Vivi Alatas
Abhijit Banerjee
Rema Hanna
Ben Olken
Matt Wai-poi
Ririn Purnamasari

Targeting the poor
Evidence from a field experiment in Indonesia

March 2014

Impact Evaluation Report 12

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) was set up in 2008 to meet growing demand for more and better evidence of what development interventions in low- and middle-income countries work and why. By funding rigorous impact evaluations and systematic reviews and by making evidence accessible and useful to policymakers and practitioners, 3ie is helping to improve the lives of people living in poverty.

3ie Impact Evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of this report, *Targeting the poor: evidence from a field experiment in Indonesia*, as partial fulfilment of requirements under grant OW3.1055, issued under Open Window 3. The content has been copyedited and formatted for publication by 3ie. Due to unavoidable constraints at the time of publication, a few of the tables or figures may be less than optimal. All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its Board of Commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Rema Hanna, at rema_hanna@ksg.harvard.edu.

Funding for this impact evaluation was provided by 3ie's donors, which include UKaid, the Bill & Melinda Gates Foundation, Hewlett Foundation and 12 other 3ie members that provide institutional support. A complete listing is provided on the 3ie website at www.3ieimpact.org/en/about/3ie-affiliates/3ie-members.

Suggested citation: Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014. *Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12.* New Delhi: International Initiative for Impact Evaluation (3ie).

3ie Impact Evaluation Report series executive editors: Jyotsna Puri and Beryl Leach
Managing editors: Stuti Tripathi and Thomas de Hoop
Assistant managing editor: Kanika Jha
Production manager: Lorna Fray
Assistant production manager: Rajesh Sharma
Copy editor: Lucy Southwood
Proofreader: Sarah Chatwin
Cover design: John F McGill
Printer: VIA Interactive
Cover photo: Jurist Tan

Targeting the poor: evidence from a field experiment in Indonesia

Vivi Alatas
The World Bank

Abhijit Banerjee
Massachusetts Institute of Technology

Rema Hanna
Harvard University

Ben Olken
Massachusetts Institute of Technology

Matt Wai-poi
The World Bank

Ririn Purnamasari
The World Bank

March 2014
3ie Impact Evaluation Report 12

International Initiative for Impact Evaluation

Acknowledgements

This project was a collaboration involving many people. We thank Jie Bai, Talitha Chairunissa, Donghee Jo, Chaeruddin Kodir, He Yang, Ariel Zucker and Gabriel Zucker for their excellent research assistance. We thank Mitra Samya, the Indonesian Central Bureau of Statistics (BPS), the Indonesian National Team for the Acceleration of Poverty Reduction (TNP2K, particularly Sudarno Sumarto and Bambang Widianto), the Indonesian Ministry of Social Affairs (DepSos), and SurveyMeter for their cooperation implementing the project. Most of all, we thank Jurist Tan for her truly exceptional work leading the field implementation. This project was financially supported by the World Bank, AusAID and 3ie, and analysis was supported by National Institutes of Health (NIH) under grant P01 HD061315.

Abstract

Governments of developing countries often lack verifiable income information for poor people and communities. This makes targeting for social programmes a challenge. This report provides results from a randomised control trial that was designed to better understand how to improve targeting in Indonesia. Specifically, during the expansion of Indonesia's real conditional cash transfer programme, Program Keluarga Harapan (PKH), we randomised three different targeting methodologies — proxy means testing, self-targeting and community targeting – across 600 villages. We found that, when poverty is defined by consumption, self-targeting identifies poorer beneficiaries than proxy means testing and it has lower administrative costs. Community targeting is less effective than proxy means testing in identifying the poor based on *per capita* consumption, but it results in higher satisfaction levels with the programme.

Contents

Abstract iii

List of figures and tables v

Abbreviations and acronyms vi

1. Background: targeting social programmes and principal interventions 1

2. Experimental design and data 3

2.1 Setting: the PKH programme 3
2.2 Sample selection 3
2.3 Experimental design 5
2.4 Randomisation design and timing 8
2.5 Power calculations 8
2.6 Quantitative and qualitative data collection 8

3. Results 9

3.1 Balance check 9
3.2 Accuracy and perceptions 11
3.3 Alternative measures of wealth 15
3.4 Subtreatments 15
3.5 Treatment heterogeneity 17
3.6 Hypothetical universal PMT 19
3.7 Elite capture 22
3.8 Observables and unobservables in self-selection 24
3.9 Impacts on poverty rates 24

4. Programme costs and policy conclusions 27

4.1 Programme costs 27
4.2 Policy conclusions 28

References 30

List of figures and tables

Figure 1 Study area 4
Figure 2 Consumption of beneficiaries under different treatments 13

Table 1 Experimental design 5
Table 2 Balance check 10
Table 3 Effect of treatments on targeting accuracy 12
Table 4 Perceptions of different targeting mechanisms (scaled 0–1) 14
Table 5 Targeting accuracy using subjective measures of wealth and welfare 15
Table 6 A Subtreatments in self-targeting 16
Table 6 B Subtreatments in community-targeting experiment 17
Table 7 A Heterogeneity tests, principal accuracy results 18
Table 7 B Heterogeneity tests, perceptions and satisfaction results 20
Table 8 Effect of treatments on targeting accuracy 21
Table 9 Elite capture in community targeting 23
Table 10 Simulated impact on poverty rates 26
Table 11 Targeting costs and summary 27

Abbreviations and acronyms

BAPPENAS	State Ministry of National Development Planning
BDT	Basis Data Terpadu
BPS	Badan Pusat Statistik (Statistics Indonesia)
DepSos	Indonesian Ministry of Social Affairs
FE	fixed effect
HH	household
IDR	Indonesian rupiah
J-PAL	Adbul Latif Jameel Poverty Action Lab
LOGIT	logistic regression
NREGA	National Rural Employment Guarantee Act
OLS	ordinary least squares
PKH	Program Keluarga Harapan (Indonesia's conditional cash transfer programme)
PMT	proxy means test/testing
PPLS	official poverty census
RT	hamlet
TNP2K	Tim Nasional Percepatan Penanggulangan Kemiskinan (National Team for the Acceleration of Poverty Reduction)
WPA	Works Progress Administration

1. Background: targeting social programmes and principal interventions

Targeted social safety net programmes have become an increasingly common tool for addressing poverty (Coady, Grosh and Hoddinott 2004). In developing countries, however, targeting the poor is often a challenge, as most potential recipients work in the informal sectors and the government lacks verifiable records of their earnings. Governments have thus been developing three types of targeting strategies that do not rely on directly observing incomes:

- **Proxy means testing (PMT):** The government collects information on assets and demographic characteristics to create a proxy for household consumption or income. This proxy is then used for targeting.
- **Community targeting:** The government allows the community or some part of it (for example, local leaders) to select the beneficiaries through a pre-specified process. Examples include the Bangladesh Food for Education programme (Galasso and Ravallion 2005) and the Albanian economic support safety net programme (Alderman 2002).
- **Self-targeting:** In economic literature, this is called an ordeal mechanism. It imposes requirements on the programme that have differing costs for poor and rich people, dissuading the rich but not the poor from participating (Nichols, Smolensky and Tideman 1971; Nichols and Zeckhauser 1982; Besley and Coate 1992). Such mechanisms are common in many contexts: welfare programmes that require manual labour, for example the Works Progress Administration (WPA) in the United States or the National Rural Employment Guarantee Act (NREGA) in India; unemployment schemes that require participants to report weekly to the unemployment office during working hours; and, subsidised food schemes that provide low-quality or less- preferred grains.

Despite these developments in targeting, inaccurate targeting continues to be a tremendous impediment to the ability of social programmes' ability to achieve their goal of poverty reduction. While targeted transfer programmes have become increasingly prevalent in the developing world, they are plagued by high error rates; in fact, it is common to observe exclusion error rates of up to 50 per cent.[1] With high error rates, small improvements in targeting accuracy may yield a large increase in social programmes' power to improve the lives of the poor.

Improving targeting outcomes is an especially important tool in Indonesian social policy today, as the country moves to adopt a unified database, Basis Data Terpadu (BDT) to administer its social programmes. The targeting strategies tested in this experiment were designed to provide insight into the united database, and therefore into ways to construct more accurate beneficiary lists and mechanisms that will facilitate dynamic database updates.

---

[1] Targeting inaccuracy has been documented in many government anti-poverty programmes that offer subsidised rice, basic commodities, health insurance and scholarships for poor households. See, for example, Olken (2006); Daly and Fane (2002); Cameron (2002); Conn *et al.* (2008), and Alatas *et al.* (2012).

Given the high policy relevance of this project, the experiment was designed and carried out in close cooperation with the Indonesian government agencies that are interested in bringing these results to bear on pressing policy decisions. Researchers from J-PAL (Adbul Latif Jameel Poverty Action Lab) and the World Bank worked closely throughout this project with Indonesia's National Team for the Acceleration of Poverty Reduction (Tim Nasional Percepatan Penanggulangan Kemiskinan; TNP2K), the official statistics body Statistics Indonesia (Badan Pusat Statistik; BPS) and the State Ministry of National Development Planning (Badan Perencanaan Pembangunan Nasional; BAPPENAS).

We compare the three targeting mechanisms discussed above in the context of applying for aid programmes in Indonesia. Specifically, we examine Indonesia's conditional cash transfer programme, known as Program Keluarga Harapan (PKH), which is aimed at approximately the poorest 6–10% of the population. Eligibility for PKH has traditionally been determined using a PMT – a weighted sum of approximately 40 easy-to-observe assets, such as house size, roof material, motorcycle ownership, etc. Beneficiaries receive about US$150 per year for six years.

