are consistent with no elite capture, but they could also be consistent with the elite so dominating the whole community meetings that both types of meetings reflect elite preferences. This second story seems unlikely: the facilitators report that a few individuals dominated the conversation in only 15 percent of the meetings, and that otherwise the meetings were a full community affair.

To probe this further, in Table 7, we examine whether or not elites and their relatives (those with high λ_e) were more likely to be on the list in both the full community and elite meetings relative to the PMT. Specifically, we estimate the following equation:

$$MISTARGET_{ivk} = \alpha + \beta_1 COMMUNITY_{ivk} + \beta_2\, HYBRID_{ivk} + \beta_3\, ELITE_{ivk} + \beta_4\, CONNECTED_{ivk} + \beta_5\, (COMMUNITY_{ivk} \times CONNECTED_{ivk}) + \beta_6\, (HYBRID_{ivk} \times CONNECTED_{ivk}) + \beta_7\, (ELITE_{ivk} \times CONNECTED_{ivk}) + \gamma_k + \varepsilon_{ivk} \quad (2)$$

where CONNECTED_{ivk} is an indicator that equals one if the household is related to any of the sub-village leaders/elites, or is one of the leaders themselves.[33] Columns 1 and 2 examine the mis-targeting rate as the dependent variable, and columns 3 and 4 examine whether a household received the transfer as the dependent variable. We find little evidence of elite capture. In fact, the point estimates suggest the opposite: elite connected households are less likely to be mis-targeted in the community and elite treatments, although the effect is not significant at conventional levels. Moreover, we find that elites are actually penalized in the community meetings: elites and their relatives are about 6.7 to 7.8 percent less likely to be on the beneficiary list in the community meetings, regardless of their income level (Columns 3 and 4). Overall, these findings suggest that the reason that mis-targeting is worse under the community method is not due to increased elite capture of the community process.

VI. Community effort

The community-based ranking process requires human effort to make each comparison. For example, ranking 75 households would require making at least 363 pair-wise comparisons.[34] One might imagine that the worse targeting in the community methods could result simply from fatigue as the ranking exercise progresses. We introduced two treatments to investigate the role of effort: randomization of order in which ranking happened and the 10 poorest treatment.

Figure 2 graphs the relationship between mis-targeting and the randomized rank order from a non-parametric Fan regression, with cluster-bootstrapped 95 percent confidence intervals

[33] Specifically, we defined an "elite connected" household as any household where 1) we interviewed the household and found that a household member held a formal leadership position in the village, such as village or sub-village head, 2) at least two of the respondents we interviewed identified the household as holding either a formal or informal (*tokoh*) leadership role in the village, or 3) a household connected by blood or marriage to any household identified in (1) or (2).

[34] The community sorting algorithm described above is called a binary insertion sort, which has been shown to require in expectation $\log_2(n!)$ pairwise comparisons to sort a list of size n, which is the theoretical lower bound for comparison sorting (Knuth 1998). In practice, the community may not perfectly implement this algorithm, so the number of comparisons may be even higher.

shown as dashed lines. Figure 2 shows that the mis-targeting rate is lowest for the first few households ranked, but then rises sharply by the 20th percentile of households. The magnitude is substantial – the point estimates imply that mis-targeting rates are between 5-10 percentage points lower for the first household ranked than for households ranked in the latter half of the meeting.

Table 8 reports results investigating these issues in a regression framework. Column 1 reports the results from estimating equation (1), including a variable that captures when in the meeting each household was randomized to be ranked. Specifically, we include a variable that varies from 0 (household was ranked first) to 1 (the household was ranked last). The point estimate is positive, indicating a higher mis-targeting rate for households ranked later, but it is not statistically significant. In Column 2, we interact the order with the hybrid treatment. The results show that in the community treatment, there is substantially more mis-targeting at the end of the process: the first household ranked is 5.9 percentage points less likely to be mis-targeted than the last household ranked (p-value 0.11). This effect is completely undone in the hybrid, where there is no systematic relationship between the rank order and the mis-targeting rate.

Columns 5 and 6 examine how the rank order affects whether a household receives the transfer. The results show that on average, households ranked at the end of the meeting are 4.9 percentage points more likely to be on the beneficiary list than those ranked at the start (significant at 10% level). The additional mis-targeting from being late in the list thus comes largely from rich households ranked late being more likely to be on the list.

The second treatment we investigated was the “10 poorest” framing treatment. The 10 poorest treatment was aimed at having the community focus on talking about the poorest households at the start of the meeting, before possible fatigue set in. As shown in Columns 3 and 4, the 10 poorest framing treatment did not substantially affect mis-targeting.

The results presented here provide some evidence that effort is important in the community method. Specifically, for households ranked early in the process, the community targeting method was identical to (or perhaps even slightly better than) the PMT, whereas PMT worked substantially better for those ranked later in the process.

VII. Does the Community Have a Different Maximand?

A third potential reason why community rankings could produce different outcomes than the PMT is that the community is actually identifying those who they believe to be poor, but that they have a different view of what constitutes poverty than the government, i.e., $u_g \neq u_c$. For example, while the PMT is based on per capita consumption, communities might recognize that children need less consumption than adults and that there are economies of scale in household production, and make their rankings accordingly (Olken, 2006). Thus, while the PMT predicts per capita consumption more accurately, the community may be identifying those who are really poor in a welfare sense. The community's social welfare function could differ from the government's in other ways as well: for example, the community might be more sympathetic towards some forms of poverty (bad luck) than others (laziness).

VII.A. Alternative welfare metrics

We begin by examining how the targeting outcomes compare not just against the government's metric of welfare u_g (captured by $r_{g,}$, the ranking based on per-capita consumption), but also against other welfare metrics. In our baseline survey, we asked eight randomly chosen members of the community to confidentially rank each other from poorest to richest. We average the ranks to construct each household's wealth rank according to the other community members, denoted r_c. To capture welfare as measured from an elite perspective, denoted r_e, we examine how the sub-village head ranked these eight other households. To measure how people assess their own poverty, denoted r_i, we asked all respondents to rate their own poverty level on a scale of 1 to 6. We computed the percentile rank of each measure in order to put them on the same scale.

