# SELF-TARGETING:
# EVIDENCE FROM A FIELD EXPERIMENT IN INDONESIA

VIVI ALATAS, ABHIJIT BANERJEE, REMA HANNA,
BENJAMIN A. OLKEN, RIRIN PURNAMASARI, AND MATTHEW WAI-POI

Abstract. In this paper, we show that adding a small application cost to a social assistance program can substantially improve targeting because of the self-selection it induces. We conduct a randomized experiment within Indonesia's Conditional Cash Transfer program that compares two of the most common methods of targeting welfare programs in the developing world: in one, beneficiaries first need to apply for the program, and then an enumerator visits them at home and determines their eligibilty based on a proxy-means asset test; in the other, they are visited directly by the enumerator and automatically enrolled if they qualify based on the same proxy-means test. When applications were required, we find that the poor are more likely to apply than the rich, even conditional on whether they would pass the asset test. On net, the villages where applications were required have a much poorer group of beneficiaries than automatic enrollment villages. However, marginally increasing the cost of applying does not necessarily improve targeting: while experimentally increasing the distance to the application site reduces the number of applicants, it screens out both rich and poor in roughly equal proportions. Estimating our model of the enrollment choice suggests that our results are largely driven by the rich forecasting that they have a very small likelihood of passing the asset test, and so not bothering to apply, which in aggregate substantially improves targeting efficiency. The results suggest that the combination of the small cost and the final screening gives this class of mechanisms the ability to achieve many of the benefits of self-selection without imposing onerous ordeals on program beneficiaries.

---

*Date*: November 2013.
Affiliations: Alatas, Purnamasari, Wai-Poi: World Bank. Banerjee, Olken: MIT, BREAD, CEPR, J-PAL, and NBER. Hanna: Harvard Kennedy School, BREAD, CEPR, J-PAL, and NBER. Contact email: bolken@mit.edu.
This project was a collaboration involving many people. We thank Jie Bai, Talitha Chairunissa, Amri Ilmma, Donghee Jo, Chaeruddin Kodir, He Yang, Ariel Zucker, and Gabriel Zucker for their excellent research assistance, and Raj Chetty, Esther Duflo, Amy Finkelstein, and numerous seminar participants for helpful comments. 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 Social Affairs Department (DepSos), and SurveyMetre 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 NIH under grant P01 HD061315. All views expressed are those of the authors and do not necessarily reflect the views of the World Bank, TNP2K, Mitra Samya, DepSos, or the Indonesian Central Bureau of Statistics.

## 1. Introduction

In designing targeted aid programs, a perennial problem is how to separate the poor from the rich. One strategy for doing so is to impose program requirements that are differentially costly for the rich and the poor, in order to induce the poor to participate while dissuading the rich from doing so (Nichols, Smolensky and Tideman, 1971; Nichols and Zeckhauser, 1982; Ravallion, 1991; Besley and Coate, 1992). These "ordeal" mechanisms are quite common: welfare programs, from the Works Progress Administration (WPA) in the United States during the Great Depression to the National Rural Employment Guarantee Act (NREGA) right-to-work scheme in India today, often have manual labor requirements to receive aid, and subsidized food schemes often provide lower quality food so that those who can afford tastier food choose not to purchase the subsidized products.

The challenge with these ordeal mechanisms is that they can be quite inefficient: in order to dissuade the rich from participating, the poor are forced to incur substantial utility costs in order to receive transfers, whether by toiling in the hot sun or eating unappetizing food. In this paper, we ask whether much smaller costs can still achieve substantial self-selection. In particular, we show that when the cost entails applying for benefits, and there is some ex-ante uncertainty in whether an application will be successful, even small application costs can generate significant improvements in targeting. The reason is that if the rich correctly foresee that they face only a very small chance of slipping through the screening procedure and receiving benefits, they may not bother to apply, and the resulting reductions in inclusion error may substantially improve the degree to which the program is targeted to the poor. Self-selection can also reduce the degree to which the poor are excluded from the program compared to alternative, top-down targeting schemes if it encourages the very poor who live at the margins of society to make themselves known to government staff by applying for the program.

Of course, it is not ex-ante obvious that these types of application costs will necessarily improve targeting. The recent movement in behavioral economics, for example, has emphasized that people tend to over-respond to small costs, and that these types of application costs may dissuade the poor from applying. For example, the hassle costs with applying for social assistance programs, such as food stamps and welfare payments, have been cited as a reason for low takeup of these programs, and there have been policy suggestions that these types of hassles should be removed from application processes to encourage program takeup (Bertrand, Mullainathan and Shafir, 2004; Currie, 2006). It is also theoretically possible that the poor face a variety of other deterrents from applying, such as self-control problems (e.g., Madrian and Shea, 2001), stigma (e.g., Moffitt, 1983), and poor access to information about government programs (e.g., Daponte, Sanders and Taylor, 1999). Ultimately, whether making people apply for programs achieves better targeting than automatic enrollment is an empirical question.

In this paper, we investigate whether in fact making people apply for programs improves or worsens targeting by conducting a randomized experiment in the context of Indonesia's Conditional Cash Transfer program, known as PKH. Conditional cash transfer programs have spread rapidly throughout the developing world and are present in over 30 countries today. In Indonesia, the PKH program provides beneficiaries with US $130 per year for 6 years, and is one of the country's largest social assistance programs, aiding about 2.4 million households. The program is aimed at the poorest 5-10 percent of the population, with eligibility determined based on a weighted sum of about 30 easy-to-observe assets (e.g., size of house, materials used to construct household roof, motorbike ownership).

Working with the Indonesian government, we experimentally varied the enrollment process for PKH across 400 villages, comparing a process that required potential applicants to apply for the program to the status quo, where eligible households were surveyed by the government and were automatically enrolled if they qualified. In both cases, eligibility was determined based on an asset screen, known as a proxy-means test (PMT), so the key difference we studied was whether households had to actively apply or were instead automatically enrolled based on a government survey. These two approaches to targeted social-assistance programs – automatic enrollment based on a top-down survey or enrollment being limited to who actively apply – represent the two most common ways of determining beneficiary lists for targeted transfer programs in the developing world (Grosh et al., 2008; Kidd and Wylde, 2011).[1]

In our study, in villages randomized to receive the application process (henceforth "self-targeting" villages), households that were interested in the program were required to go to a central registration site to take an asset test administered by the statistics office. This entailed both traveling a few kilometers to the application site and waiting in line to apply. Within these areas, we randomly varied the application costs along two dimensions: the distance to the application site and whether one or both spouses needed to be present to apply. These application costs – missing about a half day's work, and traveling a few kilometers – pale in comparison to the benefits on offer, which amount to $130 per year for 6 years. In control areas, the status quo procedure – automatic enrollment – was followed: the statistics office, working with local government officials, drew up a list of potential beneficiaries; interviewed everyone at their homes; and then automatically enrolled those who passed using the same asset test that was used in self-targeting.

We begin with a description of the experiment and the data. We then ask what we would expect from such an experiment on purely *a priori* grounds. Specifically, we adapt the classical theory of self-selection into social programs developed by Nichols, Smolensky and Tideman (1971), Nichols and Zeckhauser (1982), Besley and Coate (1992) and others to a context where after selecting into applying, one receives the program stochastically, with the probability of receiving the program declining with income. The fact that receiving benefits is stochastic but declining with income captures the fact that most screening mechanisms, including but not limited to proxy-means tests, do differentiate between rich and poor, but not perfectly, so that people cannot exactly forecast before applying whether they will in fact be eligible.[2] The standard Nichols and Zeckhauser (1982)

[1]Examples of automatic enrollment PMTs include the Mexican Progresa program, the Columbian social assistance programs, and the Indonesian cash transfer programs; examples of self-selection based PMTs include the expansion of Progresa under the name Oportunidades to urban areas, the Chilean social assistance system, the Costa Rican SIPO system, and the Mongolian Child Money Program (Castaneda and Lindert, 2005; Hodges et al., 2007; Coady and Parker, 2009; Martinelli and Parker, 2009).

[2]The fact that people cannot perfectly forecast program eligibility before applying is not just limited to developing country contexts; in the Oregon Health Insurance Experiment, for example, about half of those who applied for the health insurance were in fact ineligible, a much lower rate than one would expect from the population at large but still substantial (Finkelstein et al., 2012).

self-selection idea depends on a single-crossing property, where the ordeal is more costly for rich than poor. Time-based ordeals are the canonical example, since the rich presumably have a higher opportunity cost of time than the poor. In this context, we illustrate that there are a number of reasons why requiring people to spend time traveling to the application site and applying need not necessarily generate single-crossing: the poor and rich may have different tools for overcoming the costs which may make them less costly for the rich; there may be income effects so that the same monetary cost imposes a differential utility cost on the poor; and the distribution of idiosyncratic costs of applying may mean that there are more poor on the margin of being deterred by increases in costs than rich. On the other hand, the fact that the probability of receiving benefits is downward sloping in income is a strong force in favor of self-targeting improving targeting: since the rich face only a very small chance of passing through the proxy-means test if they apply, they may not bother, even if the costs are relatively small.

Our empirical analysis then proceeds in four stages. First, we begin by examining who chooses to self-select into applying for the program in the 200 villages with the application-based process. To do so, we utilize data on households' per capita consumption that we collected before the program was announced or targeting began. We find that the probability of self-selecting to apply for the program is decreasing in a household's per capita consumption, i.e., that the poor are always more likely to apply than the rich. Decomposing consumption into that which is potentially observable to the government (i.e. the part that can be predicted based on observable assets) and the unobservable residual, we show that those who apply are poorer on both observables and unobservables than those who choose not to apply. This implies that self-selection can not only potentially save resources (since many who would fail the asset test, i.e., have high observables, are no longer tested), but that it also has the potential to improve targeting even over a universally-administered asset test (since those that apply are poorer on unobservables than the population at large). However, we also find evidence for the view that self-targeting may screen out some of the poor: for example, only about 60 percent of the very poorest apply under self-targeting.

The question, though, for most governments is not necessarily how self-targeting would perform relative to a counterfactual of no error, but rather how it would compare against the next best alternative targeting strategy. The second step of our empirical analysis is to use the experiment to compare self-targeting with the status quo of automatic-enrollment, in which the government conducted the asset test for all potential beneficiaries (chosen through prior asset surveys and consultations with village leadership) at their homes and automatically enrolled those that passed. Compared against this real alternative, we find that per capita consumption was 21 percent lower for beneficiaries in the self-targeting villages than those under the status quo. Moreover, exclusion error was actually less of a problem in self-targeting than in the status quo: the very poorest households were twice as likely to receive benefits in self-targeting than in control areas. These findings are not entirely driven by the fact that the government ineptly chose whom to interview under the status quo: supplementing the government's asset test data in the automatic enrollment villages with asset data that we independently collected for those not interviewed, we find that the beneficiaries under self-targeting would still be, on average, poorer than those under a "hypothetical" universal automatic enrollment system where everyone is interviewed for the asset test. Intuitively, this is possible because – as we showed above – self-selection includes selection on unobservables. That is, conditional on passing the asset test, those that self-select into applying have lower consumption than the average person in the population.

The third step in our empirical analysis is to consider whether marginal increases in the severity of the ordeal further increase targeting performance. We examine the results from experimentally varying the distance to the registration site (increasing travel costs) and the number of household members required to be present at the application site (increasing opportunity costs of time for the family). We find no evidence that these marginal increases in application costs further improve selection. In some cases, they reduce overall takeup, but they do not differentially discriminate between rich and poor.

The theoretical model outlines a number of reasons why marginal increases in the extent of ordeals might not necessarily improve targeting. To understand which factors are in fact empirically relevant for takeup, the final step of our empirical analysis uses Generalized Method of Moments to estimate a CRRA utility version of our model with logit shocks. We use the average show-up rates in the far distance treatment for each income quintile as moments. Since we estimate the model using only one experimental sub-treatment and cross-sectional differences in distances to fit the model, not the experimental variation, we can check that the model's predictions provide a reasonable approximation of the experimental findings, which indeed they do.

We use the estimated version of the model to see which of the various mechanisms we outlined in the theoretical section lie behind the fact that marginal increases in the extent of the ordeal do not seem to differentially improve selection. Simulations from the estimated model suggest that, of the theoretical mechanisms we outline, neither curvature of the utility function nor differential travel technology is driving that result. Instead, the model suggests that the key driver of selection is the fact that rich households forecast that they have a very small likelihood of receiving benefits conditional on applying and therefore do not bother to apply if there is any cost of applying. This helps explain both why small costs can produce substantial selection but marginally increasing the intensity of the costs reduces overall application rates without substantially improving targeting.

The remainder of the paper is organized as follows. Section 2 discusses the setting, experimental design, and data. Section 3 introduces our model,which revisits the standard screening model in light of curvature in the utility function, differential access ways of dealing with costs, and idiosyncratic shocks. Section 4 examines the self-targeting data to ask who chooses to apply for the program. Section 5 uses the experiment to compare self-targeting with the status quo PMT-based approach. Section 6 examines the marginal effect of targeting when the ordeal is changed experimentally. Section 7 estimates the model to help shed light on which of the possible theoretical mechanisms we outline best explains the results. Section 8 concludes.

## 2. Setting and Experimental Design

2.1. **Setting: The PKH Program.** This project explores self-targeting mechanisms within the context of Program Keluarga Harapan (PKH), a conditional cash transfer project administered by the Ministry of Social Affairs (DepSos) in Indonesia. The program targets households with per capita consumption below 80 percent of the poverty line (approximately the poorest 5 percent of the population we study) and that meet the demographic requirements of having a pregnant woman, a child between the ages of 0 to 5, or children below the age of 18 years old that have not finished the nine years of compulsory education. Program beneficiaries receive direct cash assistance ranging from Rp. 600,000 to Rp. 2.2 million (US$67-US$250) per year—or about 3.5 to 13 percent of the average yearly consumption by very poor households in our sample—depending on their family composition, school attendance, pre/postnatal check-ups, and completed vaccinations.[3] The payments are disbursed quarterly for up to six years. In 2013, approximately 2.4 million households are enrolled in the program. Determining whether households fall below the consumption requirement ("targeting") is difficult because per capita consumption is not easily observed by the government. Instead, PKH uses a proxy-means test (PMT) approach with automatic enrollment for all households that meet the demographic requirements and pass a proxy-means test. Specifically, every three years, enumerators from the Central Statistical Bureau (BPS) conduct a survey of households nationwide who are potentially eligible for anti-poverty programs, including but not limited to PKH. They survey all households that were included on previous surveys (regardless of whether they previously qualified or not) and supplement this list with recommendations from local leaders and their own observations of the kinds of houses that the households inhabit. After passing an initial pre-screening, each household is asked a series of about 30 questions, including attributes of their home (e.g., wall type, roof type), ownership of specific assets (e.g., motorcycle, refrigerator), household composition, and the education and occupation of the household head. These measures are combined with location-based indicators, such as population density, distance to the district capital and access to education. Using independent survey data, the government then estimates the relationship between these variables and the household per capita consumption to generate a district-level formula for predicting consumption levels based on the responses to the survey. Individuals with predicted consumption levels below each district's very poor line were eligible for the program.