We compare the self-targeting and community-based methods against Indonesia's current targeting policy (PMT on a preselected lists of households). In self-targeting, if the application process is time-consuming and unlikely to result in a rich person getting benefits, the rich might choose not to apply, potentially saving the government the cost of screening out the rich. On the other hand, it is possible that the complicated application process may also dissuade the poor. For example, if the time costs are substantial, a large fraction of the poor may choose not to apply. In such a case, the programme could end up with a less pro-poor distribution of beneficiaries. This take-up problem has been documented in a wide variety of settings (see Currie 2006 for a review).

Community targeting gives local leaders and communities the power to select beneficiaries, and works under the presumption that it is harder to hide wealth from one's neighbours than from the government. The choice between PMT and community-targeting approaches is generally framed as a trade-off between the better information that communities might have versus the risk of elite capture in the community process. By focusing on assets, PMT captures the permanent component of consumption, but misses out on transitory or recent shocks. For example, a family may fall into poverty because one of its members is ill and cannot work, but they may live in a large house so PMT would classify the family as non-poor. Neighbours, on the other hand, may know the family's true situation from regularly observing the way they live. If the community perceives that the PMT is wrong, this may lead to a lack of legitimacy and political instability. On the other hand, while community targeting allows for the use of better local information, targeting decisions may be based on factors beyond poverty as defined by the government. This may be due to genuine disagreements about what poverty actually means: central government typically evaluates households based on consumption, whereas local communities use a utility function that may include other factors, such as earning potential, non-income dimensions of poverty or number of dependents. Likewise, government and local communities may place a different weight on the same variable when predicting consumption. Moreover, the community process could favour friends and relatives of the elite, and therefore lack legitimacy. Given the trade-offs involved, deciding which method works best is an empirical question.

2. Experimental design and data

2.1 Setting: the PKH programme

This project explores self-targeting mechanisms within the context of PKH, a conditional cash transfer project administered by the Ministry of Social Affairs (DepSos) in Indonesia. Target beneficiaries are households with *per capita* consumption below 80 per cent of the poverty line who meet the demographic requirements of having: at least one pregnant woman; a child aged 0 to 5; or a child under 18 who has not finished the nine years of compulsory education. Programme beneficiaries receive direct cash assistance averaging 1.4 million Indonesian rupiah (IDR) (approximately US$150[2]) per year, depending on their family make-up, school attendance, pre- or post-natal check-ups and completed vaccinations.[3] Around 1.12 million households are currently served by the programme.

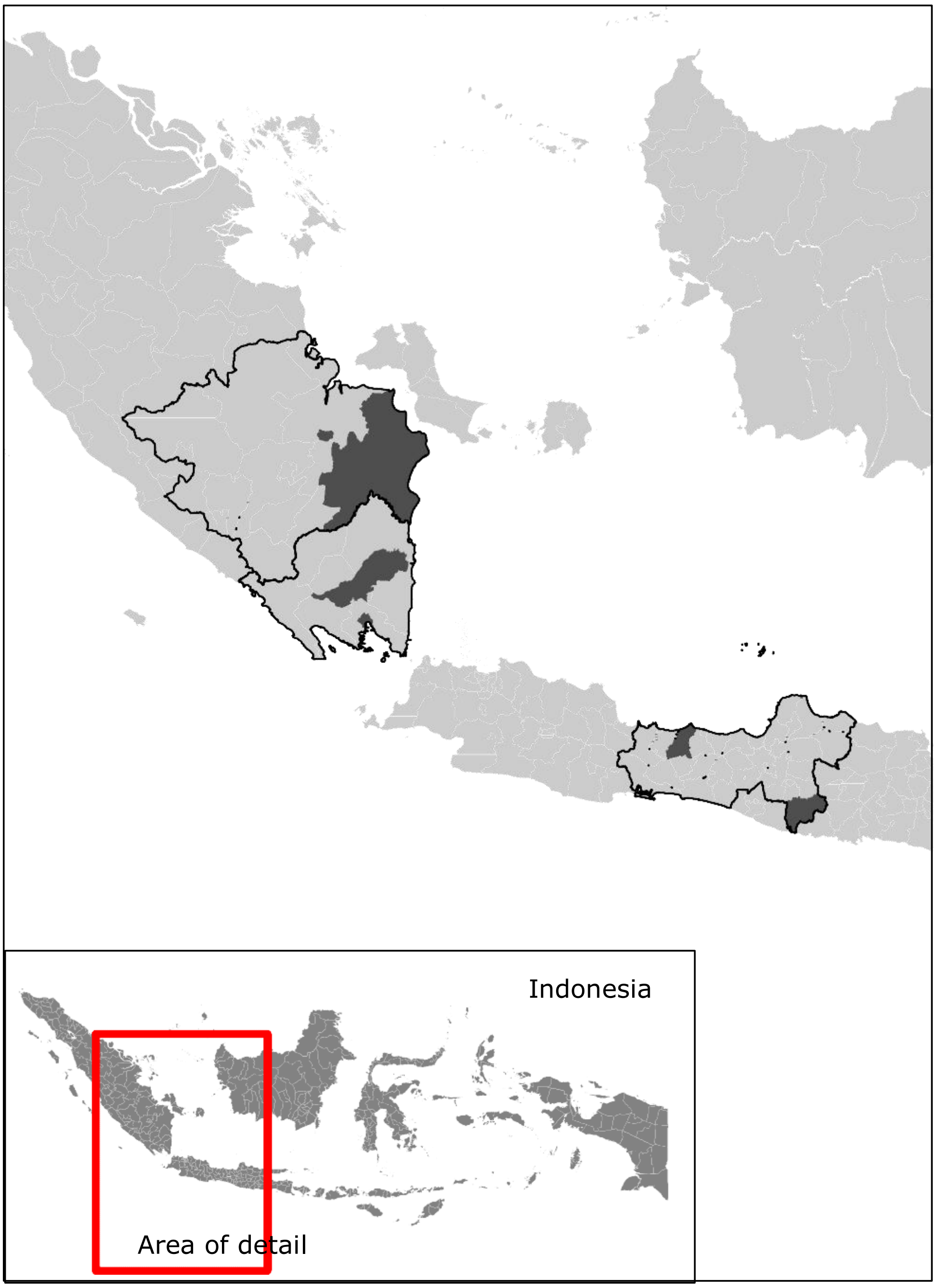
2.2 Sample selection

This project was carried out during the 2011 expansion of PKH to new areas. We chose six districts (two each in Lampung, South Sumatra and Central Java provinces) to include a variety of cultural and economic environments (see Figure 1). To understand how the different targeting methodologies worked within the context of a real programme, we chose our sites from locations where the government was rolling out the programme. Then, to ensure that the results are externally valid for the entire population of Indonesia, we stratified the sample along two key dimensions. First, we included districts both on and off Java, home to about 60 per cent of the population. Second, we ensured that 30 per cent of the sample units were located within urban areas (we would have preferred a 50:50 urban–rural split, but we were constrained by the locations where the programme was expanding).

Within each village, we randomly selected one sub-village for our surveys. These sub-villages are best thought of as neighbourhoods, consisting of less than 150 households. Each has an elected administrative head, whom we refer to as the sub-village head.

[2] This is based on an exchange rate of IDR9,535 = USD1 (2 October 2012).
[3] Note that, although eligibility for PKH transfers is officially dependent on recipients taking up healthcare and enrolling children in school, these conditions are not always enforced in practice.

Figure 1 Study area



2.3 Experimental design

Each of the 600 villages selected for the experiment was randomly allocated to one of the three methods for determining which households would be programme beneficiaries: self-targeting, community targeting or the status quo, where households are automatically enrolled in PKH based on their PMT score. This section describes each of these treatments in detail, and is summarised in Table 1.

Table 1 Experimental design

Main treatment	Subtreatment axis 1		Subtreatment axis 2		Total
Community targeting			Elite meeting	Full community meeting	
	One in, one out		50	50	100
	Addition		50	50	100
		Total	100	100	**200**
Self-targeting (ordeal mechanism)			Both spouses	Either spouse	
	Close sign-up		50	50	100
	Far sign-up		50	50	100
		Total	100	100	**200**
Automatic enrolment (PMT, status quo)				200	**200**
				TOTAL	**600**

Automatic enrolment treatment: the status quo

For each of the 200 villages in this treatment, targeting used a PMT approach that automatically enrolled all households which met the demographic requirements and passed the PMT.

BPS enumerators arrived at each village with a pre-printed list of households from the last targeting survey to interview (PPLS 2008). They asked the village leadership to add any households they thought had been inappropriately excluded, and they could add households to the list of potential interviewees if their own observations suggested that they were likely to be quite poor. Once on the interview list, households still had to undergo the PMT process.

After passing an initial pre-screening, each household was asked a series of 47 questions, ranging from attributes of their home (for example, wall type, roof type), ownership of specific assets (such as motorcycle, refrigerator), household composition and the household head's education and occupation. These measures were combined with location-based indicators, such as: population density; distance to the district capital; and access to education, healthcare facilities and a semi-permanent marketplace. Using pre-existing surveys (SUSENAS 2010 and PODES 2008), the government then estimated the relationship between these variables and the household's *per capita* consumption to generate a district-level formula for predicting consumption levels based on survey responses. Individuals with predicted consumption levels below each district's very poor line were eligible for the programme.

The automatic enrolment methodology is the one used by the Indonesian government, and we can use the results to compare the status quo with the policy alternatives discussed below. However, it is important to note that this initial screening may be more or less effective than a policy in which everyone is interviewed, depending on who the village leaders and enumerators add to the list. Therefore, we also conduct a simulation to understand how the potential targeting policies compare to a full census PMT (what we call the hypothetical PMT below).

Self-targeting treatment
In the self-targeting treatment, the enrolment criteria was the same, but rather than being automatic, interested households had to apply to join the programme at a central application. Self-selection meant that households which might have been automatically enrolled previously could miss out on benefits because they chose not to apply. Conversely, households which may have previously been passed over could apply to join the programme and ultimately receive benefits.