Table 9 presents the matrix of rank correlations between these alternative welfare metrics. The correlation matrix shows that while all of the welfare metrics are positively correlated, they clearly pick up different things. Of particular note is the bottom row, which shows the correlations with self-assessments. While the rank correlation of self-assessments (r_i) with consumption (r_g) is only 0.26, the rank correlation of self-assessments with community survey ranks (r_c) is 0.45 and the correlation with the sub-village head survey ranks (r_e) is 0.41. This suggests that the community and sub-village head ranks may capture how individuals feel about themselves more accurately than per capita consumption.

To assess the poverty targeting results against these alternative welfare metrics, we compute the rank correlation between targeting rank list derived from the experiment and each of four welfare metrics. We then examine the effectiveness of the various targeting treatments against these different measures of well-being by estimating:

$$\text{RANKCORR}_{vkw} = \alpha + \beta_1 \text{ COMMUNITY}_{vk} + \beta_2 \text{ HYBRID}_{vk} + \gamma_k + \varepsilon_{vkw} \quad (3)$$

where RANKCORR_{vkw} is the rank correlation between the targeting rank list and the well-being measure w in sub-village v. Stratum fixed effects (γ_k) are included. The results are reported in Table 10. As the data is aggregated to the village level, each regression has 640 observations.[35]

[35] Self-assessments have 637 observations due to non-response on the self-assessment question in several villages.

The results provide striking evidence that the community holds a different view of welfare than per-capita consumption. Column 1 confirms the mis-targeting results shown in Table 4: both the community and hybrid treatment result in lower rank correlations with per-capita consumption than the PMT. Specifically, they are 6.5-6.7 percentage points, or about 14 percent, lower than the rank correlations obtained with PMT. However, they move away from consumption in a very clear direction – the community treatment increases the rank correlation with r_c by 24.6 percentage points, or 49 percent above the PMT level. The hybrid also increases the correlation with r_c but the magnitude is about half that of the community treatment. Thus, the verification in the hybrid appears to move the final outcome away from the community's perception of well-being. These differences are statistically significant at the 1 percent level. Results using the rank list obtained in the survey from the sub-village head (our measure of u_e) are virtually identical to the survey list obtained by the community, and are also statistically significant at the 1 percent level, which provides further evidence that the community at large and the elite share broadly similar assessments of welfare.

Perhaps most important, we examine how the different targeting methods affected the correlation between the targeting rank list and individuals' perceptions of their own welfare (r). The community treatment increases the rank correlation between the treatment outcomes and the individual self-assessments by 10.2 percentage points, or about 30 percent of the level in the PMT (significant at 1 percent). The hybrid treatment increases the rank correlation with the self-assessments by 7.5 percentage points. Thus, the community targeting methods are more likely to identify as poor those who actually self-identify as poor.

VI.B. Broadly shared preferences?

The results above suggest that the ranking exercise moves the targeting process towards a welfare metric identified by community members. An important question is the degree to which there is agreement among various subgroups of the community as to whom is poor. We designed one of the experimental sub-treatments to get at precisely this question. In Table 10, we test the effect of changing the composition of the meeting. Specifically, we explore the effect of having the meeting during the day, when women are more likely to be able to attend. We also consider the other sub-treatments (elite and 10 poorest) in this analysis, as they could also plausibly have affected the welfare weights of those at the meeting.

Columns 1 and 3 investigate the impact of having a day-time meeting on attendance. Columns 1 and 2 show that having a daytime meeting does not change the overall share of households in the village represented at the meeting. However, Column 3 confirms that the share of households represented by women increases by 11 percentage points in the day meetings, about a 49 percent increase over meetings held at night.

Although the day meeting treatment affected the gender composition of the meetings, Columns 4 - 8 show that the day meeting treatment did not affect the targeting outcomes. The elite treatment also did not affect the rank correlations with any of the various welfare metrics. Interestingly, the only sub-treatment that affected the rank correlations was the 10 poorest treatment, which increased the correlation of the treatments with ranks from self assessments. Overall, the findings suggest that community-based targeting is not sensitive to who exactly within the community is doing the ranking.

VII.C. What is the community maximizing?

Thus far, the evidence suggests that the community has a systematic, and broadly shared, notion of welfare which is not per-capita consumption – and that the community-based targeting methods reflect this different community concept of welfare. This raises two key questions: in what ways does the community welfare function differ from per-capita consumption? And, do the differences between community targeting and PMT methods simply reflect the differences in the social welfare function used, or does the community weigh the welfare of some households more than others due to other social or political considerations– i.e., are the differences in targeting just because $r_g \neq r_c$, or is it also because $\lambda_g \neq \lambda_c$?

Table 12 explores the relationship between the welfare metrics (community survey rank r_c, elite survey rank r_e, and self-assessment rank r_s), the targeting results in PMT, community, and hybrid villages, and a variety of household characteristics that might plausibly affect either welfare functions (u) or the social welfare weights used in targeting (λ). In Columns 1 - 3, we present results where the dependent variable is the within-village rank of each household in our survey according to different survey-based welfare metrics. In Columns 4 – 6, the dependent variable is the treatment rank, put on a corresponding metric where the lowest ranked (poorest) household in our dataset in each village is ranked 0 and the highest ranked (richest) household in our dataset in each village is ranked 1.[36] We control for log per capita consumption in all regressions, so the coefficients should be interpreted as conditional on per-capita consumption.

We find several dimensions on which the utility function used by villagers appears different from the PMT. First, we find adjustments for equivalence scales. The PMT in our

[36] Note that some of the variables included as explanatory variables – including household size, share of kids, household head education, and widowhood – were explicitly included in the PMT regression, which may explain why some of these variables are significant predictors of targeting in the PMT regressions.

setting is explicitly defined using per-capita consumption. Thus, it makes no adjustment for economies of scale in the household. By contrast, all of the community welfare functions (Columns 1-3) reveal that the community believes that there are household economies of scale, so that conditional on per-capita consumption, those in larger households are considered to have higher welfare (as in Olken 2005). Likewise, the same is true for the community ranking – which assigns almost an identical household size premium (Column 5). Interestingly, for a given household size and consumption, all methods rank households with more kids as poorer, even though children generally cost less than adults (Deaton 1997).