Figure 1 shows the probability of passing the asset test and being determined eligible for the program as a function of log per-capita income, as estimated from our baseline data. Note that the particular function used to map assets to eligibility is estimated by the government separately for each district and for urban and rural areas, which is why several different downward-sloping curves are visible in the Figure. Several key points are worth observing about this function. First, it is strongly downward sloping – the poor are much more likely to receive benefits than the rich. Second, there is substantial noise in the process, driven by how hard it is to accurately estimate consumption from assets and the fact that PKH targets the bottom approximately 5 percent of the population – even the very poorest rarely have more than a 40 percent chance of receiving benefits, and even those with incomes more than twice the target threshold (i.e. about 13 log points, as opposed to the cutoff of about 12.3 log points) still have as much as a 5 to 10 percent chance of receiveing benefits.[4]

[3]Note, however, that although PKH is formally a conditional cash transfer program, with transfers dependent upon health takeup and school enrollment, these conditions are typically not enforced in practice, so this can be thought of as closer to a 'labeled' cash grant, as in Benhassine et al. (2013).

[4]The regressions of income on assets underlying Figure 1 have an R2 of between 0.4-0.6. One might be concerned that the relatively poor prediction in Figure 1 is simply because the government is using the wrong algorithm. This does not, however, appear to be the case. In separate ongoing work, we have examined a wide range of non-linear

2.2. **Sample Selection.** This project was carried out during the 2011 expansion of PKH to new areas. We chose 6 districts (2 each in the provinces of Lampung, South Sumatra, and Central Java) from the expansion areas to include a wide variety of cultural and economic environments. Within these districts, we randomly selected a total of 400 villages, stratified such that the final sample consisted of approximately 30 percent urban and 70 percent rural locations.[5] Within each village, we randomly selected one hamlet to be surveyed.[6] These hamlets are best thought of as neighborhoods that consist of about 150 households and that each has its own administrative head, whom we refer to as the hamlet head.

2.3. **Experimental Design.** We randomly allocated each of the 400 villages to one of two targeting methodologies: self-targeting or an automatic enrollment system, i.e. the status quo.[7]

2.3.1. *Automatic Enrollment Treatment.* In Indonesia, the automatic enrollment treatment is the status quo, and the procedure discussed in Section 2.1 was followed. For each hamlet in this treatment, the government Bureau of Statistics (BPS) enumerators were given a pre-printed list of households from the last targeting survey (PPLS, 2008). When they arrived at a village, the enumerators showed the list to the village leadership and asked them to add any households to the list that they thought were inappropriately excluded. The enumerators also had the option of adding households to the list of interviewees if they observed that a household was likely to be quite poor. For each household that was interviewed, a computer-generated poverty score was generated using the district-specific PMT formulas.[8] A list of beneficiaries was generated by selecting all households with a predicted score below the score cutoff for their district. [9]

machine learning algorithms to see if they can improve the predictive power in Figure 1, which is based on OLS. We found only very small improvements appear possible, so the imperfect prediction appears a fundamental challenge of predicting consumption from assets, not a function of the algorithm used.

[5]The sampling unit was a *desa* in rural areas and a *kelurahan* in urban areas. For ease of exposition, we henceforth refer to both as villages.

[6]Both desa and kelurahan are administratively divided by neighborhood into sub-villages known variously as *dusun*, *RW*, or *RT*. For ease of exposition, we henceforth call them "hamlets." In rural areas, each hamlet ranges from about 30-330 households, while in urban areas, they each range from 70-410 households.

[7]We also randomly assigned an additional 200 villages to a "hybrid treatment" (see Alatas, Banerjee, Hanna, Olken, Purnamasari and Wai-poi (2012)).

[8]The PMT formulas were determined using household survey data from SUSENAS (2010) and village survey data from PODES (2008). On average, these regressions had an R-squared of 0.52. The questions chosen for the PMT survey were those that the government was considering for the next nationwide targeting survey (the "PPLS 11").

[9]The only difference between the PMT formula used in the automatic enrollment treatment and the self-targeting treatment was that, in automatic enrollment, for each potential interviewee, the enumerator conducted an initial five question pre-screening; those households who passed the pre-screening were given the full PMT survey.The prescreening consists of 5 questions: is the household's average income per month in the past three months more than Rp. 1,000,000 (USD$110); was the average transfer received per month in the past three months more than Rp. 1,000,000 (USD$110); did they own a TV or refrigerator that cost more than Rp. 1,000,000 (USD$110); was the value of their livestock productive building, and large agricultural tools owned more Rp. 1,500,000 (USD$167); did they own a motor vehicle; and did they own jewelry worth more than Rp. 1,000,000 (USD$100). Households that answered yes on either four or five of the questions were instantly disqualified and the survey ended. Of the 6,406 households on the potential interviewee list, 16 percent were eliminated based on the initial screen, and 5,383 households (or about 37.8 percent of each hamlet) were given the full PMT survey of 28 questions. The idea is that these thresholds are so high that any household answering yes to four or more of these questions would have been eliminated anyway by the PMT. To verify that this does not affect the results, we have rerun the main experimental analysis (e.g. Tables 4 and 5) dropping from our sample any household in either treatment that would have failed this pre-screen, using answers to the same questions in the baseline survey, so that in this sample the PMT used in automatic enrollment and self targeting were exactly identical. The results are virtually unchanged when we use this sample.

2.3.2. *Self-Targeting Treatment.* The enrollment criteria for both the demographic and consumption criteria under the self-targeting mechanism were the same as in automatic enrollment, but households were required to apply at a central registration station if they were interested in the program. The fact that households needed to self-select means that some households who might have been automatically enrolled would not receive benefits because they chose not to apply. Conversely, some households who may have been forgotten or passed over when the government compiled the list of households to be interviewed could apply and ultimately receive benefits.

The self-targeting treatment proceeded as follows: to publicize the application process, a community facilitator from a local NGO (Mitra Samya) visited each village to inform the village leaders about the program, to brainstorm about the best indicators of local poverty with the leaders, and to set a date for a series of hamlet-level meetings that were aimed at the poor.[10] In these hamlet-level meetings, the facilitator described the PKH program and explained the registration process. In particular, the facilitators stressed that the program was geared towards the very poor. For example, they listed examples of questions that would be asked during the interview (e.g., type of house, motorbike), informed households that there would be a verification stage post-interview, and highlighted a set of local poverty criteria (the criteria that locals would typically use to characterize very poor households). Though they did not convey the exact criteria used, the goal was to ensure that the households generally understood that their chances of obtaining PKH conditional on showing up to be interviewed would be much higher for the poor than for the rich.

Registration days for each area were scheduled in advance based on the number of predicted applicants and their relative proportion within the hamlet.[11] During the registration days, the BPS enumerators were present at the registration station from 8AM to 5PM. Households were required to come to the registration site if they wished to apply. Once they arrived, they were signed in and given a number in the queue. When their number was called, BPS interviewed the household to collect the same data that was conducted in the PMT interview.

Households that applied were subsequently categorized by eligibility based on the PMT regression formula and the district-specific very poor line, using the same PMT formula and questions as in the automatic enrollment treatment. Any household that was both classified as very poor based on the assets they disclosed in their interview and which had also been visited by government enumerators in the 2008 poverty census and found to be very poor (about 37 percent who passed the interview at the registration site) was selected as a PKH recipient. All other households that classified as very poor based on their interview were subjected to a verification process: Government surveyors went to their homes to collect data on the same set of asset questions. The results of this home-based survey were used, with the same PMT regression formula and poverty lines, to determine the final list of beneficiaries. About 68 percent of those who got to the verification stage were ultimately

[10]The local poverty indicators generated in the meeting were not used for targeting, but were instead used by community members in the socialization process to help villagers understand how the PMT screening would operate.

[11]Specifically, we estimated the predicted number of people who would show up to be interviewed using the pilot data. We regressed the number of people who showed up on the number of households in the village and the number of poor households. BPS staff were assigned based on these predicted show-up rates, assuming a capacity to interview 34 households per day and a 25 percent buffer.

considered eligible after the verification.[12] Within self-targeting treatment villages, we varied how the self-targeting was conducted in order to vary the costs of registration. Specifically, we conducted two sub-treatments:

(1) *Distance sub-treatment*: We experimentally varied the distance to the registration site. The idea was to vary the time and travel costs required to sign up, while ensuring that all locations could still potentially be reached by walking, so as not to impose substantial financial transportation costs on poor households.[13] In the urban areas, we randomly allocated villages to have the registration site at the sub-district office (far location) or the village office (near location). In rural areas, where distances are greater than the urban areas, villages were randomly allocated to have the registration site at the village office (far location) or in the sub-village (near location).[14]

(2) *Both spouse sub-treatment*: We experimentally varied whether one or two household members were required to come to the registration site. In half of the self-targeting treatment villages, households were told that any adult in the household could do so. Given that the program was geared towards women, we expected that mostly women would apply. In the other half of the villages, we required that both the husband and the wife jointly apply at the registration site. Note that there was a fear of screening out poor households where the primary wage earner had migrated for work. Thus, if the spouse was unable to attend due to a pre-specified reason (illness, out of village for work), the household was required to bring a letter signed by the hamlet head providing the reasons for the spouse's unavailability, the rationale being that obtaining the letter in advance would still be costly to households. On average, 29 percent of applicants in spouse sub-treatment villages provided such a letter.

On net, these application costs are small relative to the potential benefits received. To see this, we can compute the costs of applying from the household survey (described in more detail Section 2.5 below) by adding up reported time and monetary costs to travel to the location where the interview will take place (which we obtain in the baseline survey for all households, even before they know about the targeting program), as well as the average time people spent waiting multiplied by an estimate of the household's likely wage rate. (See Section 7 for more details on this calculation.) On average, the total time and monetary cost of applying is about Rp. 17,000 (US $1.70) per household, with costs being higher for wealthier households with higher implied wage rates. By contrast, the per-household benefits average Rp. 1.3 million (US $130) per year for 6 years. For households with very small probabilities of receiving benefits conditional on applying, it may not make sense to apply, but for those with high probabilities of receiving benefits, the expected return from doing so appears substantial.

[12]The fact that there was substantial under reporting of assets in the initial interview, and therefore that only $\frac{2}{3}$of households passed the home-based asset verification, is consistent with the Mexican experience with targeting in Progresa (Martinelli and Parker, 2009).

[13]Thornton et al 2010 found that reducing distance to enrollment for health insurance increased enrollments substantially, but did not examine selection effects.

[14]The distance sub-treatment was violated in four villages: in the first village, a longstanding ethnic tension caused a large subset of the village to refuse to participate in interviews in a certain hamlet, so the interviewers held interviews for a day in another hamlet; in the second village, a hamlet was a 4-5 hour walk away from the village office, so the interviewers set aside a day to go to that hamlet; in the third and fourth villages, the village leaders insisted the registration site be moved closer to the village. All analysis reports intent-to-treat effects where these four villages are categorized based on the randomization result, not actual implementation.

2.4. **Randomization Design and Timing.** We randomly assigned each of the 400 villages to the treatments (see Table 1). In order to ensure experimental balance across geographic regions, we stratified by 58 geographic strata, where each stratum consisted of all the villages from one or more sub-districts and was entirely located in a single district. We then randomly and independently allocated each self-targeting village to the sub-treatments, with each of these two sub-treatment randomizations stratified by the previously defined strata and the main treatment.

From December 2010 to March 2011, an independent survey firm (Survey Meter) collected the baseline data from one randomly-selected hamlet in each village. After surveying was completed in each sub-district, the government conducted the targeting treatments. The targeting treatments thus occurred from January through April 2011.[15] SurveyMeter conducted a first follow-up survey in early August 201l, after the targeting was complete, but before the beneficiary lists were announced to the villages. Fund distribution occurred starting in late August 2011.[16] Finally, we conducted a second endline survey in January 2012 to March 2012, after two fund distributions had occurred.

2.5. **Data, Summary Statistics and Balance Test.**

2.5.1. *Data Collection.* We collected three main sources of data:

*Baseline Data*: The baseline survey was completed in each sub-district before any targeting occurred, and up to this point, there was no mention of the experiment in the villages. Within each village, we randomly selected one hamlet, and within that hamlet, we randomly sampled nine households from the set of those who met the demographic eligibility requirements for PKH, as well as the sub-village head, for a total of 3,998 households across the 400 villages. The survey included detailed questions on the household's consumption level, demographics, and family networks in the villages. We also collected data for all of the variables that enter the PMT formula, so that we could calculate PMT scores for each surveyed household.

*Targeting Data*: We obtained all of the targeting data from the government, including who was interviewed, all data from the interview (either at interview site or at home, or both), each household's predicted consumption score, and whether the household qualified to receive PKH. For the self-targeting villages, we additionally asked the government to record data on each step of the process (e.g. where and when the registration meetings occurred, how the socialization was done in each village, etc.).

*Endline Surveys*: We administered two endline surveys, both of which were conducted by SurveyMeter. The first endline survey occurred in August 2011, prior to announcements of the beneficiary lists. We surveyed up to three beneficiary households per village and revisited one household from the baseline survey per village in 97 randomly chosen automatic enrollment villages and 193

[15]There was no mention of the targeting process until SurveyMeter had completed the baseline survey in the entire sub-district. The mean time elapsed between the baseline survey and the commencement of targeting activities was 22 days.

[16]Note that after the experiment selected beneficiaries, the Department of Social Affairs realized it had additional funds available and decided to increase the number of people who received the program to also 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. This was completely unanticipated by all involved during the targeting experiment. In calculating "beneficiaries" for purposes of experimental evaluation below, we do not include these additional households.

self-targeting villages, for a total sample of 1,045 households.[17] In this survey, we collected detailed data on the households' consumption level, as well as respondents' experience and satisfaction with the targeting process (e.g., whether they applied, how long they waited to be interviewed). In addition, for all beneficiary households, we collected additional data on demographics, family networks, relationships with local leaders, and employment. We conducted the second endline from January 2012 to March 2012, after two rounds of PKH fund distribution. In this survey, we revisited all ten of the baseline households, collecting consumption data, as well as data on satisfaction with PKH.

2.5.2. *Summary Statistics and Experimental Validity.* Table 2 shows the flow of surveyed households through the experiment. Column 1 shows the total number of households in the baseline survey in each of the two primary treatments. The next columns show the number of households who applied to be interviewed for self targeting (754 out of 2,000, or 38 percent) or were interviewed as part of the automatic enrollment treatment (706 out of 1,998, or 35 percent). Column 3 shows the number of baseline households who were ultimately chosen as beneficiaries (73 out of 2,000, or 3.65 percent, in self-targeting; 86 out of 1,998, or 4.3 percent, in automatic enrollment).