To publicise the application process, a community facilitator from Mitra Samya, a local non-governmental organisation (NGO), visited each village to inform the head and other leaders about the programme. They also held a community meeting to brainstorm the best indicators of local poverty and set a date for a series of neighbourhood-level meetings where the facilitator would inform households about the PKH programme and explain the registration process and application date. Facilitators would stress that the programme was geared towards the very poor. They would give examples of the type of questions that would be asked during the interview, explain the post-interview verification stage and highlight the criteria that locals would typically use to characterise very poor households. This was to ensure that households understood their chances of getting PKH and to make self-selection efficient.

Registration days were scheduled in advance, based on the relative size of the sub-villages. BPS enumerators were at the registration location from 8am to 5pm on the day. Householders wishing to apply for the programme would be signed in and given a number in the queue. When their number was called, they were interviewed by the enumerators, who collected the same data that was conducted in a PMT interview. In total, 48,794 households (about 19 per cent of the population) were interviewed across 200 villages.

Applicant households were divided into very poor or not very poor based on the PMT regression formula and the district-specific very poor line. The PMT formula and questions used were the same as those in the automatic enrolment treatment. Anyone classified as very poor, based on the assets they disclosed at interview, and who was also listed in the 2008 poverty census as very poor (about 3.4 per cent of interviewees), was automatically selected as a PKH recipient. Other households were subjected to a verification process: surveyors visited their homes to collect data on the same set of asset questions. The results were used, with the PMT regression formula and poverty lines, to determine the final list of beneficiaries.

Note that about half the households that were verified were subsequently taken off the list for failing to pass the asset test during verification. Only three households were incorrectly screened out during this process suggesting that verification, on net, helped to reduce inclusion error.

Within self-targeting villages, there were two subtreatments, to vary the costs of registration.

<u>Distance subtreatment:</u> We experimentally varied the distance to the sign-up location, to vary the time cost in applying to join the programme. We ensured that all locations could still be reached on foot, so as not to impose transport costs on very poor households. In

urban areas, villages were randomly allocated a registration site at the sub-district office (far location) or the village office (near location). In rural areas, where distances are greater, villages were randomly allocated a registration site at the village office (far location) or in the sub-village (near location).[4]

Both spouse subtreatment: To vary the opportunity cost of signing up, we experimentally varied the requirement for one or two household members to attend registration. In half the self-targeting villages, any adult in the household (household head or spouse) could go to register. Given that the programme was geared towards women, we expected that mostly women would sign up. In the other half of villages, we required both wife and husband to attend.[5]

Community-targeting treatment

In the community-targeting villages, beneficiaries were not determined through an asset-based test, but through a community meeting with no additional verification. Those attending the community meeting in each sub-village determined the list of beneficiaries through a poverty-ranking exercise. After explaining the PKH programme and its purpose, the facilitator displayed index cards listing the poorest households in the sub-village according to the official poverty census (PPLS 2008). This is the same data source used in the status quo, asset-based treatment. The number of cards shown was roughly equivalent to 75 per cent of the sub-village's quota.

Working with the community members at the meeting, the facilitator then removed households with inaccurate information – in other words, those who had moved away or did not match at least one of the three PKH criteria. The facilitator then asked participants to brainstorm a list of additional households they thought to be the most deserving of PKH in their sub-village, up to 100 per cent of the sub-village's quota. The facilitator then led participants through a process of ranking households on both lists – the initial PPLS set and the additional households brainstormed at the meeting. The recipient list was finalised using the ranking determined at the meeting, with no further government verification.

To vary levels of control at the meetings, we randomly assigned villages to two subtreatments:

Addition subtreatment: To vary the level of community control, we randomly assigned some villages to an addition treatment, in which the PPLS households had to receive the benefit in addition to any brainstormed households. In the other villages, we used a one in, one out treatment, in which PPLS households could be substituted out. Meeting participants thus had complete control over the list.

Elite subtreatment: To vary the level of elite control in meetings, we randomly varied who was invited: in half of the villages (randomly selected), we asked the local sub-village

[4] The distance subtreatment was violated in four villages: in one, a large subset of the village refused to participate in interviews in a certain sub-village due to longstanding ethnic tensions, so we held interviews in another sub-village for one day; in the second, one sub-village was four to five hours' walk from the village office, so interviewers set aside a day to go to that sub-village; in the third and fourth villages, local leaders insisted that the interview site be moved closer to the village. All analysis reports intent-to-treat effects where these four villages are categorised based on the randomisation result, not actual implementation.

[5] If the spouse was for some reason unable to attend, we required that they bring a letter signed by the head of the neighbourhood providing reasons for the spouse's unavailability.

head to invite between five and eight local leaders, both formal and informal. In the other half, we invited the whole community, in order to provide a potential check on the power of the elites to capture the targeting process. In the full community villages, the facilitator and sub-village head heavily advertised the meeting to encourage full attendance; in many cases, the facilitators made door-to-door visits. On average, 15 per cent of households attended meetings in the elite subtreatment, while 59 per cent did so in the community subtreatment.

2.4 Randomisation design and timing

We randomly assigned each of the 600 villages to one of the main treatments (see Table 2) by computer. In order to ensure experimental balance across geographic regions, we divided it into 58 geographic strata. Each stratum consists of all the villages from one or more sub-district (*kecamatan*), and is entirely located in a single district (*kabupaten*). We then randomly and independently allocated each self-targeting and community-targeting village to the subtreatments, with each of these two subtreatment randomisations stratified by the previously defined strata and the main treatment.

2.5 Power calculations

We based our power calculations for this experiment on our previous targeting experiment in Indonesia. In that experiment, we had 200 villages per treatment group. To estimate mistargeting, we were able to distinguish PMT groups from community and hybrid ones, but were unable to distinguish between the community and hybrid groups (Alatas *et al*. 2012).

To estimate the treatment effect, we used the mean mistargeting rate for the unconstrained PMT, which is 0.27, and assumed the constrained mistargeting community rate, 0.33. We assumed nine households per village, and an intra-cluster correlation of 0.1. Setting alpha=0.05 and beta=0.85 (standard assumptions in the literature), we get a sample size of 432 for two treatments, or 216 villages in a treatment group. Stratifying by sub-district and controlling for strata, we were able to improve power such that our 200-village groups provided reasonable power.

Given our constraints, we did not conduct power calculations for our subtreatments. However, we know that these subtreatments did change participants' behaviour, as we discuss below.

2.6 Quantitative and qualitative data collection

We collected several datasets for this study. From December 2010 to March 2011, an independent survey firm, SurveyMeter, collected baseline data from nine randomly selected households and the sub-village head in one randomly selected sub-village in each village. The government conducted targeting treatments and created the beneficiary lists between January and April 2011 once the surveying was complete in each district.

SurveyMeter conducted a first follow-up survey in early August 2011, after the targeting was complete but before the beneficiary lists were announced. Fund distribution started in late August 2011.[6] A second endline survey was conducted between January and March 2012, after the first and second sets of funds had been distributed. The survey included data on consumption and income, as well as the full set of asset and demographic measures that comprise the PMT's predicted consumption score. We also collected: data on elite-relatedness to monitor the effects of elite capture in the community-targeting treatment; historical data on access to a variety of other targeted programmes; and qualitative data on respondents' perceptions and feelings about the targeting strategies within these surveys.

In addition, we collected extensive qualitative data on programme functioning and stakeholder beliefs. During the experiment, J-PAL staff visited the field 15 times to monitor the functioning of the various treatments in the various districts, and to interview beneficiaries and other stakeholders in the experiment. During field visits, J-PAL staff typically visited between five and ten villages, attending community meetings in the community-targeting treatment and sign-up centres in the self-targeting treatment.

After each field trip, the project team wrote up a summary of both their observations and the interviews that they conducted. There was considerable discussion during these trips about whether community meetings should have more detailed poverty discussions than they currently had, while stakeholders' discussions focused on logistics – for example, whether there should be more enumerators present or more days.

In short, the visits threw light on what was informing decisions in community meetings and what mechanisms were at play in determining who showed up to the self-targeting sign-up sites. The outcomes from these visits were used to refine our endline survey instrument, to ensure it would capture such subtleties. J-PAL staff also conducted monitoring trips to oversee trainings for facilitators and enumerators.

3. Results

3.1 Balance check

The variables for the balance check were chosen prior to obtaining the data. Table 2 shows the balance checks from the baseline survey and reveals that our randomisation was successful. Only two of the differences that we consider are statistically significant at standard significance levels, which is consistent with what we would expect by chance.

[6] Note that, following the selection process, the Department of Social Affairs realised it had additional funds available and increased the number of programme beneficiaries to include households that did not pass the selection process in our experimental treatments but had been classified as very poor under the 2008 poverty census. We do not include these additional households when calculating beneficiaries for experimental evaluation purposes, but it is important to keep these extra households in mind when evaluating the programme at the endline.