Second, we find that the community utility functions – as well as the community targeting method – appear to make some adjustments for earning potential. For example, households in which the household head has a primary education or less rank 4-6 percentage points poorer, conditional on their actual consumption. Similarly, households headed by a widow, those with a disability, and those where there is a serious illness are all rated poorer, conditional on actual consumption. The adjustments for education and widowhood are also reflected in the community treatment ranking, although the disability adjustment is not (Column 5).[37]

Third, we might be concerned that community methods may exclude specific sub-groups of the population, such as ethnic or religious minorities.[38] There is no evidence that ethnic minorities are discriminated against. Indeed, ethnic minorities are more likely to be ranked as

[37] There are, of course, two interpretations of these findings. One interpretation is that households are conditioning on earnings ability – i.e., if you are highly educated but do not earn much, that is your fault and you should not receive subsidies for it. Another interpretation, however, is that education is merely another signal of poverty that is more easily observable to the community than actual consumption, though communities would need to be over-weighting this signal for this effect to produce a negative coefficient conditional on actual consumption.

[38] Note that since the sample area includes both Christian-majority areas (most of North Sumatra and parts of South Sulawesi) and Muslim-majority areas (everywhere else) as well as areas with many different majority ethnic groups (e.g., Javanese in Java, Batak in North Sumatra, and Bugis and Toraja in South Sulawesi), the definition of a minority varies from region to region.

poor in the community treatment, suggesting perhaps that extra care is paid to them in the interest of social harmony (Column 5). Instead, communities discriminate against elite connected households, which the community views as richer than they actually are – the community survey ranks place about a 10 percentage point premium on elite connectedness, and even the elite and self-assessed survey ranks place a 4.7 and 2.6 percentage point premium, respectively. The community treatment ranks place a 6.3 percentage point premium on elite connectedness.

VII.D. Information

Although the findings presented thus far suggest that the community has a systematically different view of poverty than per-capita consumption, it is also possible that the results differ because the community has different information than that obtained under the PMT.

Although we cannot definitively show that the community has all of the information available in the PMT, we can examine whether the community has additional information beyond that in the PMT. To do so, we estimate the following equation:

$$\text{RANKIND}_{ijvk} = \alpha + \beta_1 \text{ RANKCONSUMPTION}_{ivk} + \beta_2 \text{ RANKPMTSCORE}_{jvk} + v_j + \varepsilon_{ijvk} \qquad (4)$$

where RANKIND_{ijvk} is household *j*'s rank of household *i* (all ranks are in percentiles), $\text{RANKCONSUMPTION}_{ivk}$ is the rank of household *i*'s per capita consumption in village *v*, and $\text{RANKPMTSCORE}_{ivk}$ is the rank of household *i*'s PMT score. Fixed effects for the individual doing the ranking are included (v_j), and standard errors are clustered at the village level. The results of this analysis are in Column 1 of Table 12. In Column 2, we add every variable that enters the PMT formula separately.

Table 13 illustrates that the community has residual information. Consumption is still highly correlated with individuals' ranks of other households from the baseline survey even though include we include the rank according to the PMT Score. Controlling for PMT score, a one percentile increase in consumption rank is associated with a 0.132 percentile increase in individual household ranks of the community (Column 1). This is significant at the 1 percent level. In the more flexible specification presented in Column 2, the correlation between consumption rank and survey rank remains positive (0.088) and significant at the 1 percent level.

The findings in Table 13 suggest that the community has residual information about consumption beyond that contained in the PMT score or even in the PMT variables. Moreover, the fact that almost all the PMT variables enter into the community ranks with plausible signs and magnitudes suggests that the community has most of the information in the PMT as well, but chooses to aggregate it in different ways, as illustrated in Table 11. While we cannot completely rule out that there is some information available in the PMT that the community does not have, the evidence presented here suggests that differences in information are not the primary drivers of the different results.

Even if there are differences in information, it is not clear that they matter from a policy perspective. If the community achieves its goals – both in terms of matching community perceptions of welfare and in terms of satisfaction levels -- and especially if the excluded households do not see themselves as being poor (and who are not in any case desperately poor), the government might have reason to prefer the community treatment even if, in per-capita consumption terms, the allocation is objectively worse. This may be especially true in our setting because the differences are not huge and the cost of any social conflict or disgruntlement that could result from imposing an unwanted list on the community may easily outweigh the gain in accuracy.

VIII. Conclusion

The debate regarding decentralization in targeting is usually framed in terms of the benefits of local information versus the costs of some form of malfeasance, such as elite capture. While we started from an experiment which took both of these ideas seriously, our results point to a third factor as being very important: the community might have a widely shared objective function that the government does not necessarily share, and while satisfying that objective might lead to other distortions, having a say in decision making leads to widespread satisfaction in the community.

From a policy point of view, the results in this paper suggest that neither the PMT nor the community-based methods of targeting is strictly dominant. If the government aims to target the poor based on per-capita consumption, then the PMT – designed explicitly with that goal in mind – performed best. However, if the government is willing to accept the local community's own definitions of poverty, the community targeting methods were superior in terms of matching better to the local views of poverty and delivering much greater satisfaction and legitimacy. The hybrid method we examined delivered the worst of both worlds – poor targeting performance and low legitimacy. The hybrid might have performed badly because its main theoretical advantage—preventing elite capture—was not important in our setting, and it is possible that alternative hybrid designs that allow the community to add some very poor households to the PMT might perform better.

The findings in this paper raise several questions for further research. First, while we found little evidence of elite capture or general malfeasance of the targeting methods, it is possible that this might change over time as individuals learn to better manipulate the system. The idea that manipulation over time has been shown to occur in some kinds of PMT systems

(Camacho and Conover 2008), but whether it would occur when the per-village allocation is fixed, and whether it would be more or less severe in community-targeted systems, are important open questions. Second, given how well the community outcomes match individual self-assessments, an important question is whether some form of self-targeting system (perhaps connected to an ordeal mechanism as in Nichols and Zeckhauser 1982), could provide a more cost-effective method of targeting the poor. We regard these as important questions for future research.