Appendix Table A.1 presents summary statistics and a check on the the experimental validity using data from the baseline survey and a village census. Note that we chose all of the variables for the table prior to analyzing the data. Column 1 presents the mean and standard deviations of each variable in villages in the automatic enrollment treatment, while this information is provided for the self-targeting villages in Column 2. Column 3 shows the difference (with associated standard errors). Column 4 shows this difference after controlling for stratum fixed effects. Only 1 of the 20 differences presented is statistically significant (at the 10 percent level), confirming that the treatment villages are balanced at the baseline. At the bottom of Columns 3 and 4, we provide the p-value from a joint test of the treatment across all baseline characteristics that we consider. The p-values of 0.99 and 0.67, respectively, confirm that the groups are balanced in the baseline.

## 3. Model

3.1. **Model Set-up.** In this section, we briefly re-examine self-selection into a welfare program based on the expected benefits and costs of applying, incorporating the fact that a household's chance of receiving benefits is stochastic but declining in income. We assume that households have a utility function $U(x)$, where $x$ is current consumption. Households vary in their per-period labor income, denoted by $y$, but for a given household this is the same number in both periods. The application cost is denoted by $c\,(l, y)$, where $l$ is the distance to the registration site. Conditional on applying, households have a probability $\mu(y)$ of passing the asset-based test and actually qualifying for the program ($\mu'(y) \leq 0$).[18] This $\mu(y)$ function captures the fact that, as shown in Figure 1, a household's chance of receiving the program is much higher for poor than rich, but still uncertain.

[17]Due to safety and travel concerns that were independent of the project, the survey company asked that that we did not return to 10 villages in endline 1 and 13 villages in endline 2. These were spread among treatment and control villages.

[18]Note that in the model, households understand the $\mu(y)$ function. Empirically, this seems plausible, as similar PMT-based exercises had been done several times in the past in these villages, in 2005 and 2008, for use in other programs.

If the household qualifies for the program, it receives an additional income $b$ in the future period (for simplicity, we assume there is just one future period). Otherwise, it receives no additional income. Finally, assume that the household starts with no assets and cannot borrow. This is consistent with the evidence that many poor, and perhaps even not so poor, households in developing countries tend to be credit constrained. This, combined with the assumption that the household discounts future utilities (the discount factor is $\delta < 1$), and the fact that in our model future consumption is always weakly higher rules out savings and tells us that consumption in a given period is just current income net of costs.

To complete the description of the model, assume that each person who applies receives a random utility shock, $\varepsilon$, that encourages him to go to register, and $F(\epsilon)$ is the distribution of $\epsilon$. These idiosyncratic shock terms will be important in Section 7 below when we estimate the model explicitly.

Taken together, the household's expected utility upon applying is:

$$U\left(y - c\left(l, y\right)\right) + \mu(y)\delta U\left(y + b\right) + \left(1 - \mu(y)\right)\delta U\left(y\right) + \varepsilon. \tag{1}$$

If the household does not apply, expected utility is:

$$U\left(y\right) + \delta U\left(y\right). \tag{2}$$

The expected gain from applying is the difference, i.e.

$$U\left(y - c\left(l, y\right)\right) - U\left(y\right) + \mu(y)\delta\left[U\left(y + b\right) - U\left(y\right)\right] + \epsilon. \tag{3}$$

It will turn out to be convenient to define:

$$g(y, l) = U\left(y - c\left(l, y\right)\right) - U\left(y\right) + \mu(y)\delta\left[U\left(y + b\right) - U\left(y\right)\right] \tag{4}$$

to denote the net gain without the shock. The household will apply if the expected utility from doing so is larger than staying home, i.e., if $-g\left(y, l\right) \leq \epsilon$. The fraction of households that will apply at a particular level of income $y$ is given by $1 - F(-g(y, l))$. We are interested in how an increase in distance, $l$, affects $1 - F(-g(y, l))$ at different levels of $y$.

3.2. **Analysis.** In this section, we start with the most basic model and add elements to the model one-by-one in order to understand how each element affects the type of household that applies.

3.2.1. *The Benchmark Case.* Suppose that the utility function is linear ($U(x) = x$) and that the time cost of applying is also linear in distance: $\tau l$.[19] For someone who earns a wage, $w$, this imposes a monetary cost of $\tau l w$. If we assume that wages are proportional to income, $w = \alpha y$, then the monetary application cost can be written as $\tau l \alpha y$. Assume also that there are no shocks ($\epsilon \equiv 0$). In this case, $g\left(y\right)$ simplifies substantially, and a household applies if:

$$-\tau l \alpha y + \delta \mu(y) b \geq 0. \tag{5}$$

[19]The linearity in time cost may be unrealistic since it includes both travel time and wait time, which are unlikely to be linear in distance (though it may be increasing in distance since the further it is, the harder it is go home and come back later if the wait time is particularly long). However, nothing really turns on the linearity assumption, and it simplifies the model.

Since the left hand side of this expression is decreasing in $y$, this expression defines a cutoff value $y^*$ such that those who have incomes less than $y^*$ apply, and those who have incomes greater than $y^*$ do not apply. Moreover, an inspection of equation (5) shows that $\frac{\partial y^*}{\partial l} < 0$, that is, making the ordeal more onerous increases the degree of selection and implies that the set of people who apply will be poorer. This simple expression captures the basic intuition for using ordeal mechanisms for selection captured by Nichols and Zeckhauser (1982).

3.2.2. *Adding Shocks.* Now, let's consider what happens if we re-introduce the utility shock term. In this case, a household applies iff:

$$\tau l \alpha y - \delta \mu(y) b \leq \varepsilon. \tag{6}$$

Consider two levels of income, $y_1$ and $y_2 > y_1$, and assume that the cutoff value of $\epsilon$ in both cases is interior to the support of its distribution. The ratio of their show-up rates is:

$$\frac{1 - F(\tau l \alpha y_1 - \delta \mu(y_1) b)}{1 - F(\tau l \alpha y_2 - \delta \mu(y_2) b)}. \tag{7}$$

This ratio is always greater than one because the rich are less likely to sign up because their costs are higher and their probability of getting the benefit is lower. Note that this ratio is a measure of how well targeted the application process is towards poorer individuals – the higher the ratio, the higher the fraction of the poor in the population of applicants. Making the ordeal tougher reduces the number of poor applicants and imposes dead-cost on everyone who applies, which are both undesirable. Therefore, the only reason to do so is that it improves the ratio of poor to rich, which may reduce the costs of the program to the government.

Taking the derivative with respect to $l$, the distance to the registration site, tells us that targeting efficiency measured by this ratio improves when $l$ increases if and only if:

$$\frac{f(\tau l \alpha y_2 - \delta \mu(y_2) b)}{1 - F(\tau l \alpha y_2 - \delta \mu(y_2) b)} \tau \alpha y_2 - \frac{f(\tau l \alpha y_1 - \delta \mu(y_1) b)}{1 - F(\tau l \alpha y_1 - \delta \mu(y_1) b)} \tau \alpha y_1 > 0. \tag{8}$$

This formula says that when costs, $l$, are marginally increased by a small amount, the share of people who are lost is proportional to the density of people right on the margin – given by the PDF $f(y)$ – to the number of people who are inframarginal, given by the $1 - F(y)$ term.

This expression shows that a sufficient condition for targeting efficiency to be improving as $l$ increases is that the hazard rate,

$$\frac{f(\tau l \alpha y - \delta \mu(y) b)}{1 - F(\tau l \alpha y - \delta \mu(y) b)} \tag{9}$$

is weakly increasing with $y$, since if this is true then clearly $\frac{f(\tau l \alpha y - \delta \mu(y) b)}{1 - F(\tau l \alpha y - \delta \mu(y) b)} \tau \alpha y$ is increasing in $y$. This property holds if $F(\epsilon)$ represents a uniform, logistic, exponential or normal distribution, but not in other relevant cases such as the Pareto distribution and other "thick-tailed" distributions. The log-logistic distribution function $F(\epsilon) = \frac{\epsilon^\beta}{c^\beta + \epsilon^\beta}$ where $c$ and $\beta$ are two known positive parameters and $\epsilon \geq 0$, exhibits declining hazard rates as long as $\beta \leq 1$, but not otherwise.

To gain intuition into the model, we provide a simple illustration. In Figure 2, we examine the simplest case of no shocks and linear utility. In Panel A, we draw the relationship between income and gains for registration sites that are closer versus farther away. Note that the gain is decreasing more steeply with income for higher distance; this is the standard single-crossing property common to all screening models. As Figure B shows, moving from lower to higher distance reduces the number of applicants, but only among the rich. Thus, targeting efficiency improves.

Figure 3 shows an example of how introducing shocks can overturn the benchmark intuition developed in Section 3.2.1 above. We consider a simple case where income $y \in [0, 5]$, we set $\tau\alpha = 0.2$ and $\delta\mu\,(y)\,b = 0.5$, choose the log-logistic parameters $\beta = c = 0.5$, and consider distances $l \in \{2, 3\}$. As shown in Panel A, at any given consumption level, show-up rates are of course still higher at lower distances, and for any distance level, show-up rates decline in income. What is important to note, however, is that in this example, the initial rate of decline in show up rate with income (once the epsilons kick in) is quite high, but then slows as incomes become high. This is a consequence of the thick tails of the log-logistic distribution, which implies that $\frac{f(y)}{1-F(y)}$ is decreasing in $y$. This implies that increasing the distance from 2 to 3 actually hurts the ratio of poor to rich show-up rates, because it has a very large impact on the takeup at low income levels (where $\frac{f(y)}{1-F(y)}$ is large) but a much smaller impact at high income levels (where $\frac{f(y)}{1-F(y)}$ is small).

What this discussion illustrates is that single crossing in the classical screening sense is not sufficient for increasing ordeals to increase targeting effectiveness. Instead, one also needs to consider the density of people who are near the threshold and, hence, who will be affected by any marginal change in ordeals.

3.2.3. *Non-linearities in the Application Cost.* Let us continue to assume linear utility, but now model a non-linearity in the cost of applying, $c\,(l, y)$. This non-linearity may be more realistic because there are different transportation modes: one can either walk or take a bus. Buses are faster, but they cost money. Given that $l$ is the distance to the registration site, walkers face a calorie cost $\gamma l$ and a time cost $\tau l w$, where $w$ is their wage rate and $\tau l$ is defined to include the waiting time. Taking a bus requires a fixed bus fare, $\nu$, plus a time cost, $\lambda l w$, where $\lambda < \tau$. Once again, $\lambda l$ includes waiting time. Assuming that the wage is proportional to income, $w = \alpha y$, the decision rule is:

$$D = \begin{cases} \text{bus} & \text{if } \nu + \lambda l \alpha y < \gamma l + \tau l \alpha y \\ \text{walk} & \text{if } \nu + \lambda l \alpha y \geq \gamma l + \tau l \alpha y \end{cases}. \tag{10}$$

Applying is optimal if and only if:

$$-\,min\{\gamma l + \tau l \alpha y, \nu + \lambda l \alpha y\} + \delta\mu(y)b \geq ln\varepsilon. \tag{11}$$

The expression on the left hand side is declining in $y$. Therefore, richer people always apply less.

To look at the effect of increasing $l$, consider two income levels $y_1$ and $y_2$, such that at $y_1$, an individual just prefers to walk if he applies, and at $y_2$, he just prefers to take a bus, so that $y_1$ and $y_2$ are separated by some small distance $\psi$. For those with income $y_1$, the cost of travel is $\gamma l + \tau l \alpha y_1$. For those at $y_2$, it is $\nu + \lambda l \alpha y_2$. The fall in utility due to an increase in distance of $\Delta l$ will be greater at $y_1$ than $y_2$: $(\gamma + \tau\alpha y_1)\,\Delta l > (\lambda\alpha y_2)\,\Delta l$. Therefore, an increase in distance can increase travel costs more for the poor than for the rich.

To see this intuitively, consider the simple illustration in Figure 4. For both the rich and poor, taking the bus is initially more expensive (i.e., no bus fare), but has a lower marginal cost. Due to the higher marginal cost of their time, the rich switch to buses at lower distance than the poor ($l^*$). Between $l^*$ and $l^{**}$ (where the poor switch to the buses), one can clearly see from the figure that the marginal travel cost when $l$ is increased is actually *larger* for the poor than the rich. As a result, even in the case where $F(.)$ has increasing hazard rates, targeting efficiency may worsen. Note from the figure that this cannot happen if both people walk or both take the bus (i.e., travel costs are locally linear), or if the difference in incomes between them is large enough.

3.2.4. *Curvature in the Utility Function.* Finally, we introduce curvature into the utility function by letting $U(x) = lnx$. To focus on one mechanism, assume that there is no utility shock ($\epsilon \equiv 0$), that the cost of travel is linear in distance ($c(l,y) = \gamma l + \tau l \alpha y)$), and that $\mu(y)$ is a constant. In this case, the net gain from applying is:

$$\begin{aligned} g(y,l) &= \ln\left(y - c\left(l,y\right)\right) + \mu\delta \ln\left(y+b\right) + \left(1-\mu\right)\delta \ln y - \ln y - \delta \ln y && (12)\\ &= \ln \frac{\left(y - c\left(l,y\right)\right)\left(y+b\right)^{\mu\delta} y^{(1-\mu)\delta}}{y y^{\delta}}. && (13)\end{aligned}$$

The household will apply when:

$$\frac{\left(y - c\left(l,y\right)\right)\left(y+b\right)^{\mu\delta}}{y y^{\mu\delta}} \geq 1. \tag{14}$$

For convenience, we will work with the following function:[20]

$$G(y,l) = \frac{\left(y - c\left(l,y\right)\right)\left(y+b\right)^{\mu\delta}}{y y^{\mu\delta}}. \tag{15}$$

There exists a $y^{min}$ such that $y^{min} - c\left(l, y^{min}\right) = y^{min} - \gamma l - \tau l \alpha y^{min} = 0$. Let's start the discussion at this value of $y$ because any $y$ below this does not make sense in our model. At just above this level of $y$, $\frac{y-c(l,y)}{y}$ is close to zero and as a result $g$ must be less than one, so those with income levels in this range will not apply. As $y$ increases, $G$ also increases, since it starts at zero and thus can only go up). Taking the derivative of $G$ with respect to $y$ yields:

$$\begin{aligned} \frac{dG}{dy} &= \frac{\gamma l}{y^2}\left(1+\frac{b}{y}\right)^{\mu\delta} - \frac{\mu\delta b}{y^2}\left(1 - \tau l\alpha - \frac{\gamma l}{y}\right)\left(1+\frac{b}{y}\right)^{\mu\delta-1} && (16)\\ &= \frac{\left(1+\frac{b}{y}\right)^{\mu\delta-1}}{y^2}\left[\gamma l\left(1+\frac{b}{y}\right) - \mu\delta b\left(1-\tau l\alpha - \frac{\gamma l}{y}\right)\right]. && (17)\end{aligned}$$

In the neighborhood of $y = y^{min}$, the expression in the square brackets is strictly positive. However, the expression in the square bracket goes down when $y$ goes up and converges to $\gamma l + \tau l \alpha \mu \delta b - \mu \delta b$. If this expression is positive, then $G$ is monotone increasing in $y$, while if it is negative, then it first goes up and then goes down.