Table 2 Balance check

	Mean in PMT	Mean in self-targeting	No FE self-targeting	With FE self-targeting	Mean in community targeting	No FE community targeting	With FE community targeting
Log *per capita* consumption	13.112	13.105	−0.007	−0.001	13.123	0.017	0.023
	(0.228)	(0.251)	(0.024)	(0.021)	(0.252)	(0.024)	(0.020)
Years of education: HH head	7.297	7.145	−0.152	−0.118	7.181	−0.116	−0.107
	(2.208)	(2.043)	(0.213)	(0.167)	(1.919)	(0.207)	(0.168)
PMT score	12.795	12.792	−0.003	0.003	12.767	−0.029	−0.027
	(0.228)	(0.251)	(0.024)	(0.019)	(0.246)	(0.024)	(0.019)
HHs in agriculture (%)	0.073	0.071	−0.002	−0.004	0.074	0.001	0.000
	(0.068)	(0.063)	(0.007)	(0.005)	(0.069)	(0.007)	(0.006)
Years of education: RT head	8.131	8.060	−0.070	−0.044	8.280	0.149	0.160
	(3.773)	(3.333)	(0.357)	(0.314)	(3.571)	(0.368)	(0.347)
Log # of HHs in RT	4.227	4.241	0.014	0.032	4.266	0.039	0.054
	(0.520)	(0.468)	(0.049)	(0.045)	(0.467)	(0.049)	(0.047)
Distance to *kecamatan*	7.434	6.404	−1.031	−1.038	7.627	0.192	0.005
	(21.919)	(8.184)	(1.654)	(1.615)	(14.509)	(1.859)	(1.806)
Log village size	4.038	3.925	−0.113	−0.129*	4.049	0.012	0.025
	(1.574)	(1.476)	(0.153)	(0.067)	(1.611)	(0.159)	(0.076)
Religious buildings per HH	0.005	0.005	0.000	−0.000	0.005	0.000	0.000*
	(0.003)	(0.003)	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)
Primary schools per HH	0.003	0.003	−0.000	−0.000	0.003	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)
Observations	200	200	400	400	200	400	400
Joint significance test:							
Coefficient			−0.0374	−0.0327		0.0200	0.0259
Standard error			0.0413	0.0301		0.0359	0.0265
p-value			0.364	0.277		0.578	0.328

Robust standard errors in parentheses. Regressions include stratum fixed effects.
***p < 0.01, **p < 0.05, *p < 0.1

3.2 Accuracy and perceptions

Table 3 compares the targeting accuracy of both methods with the status quo on a variety of outcomes. We code a household as incorrectly targeted if its *per capita* consumption level is below the poverty line (defined as 80 per cent of the 16th percentile of consumption in the district) and it was not chosen, or its *per capita* consumption is above the poverty line and it was chosen. Impacts on the error rates are estimated using the following equation:

$$Outcome = \alpha + \beta_1 SelfTargeting + \beta_2 CommunityTargeting + \gamma_k + \varepsilon$$

where *SelfTargeting* and *CommunityTargeting* are dummies for the treatment status of each village, *k* represents a stratum, and γ_k are stratum fixed effects.

Standard errors are clustered at the village level. In all of these regressions, PMT is the omitted category, so it can be interpreted as the impact of self-targeting and community targeting relative to the PMT status quo.

Column 1 provides a test of the treatments' impact on the beneficiaries' overall consumption levels. We show that self-targeting produced a significantly poorer group of beneficiaries than the status quo, while community targeting had an insignificant effect.

Column 2 indicates overall error in assigning beneficiaries. We find that self-targeting outperforms the other treatments in targeting error, while PMT outperforms the community-targeting treatment.

Columns 3 and 4 report exclusion and inclusion error, respectively. Due to the smaller size, however, the exclusion error coefficients are insignificant.

Column 5 breaks down the results from columns 1–4, disaggregating by consumption quintile (quintile 5, containing the richest households, is omitted). The results show that community targeting generally results in beneficiaries from the lower to middle quintiles, while self-targeting primarily results in beneficiaries from the lowest quintile.

Column 6 presents results similar to column 5, using measured consumption to predict benefit receipt. Unsurprisingly, self-targeting significantly outperforms PMT in the correlation of consumption and benefit receipt, while community targeting insignificantly outperforms PMT.

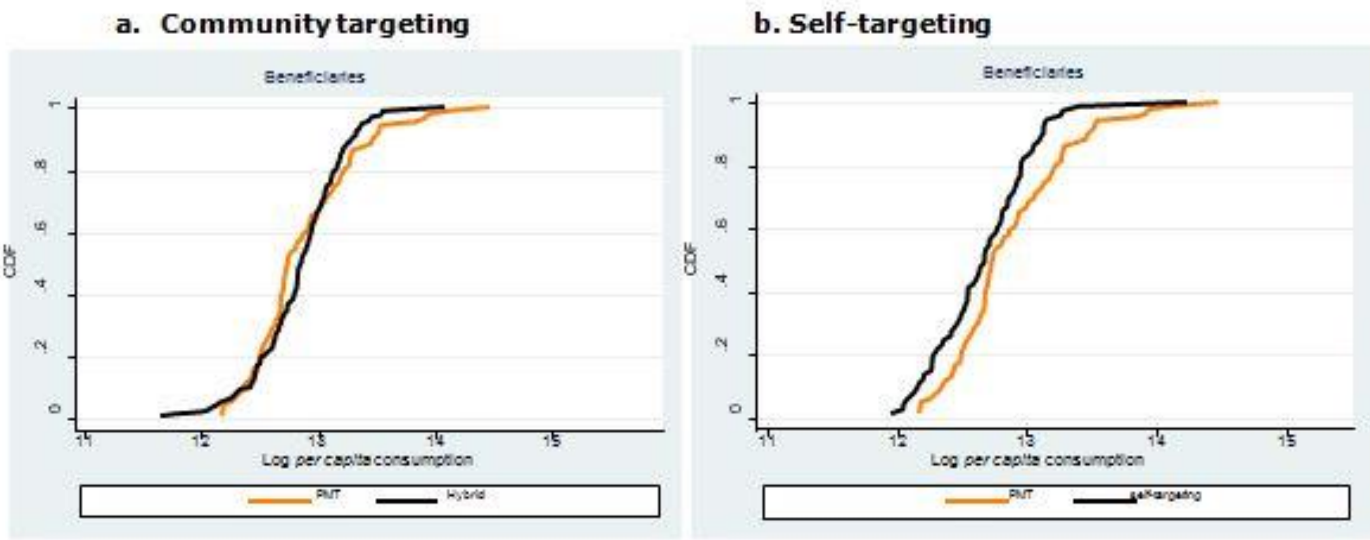
Table 3 Effect of treatments on targeting accuracy

	Log consumption of beneficiaries (OLS) (1)	Mistargeted (LOGIT) (2)	Exclusion error (LOGIT) (3)	Inclusion error (LOGIT) (4)	Get benefit (LOGIT) (5)	Get benefit (LOGIT) (6)
Self-targeting	−0.123*	−0.223**	−0.515	−0.328*	−2.061*	13.962**
	(0.071)	(0.114)	(0.445)	(0.183)	(1.074)	(4.837)
Community targeting	0.055	0.316***	−0.080	0.694***	−0.439	3.483
	(0.067)	(0.103)	(0.457)	(0.150)	(0.593)	(4.004)
Self * consumption level 1					2.329**	
					(1.106)	
Self * consumption level 2					1.670	
					(1.121)	
Self * consumption level 3					2.186*	
					(1.154)	
Self * consumption level 4					0.699	
					(1.256)	
Community * consumption level 1					0.905	
					(0.644)	
Community * consumption level 2					1.262*	
					(0.649)	
Community * consumption level 3					1.794***	
					(0.691)	
Community * consumption level 4					0.866	
					(0.706)	
Consumption level 1					1.558***	
					(0.432)	
Consumption level 2					1.279***	
					(0.436)	
Consumption level 3					0.401	
					(0.493)	
Consumption level 4					0.530	
					(0.485)	
Self * log consumption						−1.106***
						(0.378)
Community * log consumption						−0.218
						(0.310)
Log consumption						−1.086***
Constant	12.819***					(0.250)
Observations	313	5,958	219	5,423	5,796	5,796
Dependent variable mean	12.82	0.101	0.806	0.0494	0.0540	0.0540

Robust standard errors in parentheses. Regressions include stratum fixed effects. Standard errors clustered at village level.
***p < 0.01, **p < 0.05, *p < 0.1

Table 3 is presented graphically in Figure 2, which confirms that on objective consumption measures self-targeting outperforms PMT. While PMT outperforms community targeting, the latter shows some gains over PMT in reaching the very poor, suggesting that it identifies an especially vulnerable population passed over by the status quo.

Figure 2 Consumption of beneficiaries under different treatments



Turning to the satisfaction results, the community-targeting treatment shows a significant advantage over the other targeting schemes. Table 4 shows impacts on 10 different measures of satisfaction by treatment. With one important caveat, it tells quite the opposite story from Table 3, with community targeting significantly outperforming PMT on nearly every category and PMT outperforming self-targeting on many of the variables.

Community targeting shows a significant improvement over PMT on eight out of 10 measures, including perceived accuracy, fairness, overall satisfaction, beneficiary poverty levels and desire to use the system again. Self-targeting underperformed PMT on almost all categories, and was significantly inferior in perceived accuracy, satisfaction, omission of deserving households and beneficiary poverty levels. It is possible, however, that the negative results for self-targeting are driven by respondents' ignorance of the PMT method compared with the self-targeting treatment.

Panel B (responsiveness) shows the same results with the dependent variable as a binary equal to 1 if the respondent had an opinion at all. The results show that respondents were significantly more likely to express an opinion about self-selection and community targeting than about PMT. If self-selection were the status quo, and less the subject of substantial socialisation in treatment villages, it is plausible that it would stir up less harsh opinions and satisfaction results would look considerably more encouraging. The same attenuation in satisfaction might also be seen in community targeting.

Panel C (effect of receiving the benefit) considers the impact of receiving PKH on people's perceptions of the programme. Unsurprisingly, results show that PKH recipients were significantly more positive about the targeting procedures across the board. Interestingly, the gulf between recipients and non-recipients narrows considerably in community-targeting villages, probably because non-recipients were more satisfied in those villages.