Works Cited

Alderman, Harold, 2002. “Do local officials know something we don’t? Decentralization of targeted transfers in Albania.” *Journal of Public Economics 83 (2)*, pp 375-404.

Bardhan, Pranab and Dilip Mookherjee, 2005. “Decentralizing antipoverty program delivery in developing countries,” Journal of Public Economics 89 (4), pp 675-704.

Camacho, Adriana and Emily Conover (2008). “Manipulation of Social Program Eligibility: Detection, Explanations, and Consequences for Empirical Research” Mimeo, University of California, Berkeley.

Cameron, Lisa A., 2002. “Did social safety net scholarships reduce drop-out rates during the Indonesian economic crisis?” Policy Research Working Paper Series 2800, The World Bank.

Castañeda, Tarsicio and Kathy Lindert (2005). “Designing and Implementing Household Targeting Systems: Lessons from Latin America and the United States.” Social Protection Discussion Paper 526. The World Bank and DFID, Washington DC.

Coady, David, Margaret Grosh, and John Hoddinott (2004), “Targeting Outcomes Redux,” *World Bank Research Observer 19* (1), pp 61-85.

Conning, Jonathan and Michael Kevane. (2002). “Community-Based Targeting Mechanisms for Social Safety Nets: A Critical Review.”*World Development 30 (3)*, pp. 375-394.

Daly, Anne and George Fane (2002). “Anti-Poverty Programs in Indonesia.” *Bulletin of Indonesian Economic Studies,* Vol. 38, No. 3.

Galasso, Emanuela, and Martin Ravallion (2005). “Decentralized Targeting of an antipoverty program.” *Journal of Public Economics* 89.

Knuth, Donald (1998) *The Art of Computer Programming, Volume 3: Sorting and Searching, Second Edition*. Addison-Wesley. pp.80–105.

Olken, Benjamin A. (2005). “Revealed Community Equivalence Scales.” *Journal of Public Economics 89* (2-3), pp. 545-566.

Olken, Benjamin A. (2006). “Corruption and the Costs of Redistribution: Micro evidence from Indonesia.” *Journal of Public Economics* 90.

Olken, Benjamin A. (2008). “Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia.” NBER Working Paper #14123.

Perdana, Ari and John Maxwell (2004). “Poverty Targeting in Indonesia: Programs, Problems, and Lessons Learned.” *CSIS Economic Working Paper Series.* WPE083, Centre for Strategic and International Studies, Jakarta, Indonesia.

Platteau, Jean-Philippe (2004). “Monitoring Elite Capture in Community-Driven Development.” *Development and Change 35 (2)*, pp 223-246.

Pritchett, Lant; Sudarno Sumarto; and Asep Suryahadi (2005). "Targeted Programs in an Economic Crisis: Empirical Findings from the Experience of Indonesia." Governance Working Papers 84, East Asian Bureau of Economic Research.

"Rapid Appraisal Implementation of the Direct Cash Transfer Program in Indonesia 2005: Primary Findings and Recommendations" (2005). SMERU Research Team, Jakarta, Indonesia.

Sumarto, Sudarno; Asep Suryahadi; and Wenefrida Widyanti (2005). "Assessing the Impact of Indonesian Social Safety Net Programmes on Household Welfare and Poverty Dynamics." *European Journal of Development Research,* Vol. 17, No.1.

"Unconditional Cash Transfer Rapid Assessment" (2005). Conducted by LP3ES Institute for Social and Economic Research, Education, and Information, Jakarta, Indonesia.

Van de Walle, Dominique (1998). "Targeting Revisited." *World Bank Research Observer* 13 (2).

The World Bank (2008), *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program,* Washington, DC.

The World Bank (2006). "Making the New Indonesia Work for the Poor."

Table 1: Randomization Design

Community/Hybrid Sub-Treatments			Main Treatments		
			Community	Hybrid	PMT
Elite	10 Poorest First	Day	24	23	
		Night	26	32	
	No 10 Poorest First	Day	29	20	
		Night	29	34	
Whole Community	10 Poorest First	Day	29	28	
		Night	29	23	
	No 10 Poorest First	Day	28	33	
		Night	20	24	
		TOTAL	214	217	209

Notes: This table shows the results of the randomization. Each cell reports the number of sub-villages randomized to each combination of treatments. Note that the randomization of sub-villages into main treatments was stratified to be balanced in each of 51 strata. The randomization of community and hybrid subvillages into each sub-treatment (elite or full community, 10 poorest prompting or no 10 poorest prompting, and day or night) was conducted independently for each sub-treatment, and each randomization was stratified by main treatment and geographic stratum.

Table 2: Summary Statistics

Variable	Obs	Mean	Std. Dev.
Panel A: Consumption from baseline survey			
Per capita consumption (Rp. 000s)	5753	557.501	602.33
Panel B: Mis-targeting variables:			
On beneficiary list	5756	0.30	0.46
Mis-target	5753	0.32	0.47
Mis-target -- nonpoor (rich + middle)	3724	0.20	0.40
Mis-target -- poor (near + very poor)	2029	0.53	0.50
Mis-target -- rich	1840	0.14	0.35
Mis-target -- midddle income	1884	0.26	0.44
Mis-target -- near poor	1076	0.58	0.49
Mis-target -- very poor	953	0.46	0.50
Panel C: Rank correlations between treatment results and…			
Per capita consumption	640	0.41	0.34
Community (excluding sub-village head)	640	0.64	0.33
Sub-village Head	640	0.58	0.41
Self-Assessment	637	0.40	0.34