[20] So $g$, defined above, is $ln$G.

Figure 5 represents the two possible configurations of $G$ in this case. Panel A provides the case where G first increases and then falls, while Panel B represents the case where G is monotonically increasing. In each case, the values of $y$ for which the $G$ curve lies above the horizontal line at $G = 1$, are those that apply. The dashed line in each figure demonstrates what happens when $l$ goes up. In both cases the $G$ curve shifts down – in Figure 5a this means that both the poorest and richest people who were applying before the increase in $l$ drop out, while in Figure 5b only the poorest people drop out. In the first case, the effect on targeting depends on whether more of the poor proportionally drop out than the rich, which in turn depends on how the population is distributed near the two cutoffs. In the second case, the effect is unambiguously negative, with the fraction of the rich among applicants increasing when $l$ goes up.

It is worth noting that so far in this discussion we suppressed the effect of $y$ on $\mu(y)$, which goes in the direction of making the $G$ function downward sloping. In particular, if there exists a $y^{max}$ such that for $y \geq y^{max}, \mu(y) \approx 0$, as seems reasonable, then above $y^{max}$, $G < 1$ and no one will apply. The more realistic case is therefore probably the case in Figure 5a, and the effect of an increase in $l$ on targeting will depend on the shape of the income distribution.

3.3. **Summary.** This exercise illustrates the complexities in designing screening mechanisms: once we introduce a number of realistic features into the model, such as utility shocks that may have thick-tailed distributions, alternative means of transportation, and diminishing marginal utility, the intuitive argument that ordeals induce self-selection because the poor have a lower opportunity cost of time is no longer automatically true. Increasing the costs of the ordeal can worsen self-selection under relatively standard assumptions (log utility, as we saw above, for example is enough). Note that we have not yet even introduced the more behavioral arguments for why the poor may not be able to access the programs that are intended for them, such as self-control problems (e.g., Madrian and Shea, 2001), stigma (e.g., Moffitt, 1983), as well as informational arguments, such as the fact that the poor may not learn about the programs that are available to them (e.g., Daponte, Sanders and Taylor, 1999). On the other hand, the fact that $\mu(y)$ is downward sloping is a strong force in favor of self-selection leading the poor to self-select.

Given the theoretical ambiguity, whether self-targeting improves targeting efficiency is ultimately an empirical question. Therefore, we now turn to the data.

## 4. Who Self-Selects?

We begin by examining whether richer or poorer households were more likely to apply for the PKH program in the 200 villages where the government implemented the self-targeting treatment. Specifically, we plot a nonparametric Fan (1992) regression of the probability of applying against baseline log per capita consumption (Figure 6). Note, again, that the consumption data was collected before any mention of targeting occurred.[21] Bootstrapped standard error bands, clustered at the village level, are shown in dashes.

Across all expenditure ranges, Figure 6 shows that the poor are more likely to apply than the rich. This is evident as the probability of applying falls monotonically with per-capita consumption. At the very bottom of the expenditure distribution, a majority of households applies. For example, 61 percent of households at the 5th percentile of the consumption distribution do so. The share applying falls rapidly as consumption increases: at the middle of the expenditure distribution, only 39 percent percent of households apply, and by the 75th percentile, only 21 percent do so. At the 95th percentile of per-capita expenditure, only 10 percent of households apply.

From the perspective of the government, self-selection could affect targeting along two distinct dimensions. First, there could be selection on characteristics that are observable to the government: that is, households that have more assets, and are therefore less likely to pass the PMT, may be less likely to show up. This type of selection could potentially save the government resources since it would reduce the number of interviews that they would have to conduct for those who are likely to fail the PMT anyway, but it would not necessarily change the poverty profile of beneficiaries compared to automatic enrollment.[22] Second, there could be selection on the unobservable component of consumption: that is, conditional on a household's PMT score, households with higher unobservable consumption might also be less likely to attend. This could arise if there is self-selection based on the opportunity cost of time (as in the model), or if households do not perfectly understand the construction of the PMT score. If this type of selection on unobservables is occurring, then introducing self-selection has the potential to lead to a poorer distribution of beneficiaries than automatic enrollment.

To investigate this, we can decompose household consumption into the observable and unobservable components:

$$LNPCE_i = X_i'\beta + \epsilon_i \tag{18}$$

where $LNPCE_i$ is the household's log per-capita consumption, $X_i$ are the observable characteristics that enter the PMT formula, $\beta$ are the PMT weights, and $\epsilon_i$ is the residual, or the unobserved component of consumption. We then examine the relationship between the probability of applying and both the observable component, $X_i'\beta$ and the unobservable component, $\epsilon_i$.

We first examine these relationships graphically, presenting non-parametric Fan regressions of the probability of showing up as a function of the observable (Figure 7, Panel A) and unobservable (Panel B) components of log per-capita consumption. Bootstrapped standard 95 percent confidence

[21] Consumption may, of course, not be a perfect measure of welfare. First, there may be measurement error in consumption. Second, there may be alternative measures of welfare that may or may not more accurately represent a household's well-being (see Alatas, Banerjee, Hanna, Olken and Tobias (2012)). We use consumption because this is often the metric that governments are trying to actually target on. Note that these measurement errors will not affect our experimental results if the variation in consumption captures relative well-being; the measurement error will simply introduce noise into our estimate.

[22] In reality, it is often too costly to interview everyone in the country, so most governments do some form of selection to reduce the number of people interviewed. In our experimental results, we compare self-targeting to another methodology that the government uses to cull the number of interviews (the current status quo for Indonesia). We will then compare the efficiency of self-targeting to that of a hypothetical, full census PMT, to explore this dimension further.

intervals (clustered at the village level) are shown in dashes, and the vertical line in the top panel shows the average eligibility cutoff for receiving benefits. Strikingly, the probability of applying is decreasing in both the observable and unobservable components of consumption.

We now formally examine these relationships in a regression framework. Table 3 provides the results from estimating the following logit equation:

$$Prob\,(showup = 1) = \frac{exp\,\{\alpha + \gamma PMT_i + \gamma \epsilon_i\}}{1 + exp\,\{\alpha + \gamma PMT_i + \gamma \epsilon_i\}}, \tag{19}$$

where $PMT_i$ is the predicted portion of a household's log per-capita consumption (equal to $X_i'\beta$ from equation (18)), and $\epsilon_i$ is the residual portion of a household's log per capita consumption from equation (18). We use logit specifications since baseline showup rates will differ substantially once we start to examine different samples, and therefore, in these settings the logit model is easier to interpret. We show in Table A.2 that the results are qualitatively similar if we use linear probability models instead. Finally, note that all standard errors are clustered by village.

Table 3 confirms the graphical analysis and shows that there is self-selection along both margins, and that both of these forms of selection occur within both poor and richer households. Column (1) provides the coefficient estimates for the full sample. Both the observable and unobservable components of consumption significantly predict applying at the 1 percent level. The relative magnitudes suggest that the observed component of consumption has about 2.5 times the impact of the unobserved component, but both are large: a doubling of the PMT score (i.e., predicted log consumption based on assets) reduces the log-odds ratio of showing up by about 1.5; a doubling of the unobserved component of consumption reduces the log-odds ratio of showing up by about 0.6. In Columns (2) and (3), we split the sample based on whether the household would have been eligible to receive the program had they chosen to apply. What is notable is that selection on unobservables occurs in both samples. Thus, even among the poorest 4 percent of households in our sample, those who are poorer on unobservables are more likely to apply. This strong selection on unobservables suggests that self-selection has the potential to result in a dramatically poorer distribution of beneficiaries than other methods.

While both PMT scores and unobservables predict show-up rates, the R-squared is of course not 100 percent, so it is interesting to examine what other factors influence show-up decisions. In Appendix Table A.3, we add additional variables to equation (19). Panel A reports the results for the entire sample; Panel B reports the result for the subset of people who would be eligible. Several results are worth noting.[23] First, a household's subjective perceptions of its own wealth influence show up – i.e. those who perceive themselves to be poorer on a subjective scale of 1 to 6 are substantially more likely to show up. Second, those households who have received previous government programs (e.g., Raskin (rice for the poor), Askeskin (health insurance for the poor), and BLT (direct cash assistance for the poor)) are also more likely to show up. Both of these results suggest that households may be basing their show-up decisions in part on their perceived likelihood of receiving programs conditional on applying (i.e. their perceptions of $\mu(y)$), an issue we will return to in Section 7 below. Third, more educated households are less likely to apply, not

[23] Appendix Table A.3 is a logit specification, similar to Table 3; OLS results are shown in Appendix Table A.4.

more, suggesting that education is not a constraint on understanding program application rules in this context.

## 5. Comparing Self-Selection and Automatic Enrollment

The self-targeting treatment generated considerable self-selection, and yet only about 60 percent of the poorest group showed up, suggesting that there was significant exclusion error. However, it is not clear that we should be comparing self-targeting to the theoretical ideal of no error because, in reality, it is very costly for the government to collect consumption data for each and every household. Instead, the government's choice is often to conduct self-targeting or to conduct an alternative targeting methodology.[24] Therefore, in Section 5.1, we compare self-targeting against the real government procedure, which consists of an automatic enrollment for those who pass a proxy-means test among those selected to be interviewed by the government and local communities. Next, in Section 5.2, we additionally compare self-selection against a hypothetical exercise where we use the data that we have collected independently to predict selection if the automatic enrollment, based on the proxy-means test, was implemented universally.

5.1. **Experimental Comparison of Self-Targeting with Status Quo Targeting.** In this section, we test whether the types of individuals selected under self-targeting and automatic enrollment (the current status quo procedure of the Indonesian government) differ. To do so, we compare the distribution of beneficiaries in the 200 villages randomized to receive the self-targeting treatment with the 200 villages randomized to receive the automatic enrollment treatment. Given the randomization, the distribution of beneficiaries and the probability of receiving benefits should be identical in the two sets of villages absent the difference in targeting, so we can ascribe the differences that we observe between the two sets of villages to the differences in targeting methodologies (see Appendix Table A.1).

We begin with a graphical analysis in which we compare the distribution of beneficiaries under the self-targeting and automatic enrollment treatments (Figure 8). In Panel A, we plot the cumulative distribution function of log per-capita consumption of the final PKH beneficiaries in both sets of villages. The beneficiaries appear substantially poorer: the CDF of beneficiaries' consumption under automatic enrollment first-order-stochastically dominates that under selection. A Kolmogorov-Smirnov test of equality of distributions yields a p-value of 0.103.[25]

While the results in Panel A imply that the distribution of beneficiaries is poorer under self-selection, it does not tell the full story. In particular, it does not tell us whether this is due to the inclusion of more poor households, the exclusion of rich households, or some combination of both. To answer this question, we present non-parametric Fan regressions of the probability of obtaining benefits as a function of log per-capita consumption in Panel B of Figure 8. Bootstrapped 95 percent confidence intervals, clustered at the village level, are shown as dotted lines. The figure shows that the probability of receiving aid is substantially higher for the very poorest households

[24] Unlike asset data, which is verifiable in an in-person interview, consumption data is completely unverifiable since it is all self-reported. Even if the government could afford to do a consumption survey for all households, it could not use such data for targeting purposes since doing so would induce people to understate their true incomes.

[25] This p-value is based on randomization inference methods accounting for clustering at the village level. Alternatively, abstracting from the village-level clustering yields an exact p-value of 0.069.

in the self-targeting treatment. For those with log per capita consumption in the bottom 5 percent, i.e. those with log per capita consumption below about 12.33, the probability of receiving benefits is more than double that in self-targeting: 16 percent of those with log per capita consumption in the bottom 5 percent receive benefits as compared with just 7 percent in the automatic enrollment treatment. This difference is statistically significant at the 5 percent level. While exclusion error is still very high – even in self-targeting, only 16 percent of these very poor households received benefits, meaning that 84 percent were excluded – the rate of receiving benefits is 4 times higher than the overall rate of 4 percent of households in the sample who receive benefits, and double what it is in the status quo, automatic enrollment villages.

Conversely, households at higher consumption levels are substantially more likely to receive benefits in the automatic enrollment treatment. Households in the top 50 percent of the per-capita expenditure distribution – none of whom should be receiving benefits – are more than twice as likely to receive benefits in automatic enrollment than in the self-targeting treatment: 2.5 percent of such households receive benefits in automatic enrollment compared with 1 percent of such households in self-targeting (statistically significant at the 5 percent level). One explanation is that there are always errors in the PMT formula that allow some fraction of ineligible households to slip through the proxy-means test. With self-targeting, however, most of these households do not apply, so many fewer of them slip through. In sum, Figure B suggests that self-targeting both increased the probability that very poor households received benefits and decreased the probability that richer households did so, relative to the current status quo.

We now more formally quantify these effects using regression analysis, the results of which are presented in Table 4. In Column (1), we compare the difference in average log per capita consumption of the beneficiary populations ($LNPCE_{vi}$) in the two treatments, by estimating by OLS:

$$LNPCE_{vi} = \alpha + \beta SELF_v + \vartheta_{vi}, \tag{20}$$

where $SELF_v$ is a dummy for village $v$ being in the self-targeting treatment, and $\vartheta_{vi}$ is the error term. Standard errors are clustered by village. We estimate this model directly (Panel A) and with stratum fixed effects (Panel B). Note that this is the regression equivalent of comparing the means of the two distributions shown in Panel A of Figure 8. As suggested by the figures, the regression analysis confirms that beneficiaries are substantially poorer under self-selection: Column (1) of Panel A reports that per-capita consumption of beneficiaries is is 21 percent lower in self-targeting as compared to automatic enrollment (significant at the 1 percent level). Including stratum fixed effects (Panel B), the difference becomes 11 percent, and the p-value increases to 0.14.[26]

To increase our precision of the difference in consumption levels of beneficiaries, as discussed above, we did an interim midline survey after the targeting was complete, but before program beneficiary status had been announced or benefits had begun, in which we oversampled beneficiaries in both PMT and self-targeting villages. In Column (2), we compare log per-capita consumption of beneficiaries in the two treatments, including both the 159 beneficiaries from our baseline sample

[26]In general, one would expect stratum fixed effects to improve precision. However, in the regressions where we only consider beneficiaries, we have so few observations (159 observations), and hence so few observations per stratum, that including the fixed effects effectively drops many whole strata from the analysis, dramatically diminishing statistical power.

and the additional 745 beneficiaries we oversampled in the midline. Since the average level of consumption may be different in these two survey rounds (for example due to seasonality), we include a dummy variable for the survey round in which the data was collected. The results in Column 2 are similar in magnitude but more precisely estimated: self-targeting selects beneficiaries who are 18 to 19 percent poorer than those selected by the PMT treatment (statistically significant at the 1 percent level).