Table 4 Perceptions of different targeting mechanisms (scaled 0–1)

	How smooth was the process?	How efficient were PKH staff?	Is the method accurate?	Are you satisfied with the process in general?	Is the process fair?	Of HHs you know are on the list, how many do you agree should be there?	How poor are HHs on the list?	Are there HHs receiving PKH who are not supposed to? (Y=0, N=1)	Are there HHs who deserve to be on the list but are not? (Y=0, N=1)	Would you like to use the process again?
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Overall perceptions										
Self-targeting	0.00	−0.02	−0.04**	−0.05***	−0.02	−0.03*	−0.00	−0.07***	−0.02	−0.04
	(0.012)	(0.011)	(0.017)	(0.017)	(0.014)	(0.014)	(0.016)	(0.026)	(0.028)	(0.030)
Community targeting	0.02*	−0.00	0.05***	0.04**	0.06***	0.03***	0.04***	0.01	0.06**	0.10***
	(0.013)	(0.011)	(0.017)	(0.017)	(0.014)	(0.012)	(0.014)	(0.023)	(0.026)	(0.029)
Constant	0.67***	0.70***	0.61***	0.55***	0.49***	0.86***	0.48***	0.80***	0.39***	0.59***
	(0.010)	(0.009)	(0.014)	(0.014)	(0.010)	(0.010)	(0.010)	(0.017)	(0.019)	(0.022)
Observations	2,481	2,249	2,654	2,690	3,443	3,723	3,729	3,500	3,396	2,796
Panel B: Responsiveness: 1=had an opinion, 0=no opinion (data missing in Panel A and C)										
Self-targeting	0.31***	0.29***	0.29***	0.29***	0.15***	0.06*	0.06*	0.05*	0.06*	0.26***
	(0.023)	(0.023)	(0.024)	(0.024)	(0.026)	(0.032)	(0.032)	(0.031)	(0.031)	(0.024)
Community targeting	0.22***	0.19***	0.21***	0.22***	0.16***	0.18***	0.18***	0.19***	0.17***	0.19***
	(0.025)	(0.023)	(0.026)	(0.026)	(0.026)	(0.028)	(0.028)	(0.027)	(0.027)	(0.025)
Constant	0.26***	0.24***	0.30***	0.30***	0.50***	0.57***	0.58***	0.53***	0.52***	0.34***
	(0.016)	(0.015)	(0.017)	(0.018)	(0.019)	(0.022)	(0.022)	(0.021)	(0.021)	(0.017)
Observations	5,682	5,682	5,682	5,682	5,682	5,682	5,682	5,682	5,682	5,682
Panel C: Effect of receiving benefit										
Self-targeting	0.00	−0.02	−0.02	−0.04**	−0.01	−0.03**	−0.00	−0.08***	−0.02	0.01
	(0.015)	(0.013)	(0.019)	(0.019)	(0.015)	(0.017)	(0.018)	(0.029)	(0.030)	(0.033)
Community targeting	0.03*	0.00	0.07***	0.06***	0.07***	0.04***	0.05***	0.01	0.07**	0.14***
	(0.016)	(0.013)	(0.020)	(0.019)	(0.015)	(0.014)	(0.015)	(0.026)	(0.028)	(0.035)
Got PKH	0.06***	0.05***	0.20***	0.22***	0.26***	0.10***	0.08***	0.14***	0.06	0.41***
	(0.018)	(0.017)	(0.023)	(0.027)	(0.021)	(0.015)	(0.020)	(0.029)	(0.045)	(0.037)
Self * got PKH	0.03	0.05**	0.02	0.07**	−0.02	0.02	0.00	0.05	−0.01	−0.04
	(0.022)	(0.021)	(0.031)	(0.033)	(0.029)	(0.023)	(0.028)	(0.043)	(0.059)	(0.047)
Community * got PKH	−0.01	−0.00	−0.06**	−0.04	−0.06**	−0.04**	−0.05*	−0.05	−0.07	−0.12**
	(0.021)	(0.021)	(0.031)	(0.034)	(0.026)	(0.019)	(0.026)	(0.041)	(0.057)	(0.049)
Constant	0.65***	0.68***	0.55***	0.48***	0.44***	0.84***	0.46***	0.77***	0.38***	0.48***
	(0.013)	(0.011)	(0.015)	(0.015)	(0.011)	(0.011)	(0.011)	(0.019)	(0.020)	(0.026)
Observations	2,481	2,249	2,654	2,690	3,443	3,723	3,729	3,500	3,396	2,796

Robust standard errors in parentheses. All regressions OLS, and all responses scaled 0 to 1. Regressions include stratum fixed effects. Standard errors clustered at village level. *** p < 0.01, ** p < 0.05, * p < 0.1

3.3 Alternative measures of wealth

Villagers have more opportunity to alter targeting outcomes when using self-targeting and community targeting than they do through PMT. When self-targeting, individuals decide whether or not to sign up, and in community targeting they help to make decisions. Thus, even if error rates according to objective measures are not drastically different from the status quo, it is possible that alternative targeting allows communities to target other aspects of poverty that map closer to their own perceptions.

We provide some insights into this situation by exploring the communities' subjective beliefs on beneficiaries' poverty rankings across the treatments. For comparability, all measures are created as rank measures spanning 0–1 by ranking all households by village and dividing by the number of households in the village.

Table 5 presents striking evidence that community targeting does indeed allow villagers to target alternative measures of wealth that are not captured in objective consumption. Despite the small samples, community targeting still appears to target more effectively on subjective measures than the PMT. These findings are consistent with our earlier study on community targeting (Alatas *et al.* 2012).

Table 5 Targeting accuracy using subjective measures of wealth and welfare

	Wealth according to other villagers (1)	**Wealth according to RT head (2)**	**Wealth according to HH (3)**
Self-targeting	0.0101	0.00529	0.0264
	(0.0410)	(0.0399)	(0.0468)
Community targeting	−0.0434	−0.0756**	−0.00345
	(0.0358)	(0.0332)	(0.0397)
Observations	313	295	313
R-squared	0.240	0.253	0.194

Robust standard errors in parentheses.
All regressions OLS, and all responses scaled 0 to 1. Regressions include stratum fixed effects.
Standard errors clustered at village level.
***p < 0.01, **p < 0.05, *p < 0.1

3.4 Subtreatments

Table 6A reports the results of the subtreatments in the self-targeting experiment. The treatments had real effects on behaviour: decreasing the distance to be travelled to sign up increased applications; requiring both spouses to attend actually increased applications. However, these treatments affected everyone and they were uncorrelated with household consumption levels.

Table 6 A Subtreatments in self-targeting

	Show up					
	(1)	(2)	(3)	(4)	(5)	(6)
Close subtreatment	0.28***	0.48	0.19			
	(0.102)	(3.057)	(0.219)			
Both spouse subtreatment				0.18*	3.33	0.38*
				(0.102)	(3.050)	(0.218)
Close* log consumption		−0.02				
		(0.235)				
Both spouse* log consumption					−0.24	
					(0.234)	
Log consumption		−1.45***			−1.34***	
		(0.165)			(0.167)	
Close* consumption level 2			−0.29			
			(0.308)			
Close* consumption level 3			0.32			
			(0.314)			
Close* consumption level 4			−0.26			
			(0.328)			
Close* consumption level 5			0.28			
			(0.374)			
Both spouse* consumption level 2						−0.32
						(0.306)
Both spouse* consumption level 3						−0.30
						(0.312)
Both spouse* consumption level 4						−0.12
						(0.323)
Both spouse* consumption level 5						−0.36
						(0.374)
Consumption level 2			−0.33			−0.33
			(0.224)			(0.219)
Consumption level 3			−0.79***			−0.47**
			(0.232)			(0.219)
Consumption level 4			−1.07***			−1.15***
			(0.229)			(0.231)
Consumption level 5			−2.27***			−1.96***
			(0.276)			(0.265)
Observations	1,960	1,960	1,960	1,960	1,960	1,960
Mean of dependent variable	0.385	0.385	0.385	0.385	0.385	0.385

Note: Standard errors in parentheses. Dependent variable is show-up rate. All regressions logit with stratum fixed effects. Standard errors clustered at village level.
***p < 0.01, **p < 0.05, *p < 0.1

Table 6B reports the results of the subtreatments in the community-targeting experiment. For the most part, the subtreatments do not appear to significantly affect the overall targeting accuracy. While the one in, one out treatment (giving more power to the community meetings) reduces the error rates and reaches poorer beneficiaries, it is not statistically significant.

Table 6 B Subtreatments in community-targeting experiment

	Log consumption of beneficiaries (OLS)	Mistargeting (LOGIT)	Exclusion error (LOGIT)	Inclusion error (LOGIT)
	(1)	(2)	(3)	(4)
Elite subtreatment	−0.005	0.050	−0.424	−0.148
	(0.073)	(0.172)	(0.721)	(0.259)
One in, one out subtreatment	−0.033	−0.106	−0.325	−0.061
	(0.073)	(0.170)	(0.687)	(0.264)
Observations	154	2,000	130	1,870
Mean of dependent variable	12.85	0.127	0.885	0.0743

Robust standard errors in parentheses. Standard errors clustered at village level.
***p < 0.01, **p < 0.05, *p < 0.1

3.5 Treatment heterogeneity

Given that the levels of information and capture may be different across localities, we examine the heterogeneity in the relative effectiveness of the different treatments across three dimensions:

a) That community methods may do worse in urban areas, where individuals might not know their neighbours as well. Our sample was stratified along this dimension to ensure that we had a large enough sample size to test this hypothesis.
b) The level of inequality in the villages could result in important differences between the two techniques. On the one hand, community targeting may work better in areas with high inequality, since it implies that the rich and poor are more sharply differentiated. On the other hand, elite capture of community-based techniques may be more severe in areas with high inequality, if rich elites are powerful enough to exclude the poor from the community decision-making process.
c) We test for different results on and off Java, which, as mentioned above, is the principal axis of heterogeneity considered by the Indonesian government in their consideration of policy relevance for the whole country.