Table 3: Testing Balance Between Treatment Groups

	Means			Differences, No Fixed Effects			Differences, Controlling for Stratum Fixed Effects		
	PMT	Community	Hybrid	Community - PMT	Hybrid - PMT	Hybrid - Community	Community - PMT	Hybrid - PMT	Hybrid - Community
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Average per capita expenditure (Rp. 000s)	558.576	550.579	564.295	-7.997	5.719	13.716	-1.331	11.980	13.312
	[245.845]	[220.237]	[337.172]	(22.728)	(28.535)	(27.416)	(20.661)	(25.973)	(24.913)
Average years of education of household	7.360	7.566	7.087	0.206	-0.273	-0.4785*	0.219	-0.255	-0.4739**
head among survey respondents	[2.616]	[2.644]	[2.627]	(0.256)	(0.254)	(0.254)	(0.204)	(0.200)	(0.209)
PMT score	12.467	12.519	12.474	0.052	0.007	-0.045	0.053	0.011	-0.043
(calculated from Baseline survey)	[0.436]	[0.414]	[0.423]	(0.041)	(0.042)	(0.040)	(0.037)	(0.037)	(0.037)
Pct. of households that are agricultural	45.827	42.887	48.438	-2.940	2.612	5.5515*	-3.7806*	1.264	5.0442**
	[34.889]	[33.789]	[35.038]	(3.343)	(3.391)	(3.318)	(2.060)	(2.096)	(2.027)
Years of education of RT head	8.856	8.860	8.604	0.003	-0.253	-0.256	0.033	-0.206	-0.238
	[4.018]	[4.244]	[3.796]	(0.402)	(0.379)	(0.388)	(0.352)	(0.336)	(0.335)
Log number of households	3.832	3.895	3.810	0.063	-0.022	-0.0853*	0.057	-0.028	-0.0846**
	[0.491]	[0.489]	[0.460]	(0.048)	(0.046)	(0.046)	(0.044)	(0.043)	(0.041)
Distance to kecamatan in km	0.444	0.416	0.482	-0.028	0.039	0.067	-0.029	0.038	0.0673*
	[0.652]	[0.473]	[0.431]	(0.056)	(0.054)	(0.044)	(0.050)	(0.046)	(0.037)
Log size of villages in heactares	3.105	3.271	3.282	0.166	0.177	0.011	0.1435*	0.1376*	-0.006
	[1.278]	[1.197]	[1.187]	(0.121)	(0.120)	(0.115)	(0.075)	(0.075)	(0.076)
Religious building per household	0.0070	0.0060	0.0060	-0.0004	-0.0004	-0.0001	-0.0004	-0.0005	-0.0001
	[0.0050]	[0.0050]	[0.0050]	(0.0005)	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0003)
Primary school per household	0.0030	0.0030	0.0030	0.0001	-0.0002	-0.0003	0.0000	-0.0002	-0.0003
	[0.0030]	[0.0030]	[0.0020]	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
P-value from joint test				0.275	0.689	0.089	0.165	0.322	0.028
Average per capita expenditure (Rp. 000s)	558.576	550.579	564.295	-7.997	5.719	13.716	-1.331	11.980	13.312

Notes: Standard deviations in brackets in columns (1) – (3); robust standard errors are shown in parentheses in columns (4) – (9).

Table 4: Results of Different Targeting Methods on Mis-targeting Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		By Income Status		By Detailed Income Status				
Sample:	Full population	Non-poor	Poor	Rich	Middle income	Near Poor	Very Poor	Per-capita consumption of beneficiaries
Community treatment	0.031*	0.045**	0.022	0.028	0.066**	0.052	-0.019	9.933
	(0.017)	(0.018)	(0.028)	(0.021)	(0.027)	(0.037)	(0.039)	(18.742)
Hybrid treatment	0.029*	0.037**	0.010	0.020	0.052**	0.036	-0.013	-1.155
	(0.016)	(0.017)	(0.027)	(0.020)	(0.025)	(0.037)	(0.037)	(19.302)
Observations	5753	3724	2029	1840	1884	1076	953	1719
Mean in PMT treatment	0.30	0.18	0.52	0.13	0.23	0.55	0.49	366

Notes: All regressions include stratum fixed effects. Robust standard errors are shown in parentheses, adjusted for clustering at the village level. All coefficients are interpretable relative to the PMT treatment, which is the omitted category. The mean of the dependent variable in the PMT treatment is shown in the bottom row. All specifications include stratum fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Satisfaction

	(1)	(2)	(3)	(4)	(5)	(7)
	Panel A: Household Endline Survey					
	Is the method applied to determine the targeted households appropriate? (1=worst,4=best)	Are you satisfied with P2K08 activities in this sub-village in general? (1=worst,4=best)	Are there any poor HH which should be added to the list? (0=no, 1 = yes)	Number of HH that should be added from list	Number of HH that should be subtracted from list	P-value from joint test
Community treatment	0.161***	0.245***	-0.189***	-0.578***	-0.554***	<0.001
	(0.056)	(0.049)	(0.040)	(0.158)	(0.112)	
Hybrid treatment	0.018	0.063	0.020	0.078	-0.171	0.762
	(0.055)	(0.049)	(0.042)	(0.188)	(0.129)	
Observations	1089	1214	1435	1435	1435	
Mean in PMT treatment	3.243	3.042	0.568	1.458	0.968	
	Panel B: Sub-village Head Endline Survey					
	Is the method applied to determine the targeted households appropriate? (0=no, 1=yes)	In your opinion, are villagers satisfied with P2K08 activities in this sub-village in general? (1=worst,4=best)	Are there any poor HH which should be added to the list? (0=no, 1=yes)	Are there any poor HH which should be subtracted from the list? (0=no, 1=yes)		
Community treatment	0.378***	0.943***	-0.169***	-0.010		<0.001
	(0.038)	(0.072)	(0.045)	(0.020)		
Hybrid treatment	0.190***	0.528***	-0.065	-0.019		<0.001
	(0.038)	(0.071)	(0.043)	(0.019)		
Observations	636	629	640	640		
Mean in PMT treatment	0.565	2.456	0.732	0.057		
	Panel C: Comment forms and fund disbursement results					
	Number of comments in the comment box	Number of complaints in the comment box	Number of complaints received by sub-village head	Did facilitator encounter any difficulty in distributing the funds? (0=no, 1=yes)	Fund distributed in a meeting (0=no, 1=yes)	
Community treatment	-0.944	-1.772***	-2.684***	-0.062***	0.082**	
	(0.822)	(0.461)	(0.530)	(0.023)	(0.038)	
Hybrid treatment	-0.364	-0.927**	-2.010***	-0.045*	0.051	
	(0.821)	(0.467)	(0.529)	(0.026)	(0.038)	
Observations	640	399	640	621	614	
Mean in PMT treatment	11.392	2.787	4.34	0.135	0.082**	

Notes: All estimation is by OLS with stratum fixed effects; using ordered probit for multiple response and probit models for binary dependent variables produces the same signs and statistical significance as the results shown.