In Column 3 of Table 4, we examine the probability of getting benefits ($Prob\left(BENEFIT_{vi}=1\right)$) across the treatments for different groups. Specifically, we provide estimates from the following logit model:

$$Prob\left(BENEFIT_{vi}=1\right)=\frac{exp\left\{\alpha+\beta SELF_v+\gamma LNPCE_{vi}+\eta SELF_v\times LNPCE_{vi}\right\}}{1+exp\left\{\alpha+\beta SELF_v+\gamma LNPCE_{vi}+\eta SELF_v\times LNPCE_{vi}\right\}}. \quad (21)$$

The coefficient of interest is the coefficient $\eta$ on $SELF_v \times LNPCE_{vi}$, which captures the degree to which there is differential targeting in the self-targeting treatment as compared with automatic enrollment (the omitted category).[27] The results confirm the overall story shown in Panel B of Figure 8: the coefficient on $\eta$ is negative, large in magnitude, and statistically significant. This implies that there is much stronger targeting by consumption in the self-targeting treatment than in the automatic enrollment treatment. The magnitudes suggest that targeting is twice as strong in self-targeting: the estimates in Panel A imply that doubling consumption decreases the log-odds of receiving benefits by 0.70 in automatic enrollment, whereas it decreases the log-odds of receiving benefits by 1.37 in self-targeting.

In Columns (4) - (6), we examine alternative dependent variables to quantify the types of inclusion and exclusion error shown in Panel B of Figure 8. In Column (4) we define the overall error rate as a dummy that is equal to 1 if either exclusion error (failing to give benefits to a very poor household) or inclusion error (giving benefits to a non-very poor household) takes place. We find that the log-odds ratio of making an error is about 0.2 lower under self-targeting (p-values of 0.08 without stratum fixed effects and 0.11 with stratum fixed effects). Column (5) examines exclusion error, defined as a dummy for a very poor household failing to receive benefits. The results in the table suggest that the log-odds of such households being excluded (i.e., failing to get benefits) are between 0.55 and 0.71 lower in self-selection, though these results are not statistically significant (p-values of 0.18 and 0.15, respectively). Likewise, inclusion error, defined as a non-very poor household that does receive benefits, is lower in self-targeting, and statistically significant in the specification with stratum fixed effects (Column (6); p-values 0.14 and 0.08, respectively).

On net, the non-parametric and parametric results combine to paint a clear picture: self-targeting leads to a poorer distribution of beneficiaries, both because the poor are more likely to receive benefits and because richer households are less likely to receive benefits.

[27] We use logit models because the baseline benefit rate differs substantially by per-capita expenditure, so proportional models make more sense. Stratum fixed effects are also much more effective in proportional models given the substantially different poverty levels across strata. Appendix Table (A.5) shows that the OLS version of the same results are qualitatively similar, and if anything, show slightly higher levels of statistical significance. We cluster the standard errors in models with no fixed effects, and all OLS specifications, by village. For the conditional logit models where we include stratum fixed effects, for computational reasons we cluster fixed effects by stratum, which is more conservative (one stratum contains multiple villages).

5.2. **Comparing Self-Targeting to a Hypothetical Universal Automatic Enrollment Treatment.** In the automatic enrollment procedure, not all households were considered for enrollment. Instead, as discussed in Section 2.3.1, households only received the full PMT interview if they passed an initial set of screens. These pre-screening criteria were designed to save the government the cost of having to conduct a complete long-form census of every household in the country every time it wanted to select beneficiaries. On net, as shown in Table 2, about 34 percent of households in the village received the full PMT interview, which is roughly comparable to the share of households who self-select to be interviewed in the self-targeting treatment. Even though the government explicitly seeks to interview the poorest by targeting anyone who has ever taken part in a poverty-related program in the past supplemented by those that local officials think may be eligible, the fact that self-targeting improved inclusion error relative to the PMT treatment suggests that the government may not always be targeting the right people to interview. This could be the case, for example, if many of the very poorest rarely come in contact with government officials, so officialsdo not realize they are present and hence are not included on the survey list.

Comparing self-targeting against the current procedure is interesting because it provides information on the different methods that are realistically within a government's choice set. However, it is also interesting to ask how self-targeting performs relative to a PMT procedure that does not have the pre-screening that occurs in the actual procedure. While this is less realistic (i.e., it is too costly to actually be conducted by the government), it provides us with a greater understanding of the margins through which self-selection occurs. Thus, in this section, we assume, hypothetically, that the government had conducted the full PMT interview on everyone in the community. Recalling the decomposition of who selects to apply in the self-targeting treatment in Section 4 into selection on observables and selection on unobservables, we know *a priori* that self-targeting will perform worse than universal automatic enrollment with respect to selection on observables, because by definition universal automatic enrollment picks up 100 percent of households with PMT scores less than the cutoff whereas self-targeting limits the beneficiaries to a subset of those who chose to apply. However, it is still possible that self-selection could still out-perform universal automatic enrollment on net if the selection on unobservables is sufficiently large.

To simulate what would have happened in universal automatic enrollment, we use our baseline data to construct PMT scores for those households not interviewed by the government as part of the PMT process. That is, for those households who were not interviewed as part of the real PMT treatment, we assume that they would have received benefits if their PMT score (according to the asset data we collected in our baseline survey) was below the threshold required to receive the program. We then repeat the same analysis in Figure 8 and Table 4, but instead of comparing self-targeting to the actual automatic enrollment treatment, we compare it to the constructed hypothetical universal automatic enrollment procedure.

The results are shown graphically in Figure 9 and in regression form in Table 5. Panel A of Figure 9 shows that the distribution of beneficiaries still looks poorer in self-selection than in the hypothetical universal automatic enrollment, though the difference between the two distributions is no longer statistically significant (p-value from the Kolomogorov-Smirnov test of equality of distributions, with randomization inference to cluster at village level, is 0.29). Panel B of Figure 9 reveals that automatic enrollment and self-targeting have similar patterns in terms of the probability of being selected at the low end of the spectrum (though error bars cannot rule out some differences between them), but that wealthier households are more likely to receive benefits under the automatic enrollment than under self-targeting. This is related to selection on unobservables shown in Figure 8b – in the automatic enrollment treatment, some high consumption people make it through the PMT screen due to errors in the PMT, whereas those people do not self-select in the self-targeting treatment.

Looking at the regressions, Columns (1) and (2) of Table 5 confirm that, even under this hypothetical universal automatic enrollment treatment, the beneficiaries are poorer in self-targeting than in automatic enrollment (though statistical significance depends on specification.)[28]) Although noisy, exclusion error looks slightly higher in self-targeting (not surprising given that if data quality is the same, the hypothetical PMT should enroll a superset of those enrolled under self-targeting).[29] Inclusion error is substantially lower in self-selection. As a result, the overall error rate in targeting is substantially (and statistically significantly) lower in self-targeting than under this hypothetical universal automatic enrollment.

5.3. **Costs of Alternative Targeting Approaches.** Self-targeting appears to perform better in identifying the poor, but it also entails costs. There is the cost of the ordeal: households lose valuable time traveling to the interview site and waiting in line to be interviewed, and often need to spend money traveling as well. In addition, both self-targeting and PMT entail administrative costs – enumerators need to be paid to conduct interviews at self-targeting application sites for self-targeting and to conduct field verification visits to assess PMT scores in both self-targeting and PMT. One of the potential benefits of self-targeting is that it reduces the number of surveys that need to be conducted compared to a universal PMT; but if those cost savings to the government were offset by commensurate increases in the waiting and travel costs paid by households, one might not be so sanguine about such a policy.

To help shed light on this issue, Table 6 presents data on costs for the 200 villages in our sample in each treatment, along with the number of eligible households that do and do not receive benefits (exclusion error), the number of ineligible households that do and do not receive benefits (inclusion error), and, by way of comparison, the total annual dollars of benefits paid out to beneficiaries. We separate costs paid by households into those paid by households that end up receiving the benefits (for whom the net cost of applying or being interviewed was therefore positive) and for those paid by households that do not end up receiving the benefits (for whom the net cost of applying or being interviewed was negative). For PMT, where we surveyed only a single neighborhood, we extrapolate to the entire village linearly; likewise, we extrapolate the costs for the hypothetical universal PMT linearly from the actual PMT costs. Finally, note that there could be economies of scale in implementing a national program. For PMT, where we indeed know the Indonesian

[28]Note that we cannot replicate the analysis using the midline oversampling of beneficiary households here (e.g., the analogue of Column 2 of Table 4, since we did not oversample those households who would have been beneficiaries under the hypothetical universal PMT.

[29]Of course, data quality may not be the same: in self-targeting, only a small number of households likely to be selected is visited at home for the PMT interview, while in automatic enrollment, a much larger number is interviewed. It is possible that in the smaller, more focused self-targeting interviews, data quality is higher.

government's costs from implementing the nationwide PMT, we report those "at scale" costs as well as those from our experiment; for self-targeting, which has yet to be done nationally, we do not have an analogous estimate.

The results show that the costs on households imposed by self-targeting for 200 villages totaled around USD$70,000. The bulk of these costs (87 percent) were borne by non-beneficiaries, both because there were more of them and because, on average, they have a higher imputed wage rate. Administrative costs added an additional $170,000, so the total costs of targeting were around $240,000. These costs compare to around $1.2 million in benefits paid out in these villages per year. Since eligible households generally receive the program for 6 years, the total targeting costs for self-targeting are about 3 percent of the total benefits given out, and the costs actually borne by households from applying amount to about 1 percent of the total benefits given out.

The PMT treatment, which interviewed a similar number of households, imposed only US$9,366 in costs on households (just the time they spent at home taking the asset survey), and if we use the national-scale administrative costs, had a total cost of $120,378. But, as shown above, it had substantially higher rates of both inclusion and exclusion error compared to self-targeting. The hypothetical universal PMT, shown in Column (3), had almost identical exclusion error to self-targeting, though it had almost double the inclusion error. The total costs imposed on households would be about $32,000 (about 45% of PMT), but the administrative costs, even using the national-scale administrative costs, are about double that of self-targeting.

This analysis suggests that, if we treat administrative costs and costs borne by households equally, self-targeting dominates the hypothetical universal PMT, in that it achieves better targeting at lower total costs. Self-targeting and the status quo, automatic-enrollment PMT lie on very different parts of the frontier: the status quo costs as much as 40 percent less than self-targeting (though this difference could be muted if self-targeting enjoyed the same nationwide economies of scale as the status quo), but has substantially higher rates of both inclusion and exclusion error. The main additional difference is that self-targeting places a higher fraction of the burden directly on households, including many who do not ultimately receive benefits. Whether the benefits of increased targeting outweigh the costs therefore depends on how one weighs costs borne by households compared with administrative costs.

## 6. Marginal Effect of a Change in the Ordeal

Thus far, the findings suggest that self-targeting outperforms the status quo PMT procedure in identifying the poor. We next explore the optimal way to design ordeal mechanisms. We showed in Section 3 that the effect of marginally increasing the intensity of ordeals on separating the rich from the poor is theoretically ambiguous. Therefore, we first experimentally test the effect of a change in the ordeal on selection. Specifically, we examine the results from experimentally varying the distance to the registration site and the number of households members required to be present at the application site, as discussed in Section 2.3. Note that these experiments were carefully designed to be within the set of policy instruments that potentially could be considered by the government in their real conditional cash transfer program, under the requirements that the ordeals could not be so onerous that they would either discourage the severely credit-constrained poor from applying or that the program would unduly impose large application costs for the poor who might still be incorrectly screened out by the asset test.

We begin our discussion by exploring the effect of increasing distance. In the self-targeting villages, we experimentally chose whether the sign-up location would be situated very close or further away from the potential applicants' households. As Appendix Table A.9a shows, moving from the far to close registration sites decreased the distance from 1.88 km to 0.27 km; a reduction of 1.61 kilometers (or 1.69 kilometers controlling for strata fixed effects). [30] If the simplest version of the theory holds (See Section 3.2.1 under the assumption that the utility shocks are uniformly distributed), we expect that there should be more applicants in the close treatment and that they should be, on average, richer. Note, however, that under different model assumptions, the effect may be negative.

Table 7 explores the impact of the close treatment on targeting outcomes by estimating the following logit equation:

$$Prob\left(showup = 1\right) = \frac{exp\left\{\alpha + \beta CLOSE_v + \gamma LNPCE_{vi} + \eta CLOSE_v \times LNPCE_{vi}\right\}}{1 + exp\left\{\alpha + \beta CLOSE_v + \gamma LNPCE_{vi} + \eta CLOSE_v \times LNPCE_{vi}\right\}} \quad (22)$$

where $CLOSE_v$ is a dummy for the close treatment in village $v$, $LNPCE_{vi}$ is household $i$'s log per capita consumption, and $CLOSE_v \times LNPCE_{vi}$ is the interaction between them. Columns (1) - (3) show results without stratum fixed effects, and Columns (4) - (6) show results with stratum fixed effects.

Increasing distance reduces the number of applicants, but does not differentially affect who applies. We first show the results from estimating equation (22) including only the $CLOSE_v$ variable. The results show that the close treatment increases the log-odds of applying by between 0.21 (Column 1, no stratum fixed effects, p-value 0.16) and 0.28 (Column (4), with stratum fixed effects, p-value 0.101).[31] Put another way, this means that moving from far to close increases the percentage of households that applies by 15 percent (5.8 percentage points).[32] When we test for differential selection by consumption (Column (5)), we are unable to distinguish the effect of the close treatment by consumption levels from zero. Given that the theory implies that there may be non-linearities in the effect on the type of individual who applies when we alter the ordeal, we next explore potential non-linearities in the effect. Specifically, Column (6) interacts the close treatment dummy with dummies for quintiles of log per-capita consumption, and once again, we find no evidence that moving the targeting closer to the households differentially changes the distribution of who showed up.

[30] Given differences in geography, the treatment effect of distance varied across rural and urban locations. In rural areas, the sign-up station in the close treatment was located in each hamlet of the village (essentially 0 distance from people's houses), whereas in the far treatment it was in the village office (an average of 1.2 km from people's houses) (see Appendix Table A.9b). In urban areas, the sign-up station in the close treatment was located in the village office (an average of 0.8 km from people's houses), whereas in the far treatment it was in the subdistrict office (an average of 3.1 km from people's houses) (see Appendix Table A.9c).