Table 7A shows the results of heterogeneity tests along both dimensions for targeting accuracy, while Table 7B shows the results for perceptions/satisfaction. Overall, there is not much evidence that the treatments have heterogeneous results. Table 7A shows that beneficiaries' log consumption is significantly lower among high-poverty villages in the community-targeting treatment, although the effect is very small for slight changes in poverty density. Exclusion error for self-targeting is also significantly higher in urban areas, suggesting perhaps that publicising the programme to the poor is a challenge in an urban environment.

Table 7 A Heterogeneity tests, principal accuracy results

	Beneficiaries' log consumption (OLS) (1)	Mistargeted (LOGIT) (2)	Exclusion error (LOGIT) (3)	Inclusion error (LOGIT) (4)
Self-targeting	0.067	−0.638**	−2.190	−0.928*
	(0.186	(0.306)	(1.488)	(0.495)
Community targeting	0.340*	0.018	0.635	0.296
	(0.182)	(0.287)	(1.434)	(0.434)
Urban	0.055	0.137	−2.405**	0.507
	(0.137	(0.215)	(1.071)	(0.317)
Poverty density	0.305	−0.143	−0.107	0.468
	(0.386	(0.584)	(3.623)	(0.863)
Self * urban	0.052	0.347	2.788**	0.324
	(0.158	(0.251)	(1.098)	(0.392)
Community * urban	−0.176	0.205	0.664	0.154
	(0.145	(0.226)	(1.043)	(0.320)
Self * poverty density	−0.578	0.730	2.860	1.087
	(0.424	(0.721)	(4.397)	(1.094)
Community * poverty density	−0.680*	0.871	−2.518	0.998
	(0.406	(0.686)	(4.101)	(0.982)
Self * Java	−0.079	0.271	0.126	0.535
	(0.157	(0.256)	(1.194)	(0.415)
Community * Java	−0.033	−0.050	−1.316	0.238
	(0.153)	(0.243)	(1.252)	(0.362)
Observations	313	5,958	215	5,430
Dependent variable mean	12.82	0.101	0.805	0.0494

Robust standard errors in parentheses. Regressions include stratum fixed effects.
Standard errors clustered at village level.
***p < 0.01, **p < 0.05, *p < 0.1

Table 7B reveals some heterogeneity in satisfaction levels, but nothing extremely notable. Households in urban areas with community targeting are significantly less likely to find the process fair or to think the beneficiary households are poor, perhaps because the community-targeting procedure relies on the close-knit culture of a rural village. Notably, community targeting still outperforms the alternatives, even with this caveat. Respondents were significantly more likely to think the targeting process left out deserving households in urban areas for both treatments (which maps to the result about urban exclusion error in self-targeting) and high-poverty areas for community targeting, although this result might be expected given the nature of the question.

However, on net, we do not see striking heterogeneity of the treatment across areas.

3.6 Hypothetical universal PMT

The automatic enrolment system used in the study, conducting PMT on preselected beneficiary groups, was the actual system typically used by the Indonesian government. One alternative is to conduct the PMT on a census of households. This may improve targeting efficiency if those preselected out are the poor, but could also make it worse because of the error inherent in the targeting formulas. To explore the impact of the PMT preselection on the treatment comparisons, we replicate the analysis having filled in the PMT score for those who were not interviewed in the automatic enrolment treatment using our baseline data. While this is not a feasible policy, it does provide a useful benchmark against which to measure the self-targeting treatment and understand the capabilities of the proxy means process.

Table 8 replicates Table 3, but uses the hypothetical PMT as a baseline. The results are fairly straightforward: the self-targeting treatment still performs roughly as well as the PMT, due largely to the fact that the PMT includes many wealthy households screened out by the self-targeting treatment. That the self-targeting is able to roughly match error rates with the significantly more costly treatment of interviewing every single household speaks strongly for the screening power of the ordeal mechanism.

Table 7 B Heterogeneity tests, perceptions and satisfaction results

	How [smooth] was the process ?	How efficient were PKH staff?	Is the method accurate?	Are you satisfied with the process in general?	Is the process fair?	Of HHs you know are on the list, how many do you agree should be there?	How poor are HHs on the list?	Are there HHs receiving PKH who are not supposed to? (Y=0, N=1)	Are there HHs who deserve to be on the list but are not? (Y=0, N=1)	Would you like to use the process again?
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Self-targeting	−0.04	−0.04	−0.06	−0.02	−0.04	−0.02	0.07*	−0.05	0.13*	0.06
	(0.028)	(0.029)	(0.039)	(0.040)	(0.035)	(0.035)	(0.041)	(0.063)	(0.067)	(0.072)
Community targeting	0.01	−0.00	0.04	0.07*	0.09***	0.03	0.11***	0.07	0.23***	0.22***
	(0.028)	(0.025)	(0.039)	(0.042)	(0.033)	(0.029)	(0.036)	(0.049)	(0.063)	(0.072)
Urban	−0.01	−0.01	−0.01	0.02	0.02	0.02	−0.00	0.06*	0.16***	0.00
	(0.022)	(0.019)	(0.028)	(0.029)	(0.024)	(0.022)	(0.022)	(0.036)	(0.044)	(0.046)
Poverty density	−0.06	−0.01	−0.15*	0.03	−0.10*	−0.17***	−0.06	−0.31***	0.17	0.04
	(0.052)	(0.046)	(0.078)	(0.075)	(0.057)	(0.059)	(0.068)	(0.101)	(0.108)	(0.137)
Self * urban	0.01	0.01	0.02	−0.02	−0.02	−0.03	0.00	−0.06	−0.20***	−0.05
	(0.026)	(0.025)	(0.035)	(0.035)	(0.030)	(0.029)	(0.034)	(0.049)	(0.058)	(0.059)
Community * urban	−0.01	0.00	−0.04	−0.05	−0.06**	−0.02	−0.05*	−0.04	−0.12**	−0.05
	(0.025)	(0.022)	(0.033)	(0.034)	(0.027)	(0.025)	(0.027)	(0.042)	(0.054)	(0.056)
Self * poverty density	0.11*	0.06	0.05	−0.09	0.07	0.03	−0.12	0.05	−0.22	−0.15
	(0.061)	(0.059)	(0.099)	(0.092)	(0.078)	(0.086)	(0.108)	(0.162)	(0.165)	(0.178)
Community * poverty density	0.02	−0.02	−0.01	−0.07	−0.07	0.03	−0.16*	−0.09	−0.42***	−0.28
	(0.064)	(0.058)	(0.107)	(0.112)	(0.080)	(0.071)	(0.093)	(0.137)	(0.161)	(0.182)
Self * Java	0.02	0.01	0.00	0.00	0.04	−0.02	−0.08***	−0.05	−0.04	−0.09*
	(0.019)	(0.019)	(0.026)	(0.025)	(0.023)	(0.025)	(0.030)	(0.044)	(0.048)	(0.046)
Community * Java	0.02	0.01	0.04	0.02	0.03	0.00	0.00	−0.08**	−0.02	−0.06
	(0.020)	(0.018)	(0.028)	(0.026)	(0.025)	(0.019)	(0.025)	(0.038)	(0.048)	(0.048)
Constant	0.70***	0.71***	0.66***	0.54***	0.51***	0.90***	0.50***	0.87***	0.28***	0.57***
	(0.021)	(0.020)	(0.028)	(0.030)	(0.024)	(0.022)	(0.023)	(0.037)	(0.038)	(0.051)
Observations	2,481	2,249	2,654	2,690	3,443	3,723	3,729	3,500	3,396	2,796

Robust standard errors in parentheses. All regressions OLS and include stratum fixed effects. Standard errors clustered at village level.
***p < 0.01, **p < 0.05, *p < 0.

Table 8 Effect of treatments on targeting accuracy

	Log consumption of beneficiaries (OLS) (1)	Mistargeted (LOGIT) (2)	Exclusion error (LOGIT) (3)	Inclusion error (LOGIT) (4)	Get benefit (LOGIT) (5)	Get benefit (LOGIT) (6)
Self-targeting	−0.061	−0.276**	0.201	−0.562***	−2.040*	8.099*
	(0.061)	(0.113)	(0.395)	(0.176)	(1.074)	(4.697)
Community targeting	0.112*	0.267***	0.590	0.467***	−0.429	−2.505
	(0.057)	(0.102)	(0.411)	(0.141)	(0.593)	(3.808)
Self * consumption level 1					1.739	
					(1.100)	
Self * consumption level 2					1.477	
					(1.119)	
Self * consumption level 3					1.906*	
					(1.145)	
Self * consumption level 4					0.625	
					(1.254)	
Community * consumption level 1					0.337	
					(0.634)	
Community * consumption level 2					1.100*	
					(0.645)	
Community * consumption level 3					1.541**	
					(0.677)	
Community * consumption level 4					0.780	
					(0.702)	
Consumption level 1					2.102***	
					(0.417)	
Consumption level 2					1.425***	
					(0.431)	
Consumption level 3					0.648	
					(0.473)	
Consumption level 4					0.590	
					(0.479)	
Self * log consumption						−0.674*
						(0.368)
Community * log consumption						0.225
						(0.296)
Log consumption						−1.518***
						(0.232)
Observations	340	5,958	230	5,430	5,796	5,796
Dependent variable mean	12.80	0.101	0.817	0.0494	0.0540	0.0540

Robust standard errors in parentheses. Regressions include stratum fixed effects. Standard errors clustered at village level.
***p < 0.01, **p < 0.05, *p < 0.1

3.7 Elite capture

One frequently cited concern in community targeting is the risk of elite capture. With the data we collected on elite networks and the experimental design of the elite subtreatment, we can empirically test for elite capture in community targeting, and provide meaningful evidence on the risk of elite capture in the Indonesian context. We provide a summary of the findings below. For expanded analysis, see Alatas *et al.* (2013a).