Table 6: Elite Treatments

	(1)	(2)	(3)	(4)	(5)	(6)
	Attendance (Meeting Data)		Attendance (Survey Data)		Mis-target Dummy	
Community treatment			0.367***	0.360***	0.029	0.043**
			(0.038)	(0.044)	(0.018)	(0.020)
Hybrid treatment	0.021	0.019	0.370***	0.378***	0.027	0.013
	(0.029)	(0.036)	(0.037)	(0.043)	(0.018)	(0.020)
Elite sub-treatment	-0.062**	-0.064	-0.301***	-0.287***	0.004	-0.024
	(0.029)	(0.041)	(0.034)	(0.050)	(0.016)	(0.023)
Elite × hybrid		0.004		-0.029		0.057*
		(0.058)		(0.068)		(0.032)
Observations	431	431	287	287	5756	5756
Mean in PMT treatment	N/A	N/A	0.11	0.11	0.30	0.30

Notes: In columns (1) – (4), an observation is a village. In columns (1) and (2), the dependent variable is the number of households attending the meeting (as observed on the meeting attendance list) divided by the number of households in the village. In columns (3) and (4), the dependent variable is the share of households surveyed in the endline survey where at least 1 household member attended a targeting meeting. The PMT mean for in columns (3) and (4) is not zero, because the question was worded generically to be about any targeting meeting, not just meetings associated with our project. The dependent variable in column (5) and (6) is the mis-targeting dummy, as in column (1) of Table 4. Robust standard errors in parentheses, and standard errors are adjusted for clustering at the village level in columns (5) and (6). All specifications include stratum fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Are Households Connected to Elites Treated Differently?

	(1)	(2)	(3)	(4)
	Mis-target dummy		On beneficiary list dummy	
Elite connectedness	-0.025	-0.025	-0.063***	-0.063***
	(0.021)	(0.021)	(0.021)	(0.021)
Elite connectedness	-0.015	-0.013	-0.067**	-0.078**
× community treatment	(0.035)	(0.038)	(0.033)	(0.036)
Elite connectedness	0.010	0.010	-0.013	-0.001
× hybrid treatment	(0.033)	(0.035)	(0.033)	(0.035)
Elite connectedness	-0.029	-0.034	0.041	0.064
× elite treatment	(0.031)	(0.047)	(0.030)	(0.042)
Elite connectedness		0.003		-0.047
× elite treatment × hybrid		(0.063)		(0.060)
Observations	5753	5753	5756	5756

Notes: All specifications include dummies for the community, hybrid, and elite treatment main effects, as well as stratum fixed effects; columns (2) and (4) also include a dummy for elite × hybrid. Robust standard errors in parentheses, adjusted for clustering at the village level. Dependent variable in columns (1) and (2) is the mis-target dummy for the full sample, as in column (1) of Table 4. Dependent variable in columns (3) and (4) is a dummy for being a beneficiary of the program. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Effort

	(1)	(2)	(3)	(4)	(5)	(6)
		Mis-target dummy			On beneficiary list dummy	
Household order in ranking	0.030	0.059			0.049*	0.048*
(percentile)	(0.026)	(0.037)			(0.026)	(0.029)
Household order in ranking		-0.056				0.001
× hybrid		(0.052)				(0.028)
Poorest 10 framing sub-treament			-0.006	-0.007		
			(0.016)	(0.023)		
Poorest 10 framing sub-treatment				0.002		
× hybrid				(0.031)		
Observations	3784	3784	3874	3874	3785	3785

Notes: All specifications are limited to community and hybrid villages. Columns (1) – (4) include a hybrid dummy and stratum fixed effects; columns (5) and (6) include stratum fixed effects since the total number of beneficiaries is constant in all treatments. The dependent variable in columns (1) – (4) is the mis-target dummy for the full sample, as in column (1) of Table 4. The dependent variable in columns (5) and (6) is a dummy for being chosen as a recipient, as in column (3) of Table 6.*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 9: Rank correlation matrix of alternative welfare metrics

	(1)	(2)	(3)	(4)
	Consumption (u_g)	Community survey ranks (u_c)	Sub-village head survey ranks(u_e)	Self-Assessment (u_s)
Consumption (u_g)	1.000			
Community survey ranks (u_c)	0.376	1.000		
Sub-village head survey ranks(u_e)	0.334	0.737	1.000	
Self-Assessment(u_s)	0.263	0.445	0.407	1.000

Notes: This table reports the correlation matrix between the within-village ranks of the four variables listed. All correlations are statistically significantly different from 0 at the 1% level.

Table 10: Assessing targeting treatments using alternative welfare metrics

	(1)	(2)	(3)	(4)
	Consumption (u_g)	Community survey ranks (u_c)	Sub-village head survey ranks(u_e)	Self-Assessment (u_s)
Community treatment	-0.065**	0.246***	0.248***	0.102***
	(0.033)	(0.029)	(0.038)	(0.033)
Hybrid treatment	-0.067**	0.143***	0.128***	0.075**
	(0.033)	(0.029)	(0.038)	(0.033)
Observations	640	640	640	637
Mean in PMT treatment	0.451	0.506	0.456	0.343

Notes: The dependent variable is the rank correlation between the treatment outcome (i.e., the rank ordering of households generated by the PMT, community, or hybrid treatment) and the welfare metric shown in the column, where each observation is a village. Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Do community meetings reflect broadly shared preferences?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Rank Correlations with:			
	Attend Meeting (Meeting Data)	Attend Meeting (HH Data)	Female Attends (Meeting Data)	Mis-target	Consumption	Community (excluding sub-village head)	Sub-village Head	Self-Assessment
Community treatment		0.349***		0.027	-0.089**	0.232***	0.180***	0.072
		(0.042)		(0.021)	(0.045)	(0.040)	(0.052)	(0.044)
Hybrid treatment	0.020	0.353***	0.008	0.026	-0.089**	0.130***	0.064	0.046
	(0.029)	(0.041)	(0.017)	(0.021)	(0.044)	(0.039)	(0.051)	(0.044)
Day meeting treatment	-0.021	0.013	0.104***	0.008	0.019	0.004	0.055	0.014
	(0.029)	(0.033)	(0.017)	(0.016)	(0.033)	(0.029)	(0.038)	(0.033)
Elite treatment	-0.064**	-0.300***	-0.085***	0.005	-0.004	-0.023	0.034	-0.017
	(0.029)	(0.033)	(0.017)	(0.016)	(0.033)	(0.029)	(0.038)	(0.033)
10 Poorest treatment	0.022	0.023	-0.010	-0.006	0.031	0.047	0.044	0.062*
	(0.029)	(0.034)	(0.018)	(0.016)	(0.033)	(0.029)	(0.038)	(0.032)
Observations	431	287	428	5753	640	640	640	637
Mean in PMT treatment		0.110		0.300	0.451	0.506	0.456	0.343