[31] The OLS version of this coefficient, which is clustered at the village level rather than the stratum level, is statistically significant at the 5% level (p-value 0.024). See Appendix Table A.7.

[32] The fact that the marginal change in costs had any effect is in contrast to the one study we know of this form in the United States. In that study, Ebenstein and Stange (2010) use cross-state variation to examine the impact of a marginal change in ordeal, where those receiving unemployment insurance could re-certify their status by internet instead of in person. They find no effect on overall takeup from the change.

Similarly, as shown in Table 8, we also do not observe significantly fewer people applying when we require both spouses to apply in person rather than allowing either spouse to apply alone.[33] Given this, it is not surprising that we find no effect either on the interaction of $BOTH$ with per-capita consumption (Column (5)), or when we interact the treatment with quintile bins of consumption (Column (6)). One potential reason why requiring both spouses did not decrease enrollments is that this treatment included a provision through which households, in which one spouse was out of town and could not attend the interview, could get a signed letter from a neighborhood leader to this effect, allowing the interview to proceed with only one spouse. A total of 28 percent of interviewees came with such a letter, suggesting that this provision may have been used to allow those with high opportunity costs to register anyway. This suggests that ordeals may in fact be hard to enforce in practice – loopholes such as this one, which the government put in place to be fair to those who, for exogenous reasons, could not possibly comply with the ordeal, can be exploited to undo the intent of the ordeal. This phenomenon seems similar to related problems observed in providing incentives to nurses in India to show up at work – a loophole that was required to exempt those who could not attend because of a legitimate outside obligation from the incentive program was expanded so much that it undid the entire impact of the incentive program (Banerjee, Duflo and Glennerster, 2008).

## 7. Using the Model to Distinguish Theories and Predict Alternative Policies

The results thus far have shown that requiring households to apply for the program substantially improves targeting to the poor compared to automatic enrollment, yet marginal increases in application costs do not seem to further improve targeting. In this section, we return to the model in Section 3, estimate the unknown parameters of the model using the cross-sectional variation in the data, and use it to shed light on which theoretical mechanisms are driving the empirical results.

To take the model to the data, we start with equation (3), and specify a functional form for the utility function $U$ and shock term $\epsilon$. We assume that utility has a CRRA form ($U(x) = \frac{x^{1-\rho}}{1-\rho}$) with unknown curvature parameter $\rho$, and that the idiosyncratic utility shocks are drawn from a logistic distribution with mean $\alpha_\varepsilon$ and standard deviation $\beta_\varepsilon$. We focus on fitting these three parameters – $\rho$, $\alpha_\varepsilon$ and $\beta_\varepsilon$. [34]

To estimate the model, we exploit the cross-sectional variation in registration costs and benefits. We use data only from the far treatment group in fitting the model, so that we can explore what happens experimentally in the close treatment group as an out-of-sample validation of the model. We define registration costs as the per-capita monetary cost, including foregone wages, of traveling

[33] In fact, the estimates suggest that requiring both spouses to attend actually increases overall applications somewhat, perhaps because requiring both spouses means that the second spouse acts as a commitment device to show up, or perhaps because it is more fun to go together.

[34] We opt to not estimate a fourth parameter, $\delta$, because it turns out to have a lot of individual level heterogeneity which makes it hard to separate from the utility shocks. Choosing a reasonable value for $\delta$ is further complicated by the fact that PKH is supposed to last six years, but not everyone necessarily knows or believes that it will continue for that long. The discount factor therefore reflects that uncertainty as well as the usual impatience. For this reason, we take our baseline estimate of an annual discount factor to be 0.5, which is much lower than most conventional estimates, but show in the Appendix Table A.10 that the results are similar with other choices of $\delta$.

to the registration site, waiting in line, and returning home. That is, for household $i$, we specify:

$$c\left(y_i, l_i\right) = wage_i * \left(traveltime_i + \overline{waittime}\right) + travelmoney_i, \tag{23}$$

where $traveltime_i$ and $travelmoney_i$ are the individuals' reports of the time and expenditure required to reach the application site, which we observe in the baseline survey for all households, regardless of whether they show up or not. We compute $\overline{waittime}$ by taking average wait times by treatment group and urban/rural designation calculated from the endline survey.[35] We calculate the household hourly wage rate $wage_i$ by dividing monthly household expenditure by hours worked by the household in a month.

Figure **10** plots a Fan regression of the total costs of applying $c\left(y_i, l_i\right)$ against per-capita consumption $y_i$. The figure shows that the actual total sign up cost exhibits some mild concavity of the sort we introduced as a possibility in Section 3.2.3.[36]

We calculate the level of benefit, $b_i$, that the household would receive if enrolled in the program based on the number of children and their respective education levels.[37] We use a probit model to predict $\mu(y_i)$, the probability of getting the benefit conditional on applying.[38] Since consumption is likely measured with error, we assume that individuals make their decisions based on their true income $y^*$, whereas we observe $y = y^* e^{\omega}$, where $\omega$ is a normally distributed error term. We use the fact that, for a random subset of our sample, we observe per-capita consumption measured 3 months apart in the two endline surveys to calibrate the standard deviation of $\omega$, which we estimate to be 0.55, suggesting measurement error in consumption is non-trivial in our setting. We use the cross-sectional variation within the far treatment in $wage_i$, $traveltime_i$, $travelmoney_i$, $b_i$ and $\mu(y_i)$ to identify the model.

We estimate the model by Generalized Method of Moments, where the moments are the mean values of the show-up rates for the five quintiles of the consumption distribution in the far treatment. This gives us five moments to estimate three parameters, so we use a standard two-step GMM procedure to compute optimal weights among the five moments. For each quintile in the far treatment, we thus match the empirical show-up rate by integrating over possible unobserved values of the utility shock $\epsilon$ and measurement error $\omega$ term as follows:

$$Prob(showup = 1) = \iint \mathbf{1}\left\{g(y_i e^{\omega}, l_i) + \epsilon > 0\right\} df_{\epsilon} df_{\omega},$$

where $g\left(y, l\right)$ is defined in equation (4) and where $U(x) = \frac{x^{1-\rho}}{1-\rho}$.

[35] We do not have sufficient data to calculate separate wait times for each village.

[36] A regression of $c\left(y_i, l_i\right)$ on $y_i$ and $y_i^2$ shows that the coefficient on the quadratic term is statistically significant at the 5 percent level. This is not driven by the outliers shown in the figure; we obtain a similar result even when we drop the 17 observations with per-capita consumption above Rp. 2,000,000 per month.

[37] The benefit is calculated as follows. Each beneficiary household receives a base benefit of Rp. 200,000 per year. This level increases by Rp. 800,000 if they have a child age less than 3 or are currently expecting, by Rp. 400,000 if they have a child enrolled in primary school, and by Rp. 800,000 if they have a child in middle school. Since all beneficiaries fall into at least one of these categories, the benefit level is therefore between Rp. 600,000 and Rp. 2.2 million per year, with a mean of about Rp. 1.3 million.

[38] We model the probability of receiving the benefit, conditional on applying, as a function of Log PCE. We include urban/rural interacted with district fixed effects, since the PMT cutoff for inclusion varies slightly for each urban/rural times district cell. The results are shown in Figure 1.

Table 9 shows the estimated parameter values. Specifically, the three estimated model parameters are $\alpha_\varepsilon = -26,126$, $\beta_\varepsilon = 26,805$, and $\rho = 0.0000$.[39] The result that $\alpha_\varepsilon < 0$ implies that the idiosyncratic utility shocks on average favor not showing up. Since utility is estimated to be linear, $\alpha_\epsilon$ is interpretable in dollar terms, so the mean $\epsilon$ term is equal to about USD\$2.50. The fact that $\rho = 0$, which implies that the households are expected income maximizers with linear utility, is somewhat surprising: perhaps it reflects the fact that on a monthly basis both the realized gains and the actual costs are relatively small numbers (per capita monthly benefit is on average 5.22 percent of monthly per capita expenditure for the entire sample, while total cost per capita is 0.72 percent of monthly per-capita expenditure for the entire sample). Given the estimated linearity of the local utility function, it is not surprising that we got a clearly downward-sloping show-up curve when we graphed show up rates against per capita consumption in Figure 6, as the potential effect of the poor having much higher marginal utility costs of signing up, as discussed in Section 3.2.4, does not appear to play a role empirically.

We then use these estimated parameters to predict the application rates under different assumptions for the cost function $c(y,l)$. For each possible $c(y,l)$, we simulate predicted application rates. To summarize what the model predicts, we repeat the same logit regressions we performed in Table 7 on the simulated data. We also calculate the predicted show-up rates for close and far sub-treatments for those above and below the poverty line.[40]

The results from this exercise are shown in Table 10, and the predicted show-up rates by quintile are graphed in Figure 11. For comparison purposes, Column (1) of Table 10 and the top-left graph of Figure 11 replicate the actual empirical results (e.g., Column (2) of Table 7). In addition to the empirical results from the logit model, in Panel B, we calculate the show-up rates for those above and below the poverty line for both near and far treatments. In Panel C, we calculate the ratio of the poor to rich show-up rates (i.e., equation (7) from the model) for both treatments, as well as the difference in this ratio between the near and far treatments (i.e., equation (8) from the model). The ratio is positive but statistically insignificant, indicating no statistically detectable differential targeting induced by moving from near to far in the experiment.[41]

In Column 2 of Table 10, we begin by estimating the effect on the simulated data of the change in $c(y,l)$ induced by the close treatment; that is, we use the actual costs $c(y_i, l_i)$ for both close and far households calculated using equation (23), and calculate each household's predicted show-up rate using the model. Since we only used the far treatment to estimate the model, comparing these

[39] Note that the estimation was constrained such that $\rho \geq 0$.

[40] In order to run the logits using the predicted application rates, we create 3,000 copies of the data. The copies of each individual are assigned to apply or not apply in proportion to that individual's predicted probability of doing so. To make the standard errors comparable to the main experiment, we apply a cluster bootstrap approach (clustered on villages) to this distribution, holding the total number of observations equal to the number of observations in the actual data.

[41] Note that the ratio is positive but insignificant, whereas the interaction term (the estimated coefficient on $[Close * LogPCE]$) in Panel A is negative and insignificant. The reason they are of different signs is that the logit model in Panel A is estimated using the linear $LogPCE$ variable, whereas the ratios in Panel C are based on a dummy variable for poor / non-poor. If we re-estimate the logit model using a dummy variable for rich, we obtain results with the same sign. Note also that the results in this table are based on the actual populations in the near and far sub-groups. Since this was randomized, these will be statistically similar, but there may be small sample differences. Appendix Table A.12 replicates the analysis in this table adjusting for these small sample differences.

simulated show-up rates to actual show-up rates serves as an out-of-sample check of the fit of the model using the experiment. We bootstrap the standard errors using sample sizes equivalent to our actual data and with village-level clustering, so that the standard errors reported for the model-generated data are equivalent to those from the actual data. The results in column 2 thus show what we would have found had the data from our actual survey been generated by the model.[42]

Comparing the actual empirical estimates in Column 1 with the estimates on the model-generated data in Column 2, we find similar results of differential targeting between the treatments. In particular, even though the model seems to over-predict show up rates in the close treatment on average, the small differential effect between rich and poor show up ratios moving from near to far in the simulated data is not statistically distinguishable from what we actually observe in the experiment (Panel C; p-value 0.602). Consistent with this, the coefficients on the close dummy interacted with log per capita consumption ($\eta$ in equation (22)), which is another way of capturing the degree of differential targeting between the close and far treatment, are also statistically indistinguishable between the actual experimental data in Column (1) and the simulated data in Column (2) (p-value 0.441). The fact that the model predictions are similar to the experimental findings provides us with greater confidence in the simulation results for alternative cost structures in the following columns.[43] A comparison of model fit can be seen by comparing the actual show up rates by quintile and treatment in the top-left of Figure 11 with the model's predicted show up rates by quintile and treatment in the top-middle of Figure 11.

7.1. **Distinguishing Alternate Theories.** Interestingly, even though there is strong evidence of self-selection (the poor are much more likely to show up than the rich, both on observables and unobservables), both the experiment and the model show no statistically significant marginal increase in the targeting ratio from increasing the severity of the ordeal (i.e., moving from near treatment to far). We can use the model to help understand why this is not occurring, and in particular, examine the various mechanisms outlined in the model in Section 3.

*Shocks.* One possible explanation developed in the theory section is that, if the distribution of shocks does not have the monotone hazard rate property, it is possible that targeting could get worse as you increase distance, because the density of poor people induced to drop out by a higher marginal change is higher than the density of richer people (see Section 3.2.2). However, the version of the structural model we estimate and use in Column (2) uses logit shocks, which do have the

[42]Over the last few years there have been several papers in the development literature that similarly use a well-identified randomized or natural experiment to provide a check of model fit. These include Todd and Wolpin (2006), Kaboski and Townsend (2011) and Duflo, Hanna and Ryan (2012). More generally, the idea of hold out samples for validation has been used in several papers in the broader applied micro literature, staring at least with McFadden (1977); see for example, Wise (1985), Lise, Seitz and Smith (2004), Keane and Mott (1998), Keane, Todd and Wolpin (2011). Other papers which combine structural methods and experimental data (without using one of the experimental groups for out of sample validation) include Attanasio, Meghir and Santiago (2012), Einav et al. (2013) and Ferrall (2010).

[43]The one aspect of the model that does not match is that the predicted show-up rates for those below the poverty line are actually higher in the far treatment than in the near treatment (69 percent vs 67 percent). We have verified that this is not due to the model, but rather due to small-sample differences in the expected benefits from obtaining the program among the poor in these two samples. In particular, the poor in the far group have (statistically insignificantly) more middle schoolers than the poor in the near group, which leads to higher show-up rates. If we simulate the impact of moving from far to close on the exact same group of beneficiaries, we indeed would obtain lower show-up rates in far than in close in both rich and poor samples. See Appendix Table A.12.

monotone hazard rate property, yet still replicates the experimental findings. This suggests that the distribution of shocks alone is not the problem.

However, the magnitude of the shocks may explain why the response is so low. Examining equation (8), which showed the derivative of the show-up ratio with respect to a change in distance $l$, one can see that increasing the variance of the shocks, which would lower the PDF $f$ at the margin for both rich and poor, would dampen the responsiveness to a marginal increase in ordeals. To assess quantitatively whether this is important, in Column (3) we re-simulate the model where we cut the standard deviation of the shocks $\epsilon$ in half. Doing so increases the point estimate of the impact of moving from close to far on the poor/rich show-up ratio – from 0.314 in the base-case model to 0.470 – but it would still not have been enough to be statistically detectable. In Column (4) we shut off the shocks entirely, so that everyone for whom $g(y, l) > 0$ shows up. This increases the estimated impact on the show-up ratio to 0.584, but again, it would not have been enough to be statistically detectable.