Table 9 tests for elite capture by regressing benefit receipt on a dummy for elite relatedness, controlling for consumption. Panel A shows this test for all leaders and their relatives in: the PMT (columns 1 and 4), community treatment overall (columns 2 and 5) and community treatment showing the differential effect of the elites-only subtreatment (columns 3 and 6). Because beneficiary lists may be tweaked during implementation, the tests are done using two different outcome variables: actual benefit receipt (columns 1–3) and presence on the original targeting list (columns 4–6).

None of the six cases in Panel A show any evidence that elites are more likely to receive benefits greater than they are entitled to given their consumption levels, even when elites make targeting decisions essentially behind closed doors (columns 3 and 6). In fact, estimates suggest that elites are, if anything, less likely to receive benefits than their consumption implies, although these effects become insignificant in some cases with additional controls (not shown). The coefficients on elite capture between PMT and community treatment (columns 1 and 2) are also indistinguishable.

Panel A treats all elites the same and thus may hide important heterogeneity between formal and informal leaders, who are subject to different incentives and constraints. Thus in Panels B and C we present the same results on subsamples of elites — formal leaders and their relatives in Panel B; informal leaders and their relatives in Panel C. Some results change in significance and magnitude, but the overall picture remains the same: elites of all kinds are, if anything, less likely to receive PKH than non-elites, even in the closed-door meetings.

Table 9 Elite capture in community targeting

	Benefit Receipt			Targeting List		
	PMT	Community	Community	PMT	Community	Community
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: All Elites			
Elite	-0.032**	-0.042***	-0.029	-0.017*	-0.030**	-0.029*
	(0.015)	(0.015)	(0.021)	(0.009)	(0.011)	(0.017)
Log Consumption	-0.096***	-0.124***	-0.124***	-0.035***	-0.074***	-0.074***
	(0.015)	(0.015)	(0.015)	(0.009)	(0.012)	(0.012)
Elite Subtreatment			-0.005			-0.013
			(0.024)			(0.019)
Elite x Elite Subtreatment			-0.027			-0.001
			(0.029)			(0.023)
Observations	1,863	1,936	1,936	1,996	2,000	2,000
Dependent Variable Mean	0.110	0.142	0.142	0.0431	0.0770	0.0770
			Panel B: Formal Elites			
Elite	-0.034**	-0.042***	-0.021	-0.017*	-0.018	-0.017
	(0.015)	(0.015)	(0.023)	(0.009)	(0.012)	(0.018)
Log Consumption	-0.097***	-0.125***	-0.126***	-0.035***	-0.075***	-0.076***
	(0.015)	(0.015)	(0.015)	(0.009)	(0.012)	(0.012)
Elite Subtreatment			-0.004			-0.013
			(0.023)			(0.018)
Elite x Elite Subtreatment			-0.042			-0.003
			(0.031)			(0.024)
Observations	1,863	1,936	1,936	1,996	2,000	2,000
Dependent Variable Mean	0.110	0.142	0.142	0.0431	0.0770	0.0770
			Panel C: Informal Elites			
Elite	-0.033*	-0.020	-0.018	-0.011	-0.040***	-0.051**
	(0.017)	(0.018)	(0.026)	(0.011)	(0.014)	(0.021)
Log Consumption	-0.097***	-0.127***	-0.127***	-0.036***	-0.074***	-0.074***
	(0.015)	(0.015)	(0.015)	(0.009)	(0.012)	(0.012)
Elite Subtreatment			-0.014			-0.017
			(0.024)			(0.019)
Elite x Elite Subtreatment			-0.004			0.022
			(0.038)			(0.029)
Observations	1,863	1,936	1,936	1,996	2,000	2,000
Dependent Variable Mean	0.110	0.142	0.142	0.0431	0.0770	0.0770

Note: Test of equality on elite related coefficient between columns (1) and (2) yields: Panel A: p-value 0.637; Panel B: p-value 0.702; Panel C: p-value 0.593.

3.8 Observables and unobservables in self-selection

In considering the impact of self-selection, it is interesting to disentangle two related effects that could be driving the strategy's efficacy, but have drastically different implications.

From the government's perspective, there are two ways in which the self-selection decision could affect targeting:

- Selection on observables: those households who have more assets, and are therefore less likely to pass the PMT, could be less likely to show up. This type of selection would save the government resources (since it would not have to interview people who are likely to fail the selection process), but it would not necessarily change the poverty profile of beneficiaries compared with automatic enrolment; and
- Selection on unobservables: conditional on a household's PMT score, those with higher unobservable consumption might be less likely to attend. This could arise if self-selection was based on the opportunity cost of time, or if households do not perfectly understand the construction of the PMT score. In these cases, introducing self-selection could lead to a poorer distribution of beneficiaries than automatic enrolment.

In Alatas *et al.* (2013b), we divide respondents' consumption into observable and unobservable characteristics to see which type of self-selection is occurring. We find that households self-select across both unobservable and observable characteristics, which suggests that both types of self-targeting have the potential to save costs in two ways:

- Observables: many households that would fail the proxy means test do not show up, saving them time and the government the cost of interviewing them; and
- Unobservables: many households may potentially pass the proxy means test despite being ineligible because of error in the proxy means test. Those that passed erroneously are less likely to show up, reducing inclusion error and saving the government the cost of paying transfers to non-eligible households.

3.9 Impacts on poverty rates

The analysis in section 3.2 showed that error rates differ significantly across the treatments. Targeting error rates, however, reflects only on intermediate outputs. Given that error rates are driven largely by those near thresholds, it is important to consider whether the treatments have differential impacts on real outcomes, such as the headcount poverty rate (the percentage of people who fall below the poverty line) and the poverty gap (the mean distance below the poverty line as a proportion of the line, counting the non-poor as having zero gap). Moreover, given that both treatments outperform PMT in targeting the very poor, it is possible that they may perform better at reducing the squared poverty gap (which places greater weight on reducing the poverty of the very poor), even if one or both perform worse in reducing the poverty headcount ratio.

We follow the methods used in Ravallion (2009) and simulate the effects of the different targeting methods on the headcount poverty rate, the poverty gap and squared poverty gap. In Table 10, we provide the results of the simulation for four transfer amounts:

- no transfer;
- the average *per capita* monthly PKH transfer in our sample (IDR20,000);

- half this average; and
- double this average.

We focus on the poor and very poor poverty lines, defining both at a low level in the consumption spectrum as is appropriate for PKH's targeting. Note that, despite the randomisation, there are statistically insignificant differences between the poverty rates in the different treatments as a result of sampling. For the simulations, we assume for all treatments the distribution of consumption from the PMT villages, so that we have exactly the same income distribution across treatments.

Table 10 shows the results of this exercise. At a *per capita* transfer size of IDR40,000 (roughly double the average transfer to beneficiaries in our sample), the three treatments reduce the poverty headcount by about 0.5 to 0.6 percentage points, led notably by the PMT (PMT: 15.01 per cent; self: 15.16 per cent; community: 15.14 per cent). Self-targeting shows a slight advantage over the other mechanisms in targeting the very poorest, as can be seen in its lower figures for the poverty gap, squared poverty gap and poverty headcount using the very poor poverty line. On many of these variables, self-targeting outperforms even the universal PMT. Note that these figures are generally similar, but less pronounced at other transfer levels.

Table 10 Simulated impact on poverty rates

Transfer size *per capita* (IDR)		Poverty line = poor				Poverty line = very poor			
		PMT	Hypothetical PMT	Self-targeting	Community targeting	PMT	Hypothetical PMT	Self-targeting	Community targeting
0	Headcount	15.62	15.62	15.62	15.62	5.98	5.9	5.98	5.98
	Poverty gap	2.80	2.80	2.80	2.80	0.93	0.93	0.93	0.93
	Poverty gap^2	0.79	0.79	0.79	0.79	0.24	0.24	0.24	0.24
10,000	Headcount	15.45	15.29	15.49	15.49	5.93	5.85	5.88	5.96
	Poverty gap	2.76	2.75	2.74	2.76	0.91	0.90	0.90	0.91
	Poverty gap^2	0.78	0.78	0.77	0.78	0.23	0.23	0.23	0.23
20,000	Headcount	15.27	15.10	15.45	15.40	5.82	5.70	5.63	5.95
	Poverty gap	2.73	2.71	2.74	2.74	0.89	0.88	0.87	0.90
	Poverty gap^2	0.77	0.76	0.77	0.77	0.23	0.22	0.22	0.22
40,000	Headcount	15.01	14.81	15.16	15.14	5.67	5.53	5.44	5.82
	Poverty gap	2.69	2.65	2.61	2.69	0.88	0.85	0.83	0.87
	Poverty gap^2	0.76	0.74	0.72	0.76	0.23	0.22	0.21	0.22

4. Programme costs and policy conclusions

4.1 Programme costs

Table 11 shows summary statistics of the various targeting treatments, including the costs of each treatment. These include administrative costs and costs incurred by households during the process (particularly in self-targeting and community targeting). We outline the costs below.

Table 11 Targeting costs and summary

	PMT (1)	Hyp Universal PMT (2)	Self-Targeting (3)	Community Targeting (4)
Eligible households that receive benefit	1,376	2,409	2,167	1,687
Eligible households that do not receive benefit	13,189	12,157	12,399	12,937
Ineligible households that receive benefit	8,946	11,122	6,621	16,813
Ineligible households that do not receive benefit	217,244	21,5068	219,569	209,377
Total annual benefits paid ($)	1,407,347	1,844,764	1,198,099	2,522,349
Total cost to households ($)	9,366	32,403	108,145	66,653
Total cost to beneficiary households ($)	1,176	1,411	13,400	10,741
Total cost to non-beneficiary households ($)	8,190	30,996	94,618	55,912
Total administrative costs in sample ($)	784,043	2,218,978	170,800	12,230
Total administrative costs, scaled ($)	120,378	340,673	-	

Note:
Estimates are totals for the 200 villages in our self-targeting sample. Columns 1 and 2 were estimated using PMT sample; column 3 using self sample; column 4 using community sample. Total population, eligible/ineligible households, annual benefits paid, costs to households and percentage of eligible households in the village are scaled in columns 1, 2 and 4 to match the figures from column 3. All monetary costs are reported in US$, using an exchange rate of IDR9,535 = USD1 (2 October 2012). Benefits per household are assumed to be IDR 1.3 million annually. Costs to households are calculated as time costs for travel, waiting, attending meetings and completing surveys (in PMT, just the cost of completing surveys) using the household average wage rate, plus transportation costs. Total administrative costs in sample are calculated based on per-village and per-neighbourhood costs cited by the Indonesian government at the time of the survey. Total administrative costs at scale in PMT are based on the actual cost of executing the PMT for an area with population 40 million. The costs of PMT are assumed to be linear in the number of households surveyed per village.