Notes: For columns (1) and (2), see notes to Table 5; for column (4), see notes to Table 5; for columns (5) – (8), see notes to Table 8. For column (3), the dependent variable is the percentage of households in the village in which a female attends the meeting, using data collected from the meeting attendance lists. *** p<0.01, ** p<0.05, * p<0.1

Table 12: What is the community maximizing?

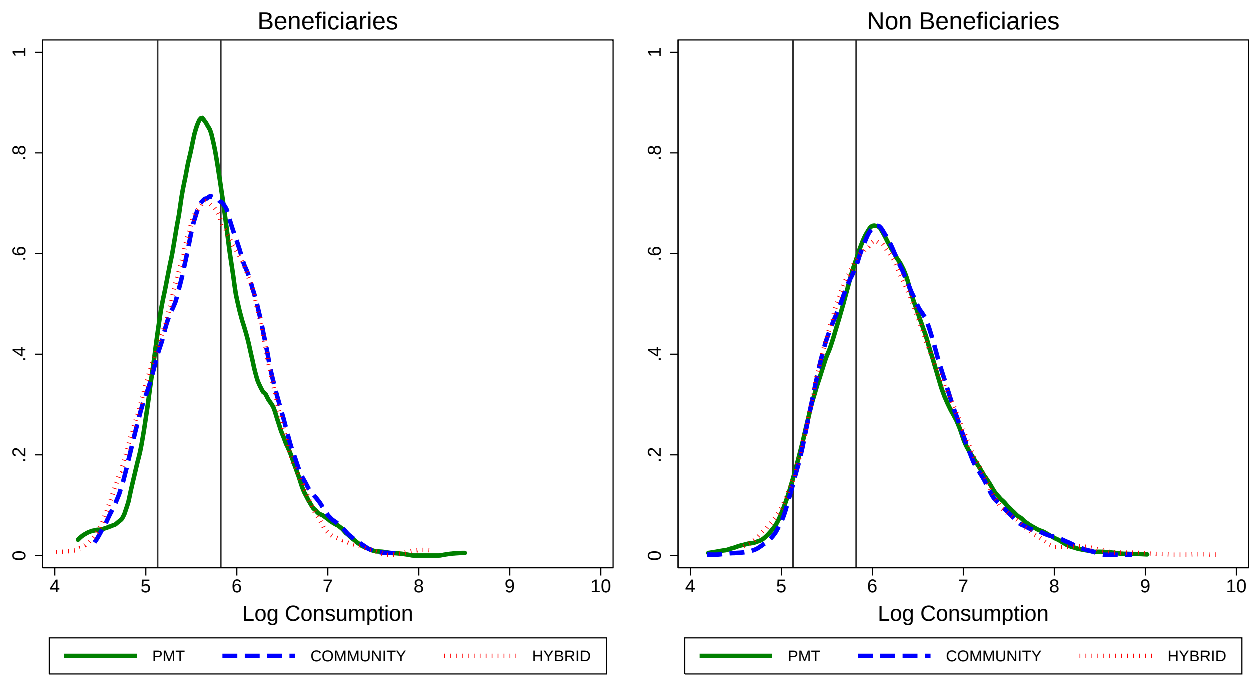
	Rank according to welfare metric…			Targeting Rank List in…		
	Community survey ranks (u_c)	Sub-village head survey ranks(u_e)	Self-Assessment (u_s)	PMT villages	Community villages	Hybrid villages
	(1)	(2)	(3)	(4)	(5)	(6)
Log PCE	0.201***	0.159***	0.100***	0.162***	0.221***	0.181***
	(0.007)	(0.007)	(0.004)	(0.011)	(0.012)	(0.013)
Log HH size	0.177***	0.144***	0.081***	-0.012	0.159***	0.092***
	(0.010)	(0.010)	(0.006)	(0.018)	(0.018)	(0.020)
Share kids	-0.136***	-0.105***	-0.043***	-0.315***	-0.053	-0.156***
	(0.021)	(0.021)	(0.012)	(0.034)	(0.041)	(0.039)
HH head with primary education or less	-0.056***	-0.045***	-0.051***	-0.133***	-0.042**	-0.096***
	(0.009)	(0.009)	(0.005)	(0.016)	(0.018)	(0.018)
Elite connected	0.094***	0.047***	0.026***	0.066***	0.063***	0.040***
	(0.008)	(0.008)	(0.005)	(0.015)	(0.014)	(0.014)
Ethnic minority	-0.019	-0.015	-0.000	0.022	-0.046*	-0.005
	(0.015)	(0.015)	(0.008)	(0.027)	(0.026)	(0.025)
Religious minority	0.013	-0.010	-0.013	-0.019	0.035	0.014
	(0.018)	(0.017)	(0.008)	(0.029)	(0.033)	(0.031)
Widow	-0.093***	-0.076***	-0.007	0.024	-0.103***	-0.026
	(0.015)	(0.014)	(0.008)	(0.028)	(0.023)	(0.028)
Disability	-0.050***	-0.045***	-0.029***	-0.091***	0.000	0.008
	(0.016)	(0.014)	(0.008)	(0.027)	(0.028)	(0.027)
Death	-0.035	-0.027	-0.008	-0.112***	-0.010	-0.045
	(0.025)	(0.024)	(0.015)	(0.041)	(0.047)	(0.043)
Sick	-0.041***	-0.042***	-0.030***	0.005	-0.022	-0.048**
	(0.011)	(0.011)	(0.006)	(0.019)	(0.019)	(0.019)
Recent shock to income	-0.004	-0.005	-0.015***	-0.021	0.002	-0.012
	(0.009)	(0.009)	(0.005)	(0.016)	(0.016)	(0.017)
Observations	5344	4686	5731	1814	1881	1891

Notes: Note that the children and household head education variables are explicitly included in the PMT regression (see Table 12). The PMT regression also includes a dummies for the household head being male, married, and male * married, which together will be closely correlated with the widow variable.