*Curvature in the Utility Function.* Another possible explanation given by the theory is that there may be curvature in the utility function, so that even though the marginal monetary cost of higher distance is greater for the rich, the monetary utility cost is greater for the poor (see Section 3.2.4). However, when we estimated the structural model, the found that the model was best fit with linear utility (i.e. $\rho = 0$), suggesting that this is not an important part of the explanation in our setting.[44]

*Different Travel Technologies s.* The third explanation suggested by the model is that there are different transportation technologies used by the poor and the rich, so that the marginal monetary cost of distance is smaller for the rich (see Section 3.2.3). Figure 10 showed that this might be a possible explanation in the data, as the total costs of travel do appear to be concave in per-capita consumption. To investigate whether this explains the lack of differential selection in response to an increase in distance, we use the model to generate simulated show up rates under the counterfactual that the poor and the rich use the same travel technology. To do so, we model travel costs (time and money) as a function of distance. Treating urban and rural populations separately, we regress reported monetary costs and reported travel time to the close and far registration places on quadratic functions of distance. We then use these predicted average travel costs – which by construction no longer allow richer households to use different transportation technologies – for all households, and re-calculate total registration costs $c(y, l)$. We then re-estimate the logit regressions and calculate the show-up rates for the simulated data using these costs instead of the actual costs. Column (5) reports the results, which appear similar to the experimental findings (p-value 0.449 in Panel A; p-value of 0.624 in Panel C). The fact that the results are virtually unchanged when everyone is constrained to use similar transportation technologies suggests that the lack of differential selection between close and far is not being driven by the fact that the rich and poor use different transport technologies. The predicted show-up rates using the same transport technology are shown in the top right of Figure 11 and confirm that technology is not the main issue.

*Probability of Receiving Benefits.* A final explanation is that most of the selection we observe in Section 4 is being driven by the fact that households anticipate that $\mu(y)$, the probability of receiving

[44] Appendix Figure A.1 shows the actual model fit, and alternatives where we impose higher values of $\rho$. As is evident from the Figure, imposing higher values of $\rho$ leads to a more convex relationship between show-up rates and income quintile than we observe in the data.

benefits conditional on showing up, is downward sloping in income.[45] To gauge the magnitude of that effect, in Column (6), we simulate what would happen if, instead of using the actual empirical $\mu(y)$ function, we assume that all households assume that they will receive benefits with some constant probability $\bar{\mu}$ equal to the population average probability of getting benefits. The results are dramatic – the coefficient on log-per-capita expenditure falls from around -1.4 (in Columns (1) and (2)) to -0.3 (in Column (5)). This suggests that about 20 percent of the selection effect is driven by the differential costs paid by rich and poor, and about 80 percent of the selection is caused by the fact that the poor and rich have differential beliefs about their probability of receiving benefits conditional on applying. Comparing the change in poor to rich show up ratios when we move from the base model to the model with constant $\bar{\mu}$, the share of the selection caused by $\mu$ as opposed to differential costs is even higher. This result is consistent with the overall empirical findings of the paper: if most of the selection is coming because $\mu(y)$ declines rapidly with income rather than $c(l, y)$ increasing rapidly with income, then even small costs can have very large selection effects, since people with very low $\mu(y)$ will not bother to sign up, but marginal increases in the costs of the ordeal $l$ impose deadweight costs without substantially improving selection.

7.2. **Simulating Alternate Policies.** The results thus far suggest that perhaps the problem is largely one of magnitudes – one might need a very large change in ordeals to impose meaningful additional self-selection. The remaining columns consider counterfactual experiments where, for the far group, we increase either the distance to the application site or the average wait time, to see just how much of an ordeal one might need for the selection to become substantial. To simulate these counterfactual costs with increased distance, we again regress travel time and monetary costs on quadratic functions of distance from the application site, but now we do it separately for each rural/urban and income quintile bin, to allow costs to be heterogeneous by income group. We then calculate the additional costs of increased distance by adding either 3 km or 6 km to the actual distance, using the estimated relationships to calculated marginal time and money costs from that additional distance, and then adding that amount to the actual time and money costs reported for each individual. To simulate counterfactual costs with increased waiting time, we simply increase the average waiting times by 3 or 6 times.

The results shown in Columns (7) and (8) of Table 10 and graphed in the second row of Figure 11, demonstrate that adding additional distance is still not enough to induce substantial differential selection – even adding 6 km of distance, almost 4 times the mean value of 1.67 km – is not enough to induce substantial additional selection. The reason is that the marginal costs of increased distance do not appear to be that high because the costs of distance are concave – given that at such far distances almost everyone (even the poor) takes some form of motorized transportation, adding 6 km of distance raises the costs of applying by only about Rp. 6,700 on average (US$.70) (see Appendix Table A.11).

The results in Columns (9) and (10), and graphed in the third row of Figure 11, show that, by contrast, dramatically increasing wait times in the far treatment could induce detectable differential

[45]Alternatively, it could be that there is a stigma from applying that is increasing with income $y$; i.e. the rich would feel embarrassed by showing up and applying for an anti-poverty program, and the poor would not. Empirically, this will look identical to a downward sloping $\mu(y)$ function.

selection. For example, when we increase wait times by a factor of 6 for the far treatment, we estimate a ratio of 2.8-1 for the poor-to-rich show up rates. This compares to a predicted ratio of 2.2-1 for the baseline model in Column (2). What is happening is that the non-poor are dissuaded from showing up – 33 percent of non-poor show up in the baseline model, compared to only 23 percent when the wait times are increased by a factor of 6, a decline of about 30 percent. By contrast, the show up rates for the poor decrease by only about 10 percent when the wait times are increased by a factor of 6. Intuitively, wait times are more effective than distance in generating selection because wait times are a pure time cost, so the monetary costs are much more differential by income, while poor and rich, after a certain distance, use motorized transportation technologies so that the marginal cost of additional distance is relatively low for both income groups.

However, it is important to note that there are problems with long wait times in practice – the estimated wait times we needed to assume in Column (10) averaged over 17 hours – almost two full work days of waiting in line. The wait times in Column (9), where we increase them by a factor of 3, are still about 9 hours. In a pilot for this study, when we experienced long wait times (although still much less than 17 hours), villagers spontaneously organized themselves and assigned queuing numbers, so that people could wait at home and come back when it was their turn to be interviewed, rather than having to spend hours waiting in line. This suggests that while theoretically long wait times could be an effective screening device, actually making applicants wait for more than a full day may be very difficult in practice.

## 8. Conclusion

Using data from a field experiment across 400 villages to examine targeting in Indonesia's conditional cash program (PKH), we showed introducing application costs meant that the poor are more likely to self-select into applying than the non-poor. Interestingly, this selection occurred on two types of margins. First, we observe selection on the component of consumption that is observable to governments. This implies ordeals have the potential to save money by not having to survey rich people who would ultimately fail the asset test. Second, ordeal mechanisms also lead to selection on the unobservable components of consumption, which means that targeting may become more pro-poor by screening out the rich who may get incorrectly screened in by an asset test. On net, introducing self-selection improved targeting as compared with the other targeting mechanisms that we considered, both the current status quo and a universal automatic enrollment system.

However, while experimentally increasing the ordeals by increasing the distance to the application site reduced the number of individuals who applied under the self-targeting regime, it did not differentially improve targeting. Put another way, the increase in distance we experimentally induced (a 1.6 kilometer increase in distance) imposed substantial enough costs on households to lower application rates, but these costs did not differentially impact poor and rich households. Estimating our model suggested that the key driver behind the improvement in targeting from application costs was the fact that the rich forecast they have a low probability of success and hence do not choose to apply.

In short, these types of administrative costs can be a powerful tool to improve targeting relative to automatic enrollment systems, but making onerous ordeals even more costly may not be the best way to improve targeting further. This suggests that one should not strictly view administrative barriers as a bar to take-up, but instead should carefully consider their power as a screening device. On the other hand, while ordeals dominate the status quo, many of the poor still do not sign up. Understanding how to design screening mechanisms to increase takeup of the poor while still discouraging sign up of the rich seems a promising direction for future work.

## References


Alatas, V., A. Banerjee, R. Hanna, B.A. Olken and J. Tobias. 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review* 104 (2):1206–1240.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari and Matthew Wai-poi. 2012. Elite Capture or Elite Benevolence? Local Elites and Targeted Welfare Programs in Indonesia. Technical report MIT.

Attanasio, O.P., C. Meghir and A. Santiago. 2012. "Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa." *The Review of Economic Studies* 79(1):37–66.

Banerjee, A.V., E. Duflo and R. Glennerster. 2008. "Putting a Band-Aid on a corpse: Incentives for nurses in the Indian public health care system." *Journal of the European Economic Association* 6(2-3):487–500.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas and Victor Pouliquen. 2013. Turning a Shove into a Nudge? A "Labeled Cash Transfer" for Education. Technical Report 19227 NBER.

Bertrand, Marianne, Sendhil Mullainathan and Eldar Shafir. 2004. "A behavioral-economics view of poverty." *The American Economic Review* 94(2):419–423.

Besley, Timothy and Stephen Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *American Economic Review* 82(1):249–61.

Castaneda, Tarsicio and Karthy Lindert. 2005. Designing and Implementing Household Targeting Systems: Lessons from Latin American and The United States. Social Protection Discussion Paper 526 World Bank.

Coady, D. and S. Parker. 2009. Targeting social transfers to the poor in mexico. Working Paper 9/60 International Monetary Fund.

Currie, J. 2006. Public Policy and the Income Distribution. In *Public Policy and the Income Distribution*, ed. David E. Quigley John M. Auerbach, Alan J. Card. Russel Sage Foundation.

Daponte, B.O., S. Sanders and L. Taylor. 1999. "Why do low-income households not use food stamps? Evidence from an experiment." *Journal of Human Resources* pp. 612–628.

Duflo, E., R Hanna and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4):1241–78.

Ebenstein, Avraham and Kevin Stange. 2010. "Does inconvenience explain low take-up? Evidence from unemployment insurance." *Journal of Policy Analysis and Management* 29(1):111–136.

Fan, Jianqing. 1992. "Design-adaptive Nonparametric Regression." *Journal of the American Statistical Association* 87(420):998–1004.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P Newhouse, Heidi Allen, Katherine Baicker et al. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year*." *The Quarterly Journal of Economics* 127(3):1057–1106.

Grosh, Margaret, Carlo del Ninno, Emil Tesliuc and Azedine Ouerghi. 2008. *For Protection & Promotion: The Desgn and Implementation of Effective Safety Nets*. World Bank.

Hodges, Anthony, Anne-Claire Dufar, Khurelmaa Dashdorj, Kang Yun Jong, Tuya Mungan and Uranchimeg Budragchaa. 2007. Child benefits and poverty reduction: Evidence from Mongolia's

Child Money Programme. Technical report Maastricht University.

Kaboski, J.P. and R.M. Townsend. 2011. "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative." *Econometrica* 79(5):1357–1406.

Keane, M.P., P.E. Todd and K.I. Wolpin. 2011. "The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications." *Handbook of Labor Economics* 4:331–461.

Kidd, Stephen and Emily Wylde. 2011. Targeting the Poorest: An assessment of the proxy means test methodology. Technical report AusAID.

Lise, J., S. Seitz and J. Smith. 2004. Equilibrium policy experiments and the evaluation of social programs. Technical report National Bureau of Economic Research.

Madrian, BC and DF Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116(4):1149–1187.

Martinelli, Cesar and Susan W. Parker. 2009. "Deception and Misreporting in a Social Program." *Journal of the European Economic Association* 7(4):886–908.

Moffitt, R. 1983. "An economic model of welfare stigma." *The American Economic Review* 73(5):1023–1035.

Nichols, Albert L. and Richard J. Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *The American Economic Review* 72(2):372–377.

Nichols, D., E. Smolensky and T.N. Tideman. 1971. "Discrimination by waiting time in merit goods." *The American Economic Review* 61(3):312–323.

Ravallion, M. 1991. "Reaching the Rural Poor Through Public Employment: Arguments, Evidence, and Lessons from South Asia." *World Bank Research Observer* 6(2):153–175.

Thornton, R.L., Hatt L.E. Field E.M. Mursaleena I. Diaz F.S. Gonzalez M.A. 2010. "Social Security Health Insurance for the Informal Sector in Nicaragua: A Randomized Evaluation." *Health Economics* 19:181–206.

Todd, P.E. and K.I. Wolpin. 2006. "Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility." *The American Economic Review* 96(5):1384–1417.

Wise, D.A. 1985. "A behavioral model versus experimentation: The effects of housing subsidies on rent." *Methods of Operations Research* 50:441–89.

TABLE 1. Experimental Design

| | | Both Spouse Subtreatment | Either Spouse Subtreatment | *Total* |
|---|---|---|---|---|
| **Automatic Enrollment** | | | | 200 (1,998) |
| **Self Targeting** | Close Subtreatment | 50 (500) | 50 (500) | 100 (1,000) |
| | Far Subtreatment | 50 (500) | 50 (500) | 100 (1,000) |
| | *Total* | 100 (1,000) | 100 (1,000) | 200 (2,000) |

Notes: This table provides the number of villages in each treatment cell. The number of households in each cell is also shown in parentheses.

TABLE 2. Descriptive Statistics for Households Surveyed in the Baseline

| | Total number of households (1) | Number of households interviewed (2) | Number of beneficiaries (3) | Percentage of households interviewed (4) | Percentage of interviewed households that received benefits (5) | Percentage of total households that received benefits (6) |
|---|---|---|---|---|---|---|
| Automatic Enrollment | 1998 | 706 | 86 | 35.34% | 12.18% | 4.30% |
| Self Targeting | 2000 | 754 | 73 | 37.70% | 9.68% | 3.65% |

Notes: This table provides information on the flow of surveyed households through the experiment.