Hypothetical universal PMT

The universal PMT significantly improves error rates, but it does so at a significant cost. The intervention cost US$340,000 in our 200 villages (note that this is before accounting for possible additional economies of scale; these programmes may be significantly cheaper when conducting nationally). Thus, while it is a useful counterfactual to judge our other interventions, the costs of universal PMT mean it is rarely conducted without some form of prior additional targeting.

Self-targeting

Compared with the PMT, self-targeting presents several advantages. Self-targeting results in a significantly poorer distribution and lower error rates across the board. It is especially effective at reaching the very poorest households. As a result, it has a notably larger impact on the headcount of households below the very poor poverty line.

Self-targeting is cheaper than PMT methodologies, but it shifts the burden of targeting onto households and away from government. Households bear 40 per cent of total targeting costs, and because richer households have greater time costs, they bear the largest portion of this. Thus, the way in which one weights administrative costs to the government versus household costs in the overall social welfare function would change the way in which one viewed the total costs (again, this depends on the government's priorities). Interestingly, it is the very act of having self-targeting (a small fixed cost to apply) that results in selection. Increasing the cost of application did not result in improved selection.

Community targeting
Community targeting embodies a very different set of trade-offs compared with self-targeting. While it has drastically lower costs, higher satisfaction and lower errors based on subjective perceptions of wealth, by using objective consumption measures it yields slightly higher error rates and higher poverty headcounts.

Community targeting is significantly cheaper than other methods. It requires one visit to the village for data collection (cheaper than going door to door) and because the full PMT survey is not entered, it requires much less data entry. The total cost of community treatment is one-third less than scaled PMT and less than 30 per cent of the cost of self-targeting. On administration, the gains are even stronger: community treatment costs 10 per cent of scaled PMT and 7 per cent of self-targeting. Furthermore, given the larger number of households selected under community treatment, administrative targeting costs fall to less than 0.5 per cent of the amount paid out in benefits, compared with 9 per cent in scaled PMT. However, some costs are transferred to households: these are around 60 per cent of those they face in self-targeting, although households who eventually become beneficiaries bear a higher portion.

4.2 Policy conclusions

In this study, we explored different types of targeting methodologies: in particular, we compared Indonesia's current targeting policy (automatic enrolment) to self-targeting and community-targeting methodologies. Our main findings are:

1. *Self-targeting methodologies are a cost-effective way to improve targeting.* Self-targeting is a cost-effective mechanism for finding the very poor, and results in lower beneficiary consumption distribution. Comparing its cost structure to universal PMT, self-targeting may be potentially useful in areas with low poverty density since it may reduce the number of ineligible household interviews the government would have to conduct. Given its lower cost, self-targeting may be a useful tool for updating the list in years in which a targeting survey is not being conducted, in order to find the newly poor. There is, however, a trade-off: it shifts programme costs on to beneficiaries and has a lower satisfaction level. Thus, future designs of self-targeting programmes should consider how to induce selection while compensating those who bear the cost. Further, methodologies that can be used to improve satisfaction levels should also be incorporated.

2. *Increasing the ordeal in self-targeting may not improve it further.* The small fixed cost of having to apply at all induces the selection we observed in the self-targeting treatment. To induce further selection, the ordeals need to be increased prohibitively high. Therefore, in designing self-targeting programmes, a small fixed cost to apply may be preferable to very large costs.

3. *Community-based methodologies may be effective in improving targeting, depending on the government's preferences.* Community methodologies are much

cheaper to implement than PMT, resulting in allocations that are closer to the community's subjective beliefs on welfare and therefore in higher community satisfaction. However, they are slightly worse at targeting through an objective measure of consumption as the measure of truth. Community targeting works best in communities that are more networked, as these communities have better information on who is poor. The use of community-based methodologies depends on the government's preferences over subjective versus objective measures of consumption as an indicator of programme success. It also depends how the government views a gain in programme satisfaction, and thus potentially how easy it is to run a programme. For example, governments may prefer community methods in areas that are prone to high levels of community discord, even if there is a loss in targeting efficiency based on the objective measure of corruption.

References

Alatas, V, Banerjee, A, Hanna, R, Olken, BA and Tobias, J, 2012. Targeting the poor: evidence from a field experiment in Indonesia. *American Economic Review,* 102(4), pp.1,206–40.

Alatas, V, Banerjee, A, Hanna, R, Olken, BA, Purnamasari, R and Wai-poi, M, 2013a. *Does elite capture matter? Local elites and targeted welfare programs in Indonesia*. No. w18798. National Bureau of Economic Research.

Alatas, V, Banerjee, A, Hanna, R, Olken, BA, Purnamasari, R and Wai-poi, M, 2013b. *Self-Targeting: evidence from a field experiment in Indonesia*. MIT Working Paper.

Alderman, H, 2002. Do local officials know something we don't? Decentralization of targeted transfers in Albania. *Journal of Public Economics*, 83(3), pp.375–404.

Besley, T and Coate, S, 1992. Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs. *American Economic Review*, 82(1), pp.249–61.

Cameron, LA, 2002. *Did Social Safety Net Scholarships Reduce Drop-Out Rates During the Indonesian Economic Crisis?* Policy Research Working Paper. Vol. 2800. World Bank Publications.

Coady, D, Grosh, M and Hoddinott, J, 2004. Targeting outcomes redux. *The World Bank Research Observer*, 19(1), pp.61-85.

Conn, K, Duflo, E, Dupas, P, Kremer, M and Ozier, O, 2008. *Bursary Targeting Strategies: Which Method(s) Most Effectively Identify the Poorest Primary School Students for Secondary School Bursaries?* Mimeo.

Currie, J, 2006. Public Policy and the Income Distribution. In: DE Quigley, JM Auerbach, AJ Card, eds, 2006. *Public Policy and the Income Distribution*. Russell Sage Foundation.

Daly, A and Fane, G, 2002. Anti-poverty programs in Indonesia. *Bulletin of Indonesian Economic Studies*, 38(3), pp.309–29.

Galasso, E and Ravallion, M, 2005. Decentralized targeting of an antipoverty program. *Journal of Public Economics*, 89(4), pp.705–27.

Nichols, AL and Zeckhauser, RJ, 1982. Targeting transfers through restrictions on recipients. *The American Economic Review*, 72, pp.372–7.

Nichols, D, Smolensky, E and Tideman, TN, 1971. Discrimination by waiting time in merit goods. *The American Economic Review*, 61(3), pp.312–23.

Olken, BA, 2006. Corruption and the costs of redistribution: Micro evidence from Indonesia. *Journal of Public Economics*, 90(4), pp. 853–70.

PODES (Village Potential Statistics), 2008. Rand Corporation [online]. Available at: <http://www.rand.org/labor/bps/podes.html> [No longer available.]

PPLS (official poverty census), 2008. *Data Collection for Social Protection Programs.* Statistics Indonesia.

Ravallion, M, 2009. How relevant is targeting to the success of an antipoverty program? *The World Bank Research Observer*, 24(2), pp.205–31.

SUSENAS (National Socio-Economic Household Survey), 2010. Rand Corporation [online]. Available at: <http://www.rand.org/labor/bps/susenas.html> [No longer available.]

Publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://www.3ieimpact.org/publications/3ie-impact-evaluations/

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, *3ie Impact Evaluation Report 1*. Martinez, S, Naudeau, S and Pereira, V (2012)

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *3ie Impact Evaluation Report 2*. Burwen, J and Levine, DI (2012)

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, *3ie Impact Evaluation Report 3*. Casey, K, Glennerster, R and Miguel, E (2013)

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A (2013)

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, *3ie Impact Evaluation Report 5.* Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G, Solís, M (2014)

The impact of daycare on maternal labour supply and child development in Mexico, *3ie Impact Evaluation Report 6*. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E (2014)

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, *3ie Impact Evaluation Report 7*. Humphreys, M, Sanchez de la Sierra, R, van der Windt, P (2013)

Paying for performance in China's battle against anaemia, *3ie Impact Evaluation Report 8*. Zhang, L, Rozelle, S and Shi, Y (2013)

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, *3ie Impact Evaluation Report 9*. Ashraf, N, Bandiera, O and Jack, K (2013)

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N (2013)

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M, Mahapatra B (2014)

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R (2014)

Developing country governments often lack verifiable household income information to identify beneficiaries for social programmes. This report analyses results from a randomised controlled trial in Indonesia to better understand how targeting could be improved. The study randomised three targeting methodologies – proxy means testing (PMT), self-targeting and community targeting – across 600 villages. This impact evaluation shows that compared to PMT, self-targeting identifies poorer beneficiaries and has lower administrative costs. Community targeting, on the other hand, does less well at identifying the poor but results in beneficiaries being more satisfied with the programme.

Impact Evaluation Series

International Initiative for Impact Evaluation
c/o Global Development Network
2nd Floor, West Wing, ISID Complex
Plot No. 4, Vasant Kunj Institutional Area
New Delhi – 110070
India

3ie@3ieimpact.org
Tel: +91 11 4323 9494

www.3ieimpact.org