Table 13: Information

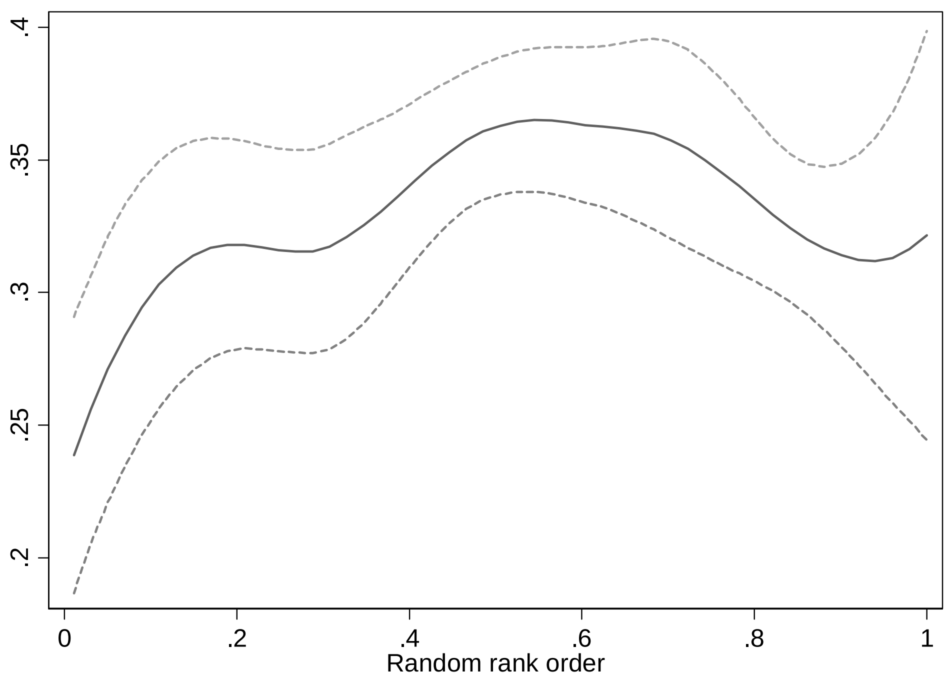
	Survey rank (1)	Survey rank (2)		Survey rank (2 continued)
Rank per capita consumption within village in percentiles	0.132*** (0.014)	0.088*** (0.012)		
Rank per capita consumption from PMT within village in percentiles	0.368*** (0.014)			
Household floor area Per capita		0.001*** 0.000	Has this Household ever got credit?	0.027** (0.011)
Not earth floor		0.060*** (0.010)	Number of children 0-4	0.000 (0.006)
Brick or cement wall		0.065*** (0.007)	Number of Children in Elementary School	0.003 (0.005)
Private toilet		0.047*** (0.008)	Number of Children in Junior High School	0.007 (0.007)
Clean drinking water		0.008 (0.009)	Number of Children in Senior High School	0.022*** (0.008)
PLN electricity		0.064*** (0.008)	Highest Education Attainment within HH is Elem. School	0.007 (0.016)
Concrete or corrugated roof		0.027* (0.014)	Highest Education Attainment within HH is Junior School	0.01 (0.016)
Cooks with firewood		0.031*** (0.008)	Highest Education Attainment within HH is Senior High or higher	0.051*** (0.017)
Own house privately		0.034*** (0.008)	Total Dependency Ratio	0.004 (0.006)
Household size		0.004 (0.006)	AC	0.049** (0.023)
Household Size Squared		-0.001 (0.001)	Computer	0.045*** (0.011)
Age of head of household		0.011*** (0.002)	Radio / Cassette Player	0.001 (0.006)
Age of head of household squared		-0.000*** 0.000	TV	0.043*** (0.010)
Head of household is Male		0.047** (0.019)	DVD/VCD player	0.017** (0.007)
Head of household is married		0.119*** (0.022)	Satelite dish	0.021* (0.011)
Head of household is male and Married		-0.043* (0.026)	Gas burner	0.030*** (0.008)
Head of household works in agriculture sector		-0.006 (0.041)	Refrigerator	0.069*** (0.008)
Head of household works in industry Sector		-0.043 (0.042)	Bicycle	-0.004 (0.007)
Head of household works in service Sector		-0.018 (0.042)	Motorcycle	0.078*** (0.007)
Head of household works in formal sector		0.071 (0.045)	Car / Mini-bus / Truck	0.116*** (0.012)
Head of household works in informal sector		0.048 (0.045)	HP	0.014* (0.007)
Education Attainment of HH Head is Elementary School		0.008 (0.008)	Jewelry	0.034*** (0.006)
Education Attainment of HH Head is Junior School		0.036*** (0.010)	chicken	-0.001 (0.006)
Education Attainment of HH Head is Senior High School or higher		0.041*** (0.011)	Caribou / Cow	0.065*** (0.012)
Observations	40398	38336		

Figure 1: PDF of log per-capita consumption of beneficiaries and non-beneficiaries, by treatment status



Notes: The left panel shows the PDF of log per-capita consumption for those households chosen to receive the transfer, separately by each treatment. The right panel shows the PDF of log per-capita consumption for those households not chosen to receive the transfer, separately by treatment. The vertical lines show the PPP$1 and PPP$2 per day poverty lines (see footnote for more information on the calculation of these poverty lines.)

Figure 2: Effect of order in ranking meeting on mis-target rate



Notes: This figure graphs the relationship between mis-targeting and the randomized rank order from a non-parametric Fan regression. The dashed lines represent cluster-bootstrapped 95th percentile confidence intervals.