TABLE 3. Probability of Showing Up as a Function of the Observed and Unobserved Components of Baseline Log Per-capita Consumption

| | Showed up | | |
|---|---|---|---|
| | All | Very poor | Not very poor |
| | (1) | (2) | (3) |
| Observable consumption ($X_i'\beta$) | -2.217*** | -0.811 | -2.283*** |
| | (0.201) | (1.981) | (0.204) |
| Unobservable consumption ($\varepsilon_i$) | -0.907*** | -1.702* | -0.878*** |
| | (0.136) | (0.877) | (0.137) |
| Stratum fixed effects | No | No | No |
| Observations | 2,000 | 72 | 1,928 |
| Mean of dependent variable | 0.377 | 0.653 | 0.367 |

Notes: Each column shows a logit regression of show up rates on PMT score and epsilon. Very poor is defined as being eligible for the program based on PMT score. Robust standard errors, clustered at the village level, shown in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4. Experimental Comparison of Targeting under Self Targeting and Automatic Enrollment Treatments

| | Log consumption beneficiaries (baseline) (OLS) (1) | Log consumption beneficiaries (baseline + midline) (OLS) (2) | Receives benefits (LOGIT) (3) | Error (LOGIT) (4) | Exclusion error (LOGIT) (5) | Inclusion error (LOGIT) (6) |
|---|---|---|---|---|---|---|
| *Panel A: No Stratum Fixed Effects* | | | | | | |
| Self targeting | -0.208*** | -0.193*** | 12.142** | -0.219* | -0.547 | -0.313 |
| | (0.076) | (0.060) | (4.894) | (0.127) | (0.403) | (0.210) |
| Log consumption | | | -1.016*** | | | |
| | | | (0.280) | | | |
| Log consumption * Self targeting | | | -0.964** | | | |
| | | | (0.383) | | | |
| Observations | 159 | 904 | 3,996 | 3,998 | 243 | 3,755 |
| Mean of dependent variable | 12.78 | 13.61 | 0.0398 | 0.0855 | 0.877 | 0.0344 |
| *Panel B: With Stratum Fixed Effects* | | | | | | |
| Self targeting | -0.114 | -0.175*** | 15.180*** | -0.239 | -0.709 | -0.334* |
| | (0.077) | (0.058) | (5.295) | (0.148) | (0.492) | (0.193) |
| Log consumption | | | -1.042*** | | | |
| | | | (0.283) | | | |
| Log consumption * Self targeting | | | -1.202*** | | | |
| | | | (0.416) | | | |
| Observations | 159 | 904 | 3,489 | 3,918 | 110 | 3,134 |
| Mean of dependent variable | 12.78 | 13.61 | 0.0456 | 0.0873 | 0.755 | 0.0412 |

Notes: Exclusion error is defined to be 1 if a household is very poor (as measured at baseline) and does not receive PKH and 0 otherwise. Inclusion error is defined to be 1 if a not-very poor household does receive PKH and 0 otherwise. Error includes either exclusion or targeting error. In Panel A, robust standard errors, clustered at the village level, are shown in parentheses. In Panel B, Columns (2) - (5), robust standard errors are clustered at the stratum level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5. Comparison of Targeting under Self-Selection and Hypothetical Universal Automatic Enrollment

| | Log consumption (beneficiaries) (OLS) (1) | Receives benefits (LOGIT) (2) | Error (LOGIT) (3) | Exclusion error (LOGIT) (4) | Inclusion error (LOGIT) (5) |
|---|---|---|---|---|---|
| | | *Panel A: No Stratum Fixed Effects* | | | |
| Self targeting | -0.133* | 6.545 | -0.271** | 0.095 | -0.541*** |
| | (0.069) | (4.710) | (0.129) | (0.350) | (0.207) |
| Log consumption | | -1.428*** | | | |
| | | (0.261) | | | |
| Log consumption * Self targeting | | -0.552 | | | |
| | | (0.369) | | | |
| Observations | 186 | 3,996 | 3,998 | 243 | 3,755 |
| Mean of dependent variable | 12.75 | 0.0465 | 0.0878 | 0.840 | 0.0391 |
| | | *Panel B: With Stratum Fixed Effects* | | | |
| Self targeting | -0.040 | 9.055* | -0.293* | 0.128 | -0.571*** |
| | (0.064) | (4.981) | (0.156) | (0.322) | (0.207) |
| Log consumption | | -1.488*** | | | |
| | | (0.271) | | | |
| Log consumption * Self targeting | | -0.749* | | | |
| | | (0.393) | | | |
| Observations | 186 | 3,489 | 3,918 | 126 | 3,134 |
| Mean of dependent variable | 12.75 | 0.0533 | 0.0896 | 0.714 | 0.0469 |

Notes: Exclusion error is defined to be 1 if a household is very poor (as measured at baseline) and does not receive PKH. Inclusion error is defined to be 1 if a not-very poor household does receive PKH. Error includes either exclusion or targeting error. Households are defined as beneficiaries of the hypothetical PMT if either their PMT score defined at baseline qualifies them for PKH or they in reality received the benefit. In Panel A, robust standard errors, clustered at the village level, are shown in parentheses. In Panel B, Columns (2) - (5), robust standard errors are clustered at the stratum level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

TABLE 6. Summary of Targeting and Costs

| | Self-Targeting (1) | PMT (2) | Hypothetical Universal PMT (3) |
|---|---|---|---|
| # of eligible households that receive benefit | 2167 | 1341 | 2347 |
| # of eligible households that do not receive benefit | 11917 | 12743 | 11737 |
| # of ineligible households that receive benefit | 6621 | 8960 | 11140 |
| # of ineligible households that do not receive benefit | 220051 | 217711 | 215532 |
| Total annual benefits paid ($) | 1198099 | 1404528 | 1838845 |
| Total cost to households ($) | 108145 | 9366 | 32403 |
| *Total cost to beneficiary households ($)* | 13400 | 1174 | 1407 |
| *Total cost to non-beneficiary households ($)* | 94618 | 8192 | 31002 |
| Total administrative costs in sample ($) | 170800 | 784083 | 2218978 |
| Total administrative costs, scaled ($) | . | 120378 | 340673 |

Notes: Estimates are totals for the 200 villages in our self-targeting sample. Column (1) is directly estimated using the self-targeting sample, and Columns (2) and (3) are estimated using the PMT sample. Total population in Columns (2) and (3) are scaled to match Column (1). For number of eligible/ineligible households, total annual benefits paid, and total cost to households, the percentage of eligible households in the village for Columns (2) and (3) are also scaled to Column (1). All monetary costs are reported in U.S. dollars, using an exchange rate of 9,535 Rp. / USD$ 1.00 (October 2, 2012). Benefits per household are assumed to be Rp. 1.3 million annually. Costs to households are calculated as the time cost of travel, waiting, and completing surveys (in PMT, just the cost of completing surveys) using the household average wage rate, as well as the cost of transportation. Note that costs to households in self-targeting also includes the time cost of attending an informational meeting on the treatment. Wage rates and beneficiary/non-beneficiary breakdown of meeting attendees based on in-sample data; meeting attendance and length based on facilitators' meeting data. All households are assumed to stay for the entire meeting. Total administrative costs in sample are calculated based on per-village and per-neighborhood costs actually incurred by the experiment for Indonesian government surveyors in both self-targeting and PMT, as well as on actual costs by an external NGO that helped to spread information about self-targeting; since PMT treatment was done in one neighborhood only, the actual costs are scaled up by the average number of neighborhoods in a village. Total at scale administrative costs in PMT are based on the actual Indonesian government cost of executing the PMT nationwide, when they surveyed approximately 16 million households. The costs of PMT are assumed to be linear in the number of households surveyed per village.

TABLE 7. Experimental Results: Probability of Showing up as a Function of Distance and Log Per Capita Consumption

| | No stratum fixed effects | | | With stratum fixed effects | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Close subtreatment | 0.205 | 1.345 | 0.195 | 0.275 | 0.485 | 0.193 |
| | (0.146) | (2.841) | (0.238) | (0.168) | (2.920) | (0.310) |
| Log consumption | | -1.434*** | | | -1.446*** | |
| | | (0.143) | | | (0.144) | |
| Close subtreatment* Log consumption | | -0.093 | | | -0.023 | |
| | | (0.217) | | | (0.218) | |
| Consumption quintile 2 | | | -0.317 | | | -0.326 |
| | | | (0.233) | | | (0.245) |
| Consumption quintile 3 | | | -0.813*** | | | -0.791*** |
| | | | (0.231) | | | (0.234) |
| Consumption quintile 4 | | | -1.084*** | | | -1.072*** |
| | | | (0.206) | | | (0.234) |
| Consumption quintile 5 | | | -2.204*** | | | -2.265*** |
| | | | (0.257) | | | (0.279) |
| Close subtreatment * Consumption quintile 2 | | | -0.271 | | | -0.292 |
| | | | (0.323) | | | (0.368) |
| Close subtreatment * Consumption quintile 3 | | | 0.255 | | | 0.321 |
| | | | (0.299) | | | (0.325) |
| Close subtreatment * Consumption quintile 4 | | | -0.385 | | | -0.261 |
| | | | (0.300) | | | (0.314) |
| Close subtreatment * Consumption quintile 5 | | | 0.174 | | | 0.277 |
| | | | (0.371) | | | (0.387) |
| Stratum fixed effects | No | No | No | Yes | Yes | Yes |
| Observations | 2,000 | 2,000 | 2,000 | 1,960 | 1,960 | 1,960 |
| Mean of dependent variable | 0.377 | 0.377 | 0.377 | 0.385 | 0.385 | 0.385 |

Notes: Each column presents a logit regression of show up on the close sub-treatment. In Columns (1) - (3), robust standard errors are clustered at the village level. In Columns (4) - (6), robust standard errors are clustered at the stratum level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8. Experimental Results: Probability of Showing up as a Function of Opportunity Cost Treatment

| | No stratum fixed effects | | | With Stratum fixed effects | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Both spouse subtreatment | 0.196 | 4.303 | 0.461* | 0.185* | 3.334 | 0.384 |
| | (0.146) | (2.840) | (0.237) | (0.099) | (2.857) | (0.243) |
| Log consumption | | -1.324*** | | | -1.343*** | |
| | | (0.145) | | | (0.144) | |
| Both spouse subtreatment * Log consumption | | -0.318 | | | -0.244 | |
| | | (0.217) | | | (0.217) | |
| Consumption quintile 2 | | | -0.292 | | | -0.327 |
| | | | (0.212) | | | (0.219) |
| Consumption quintile 3 | | | -0.478** | | | -0.470** |
| | | | (0.190) | | | (0.184) |
| Consumption quintile 4 | | | -1.157*** | | | -1.146*** |
| | | | (0.185) | | | (0.205) |
| Consumption quintile 5 | | | -1.871*** | | | -1.962*** |
| | | | (0.271) | | | (0.289) |
| Both spouse subtreatment * Consumption quintile 2 | | | -0.348 | | | -0.316 |
| | | | (0.322) | | | (0.380) |
| Both spouse subtreatment * Consumption quintile 3 | | | -0.416 | | | -0.305 |
| | | | (0.292) | | | (0.344) |
| Both spouse subtreatment * Consumption quintile 4 | | | -0.237 | | | -0.116 |
| | | | (0.305) | | | (0.328) |
| Both spouse subtreatment * Consumption quintile 5 | | | -0.514 | | | -0.356 |
| | | | (0.369) | | | (0.347) |
| Stratum fixed effects | No | No | No | Yes | Yes | Yes |
| Observations | 2,000 | 2,000 | 2,000 | 1,960 | 1,960 | 1,960 |
| Mean of dependent variable | 0.377 | 0.377 | 0.377 | 0.385 | 0.385 | 0.385 |

Notes: Each column presents a logit regression of show up on the both spouse sub-treatment. In Columns (1) - (3), robust standard errors are clustered at the village level. In Columns (4) - (6), robust standard errors are clustered at the stratum level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 9. Estimated Parameter Values for the Model

| $\alpha_\varepsilon$ | $\beta_\varepsilon$ | $\rho$ |
|---|---|---|
| -26126 | 26805 | 6.09E-15 |
| (5445.492) | (8224.896) | (0.16011) |

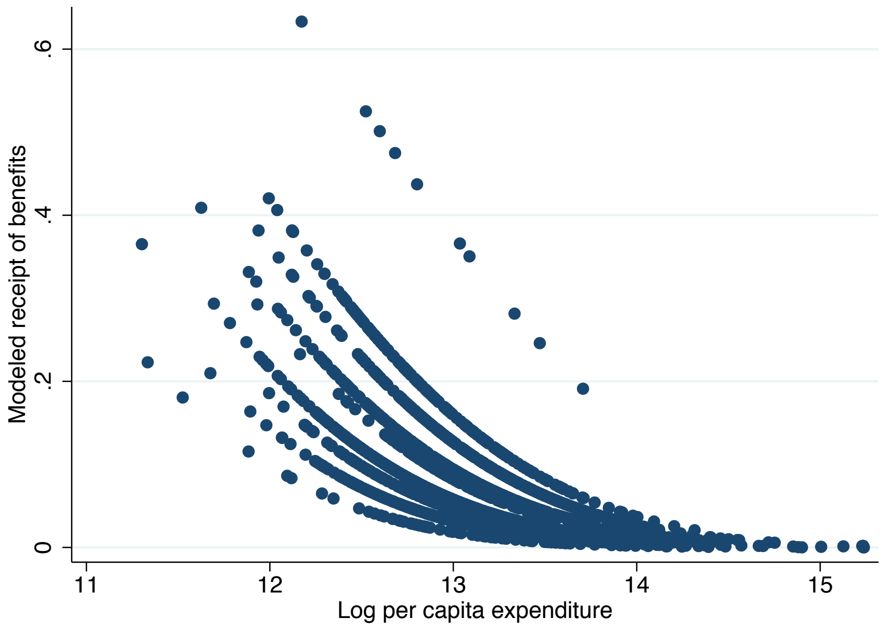
Notes: This table reports the mean and variance of the cost shock ($\varepsilon$) and the coefficient of relative risk aversion ($\rho$). The parameters are estimated using two-step feasible GMM. The moments are defined as the average show up rates within each consumption quintile. These five moments are fit only in the far treatment villages, assuming an annual discount factor of 0.5. Bootstrapped standard errors are in parentheses.

TABLE 10. Modeled Effects of Time and Distance Costs on Show Up Rates

| | Show Up (Exp.) | Predicted Show Up (Model)† | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Reported | Reported | Reported | Assuming No | Reported | Additional Distance | | Inflated Wait Time | |
| | | Total Cost | Total cost, SD[eps]/2 | total cost, SD[eps]=0 | Differential Travel Cost | total cost, constant mu | Distance + 3km | Distance + 6km | Wait Time*3 | Wait Time*6 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| *Panel A: Logistic Regressions* | | | | | | | | | | |
| Close | 1.563 | -1.654 | -2.203 | -2.378 | -1.545 | -1.690 | -1.785 | -1.614 | -4.659 | -7.307** |
| | (2.813) | (3.019) | (3.395) | (3.540) | (2.919) | (2.356) | (3.038) | (2.833) | (2.974) | (3.245) |
| Log per capita expenditure | -1.419*** | -1.450*** | -1.955*** | -2.208*** | -1.442*** | -0.328** | -1.465*** | -1.454*** | -1.700*** | -1.927*** |
| | (0.145) | (0.164) | (0.183) | (0.200) | (0.164) | (0.128) | (0.168) | (0.164) | (0.171) | (0.194) |
| Close * Log per capita expenditure | -0.109 | 0.134 | 0.178 | 0.191 | 0.125 | 0.138 | 0.149 | 0.139 | 0.385 | 0.611** |
| | (0.215) | (0.231) | (0.261) | (0.273) | (0.223) | (0.180) | (0.232) | (0.217) | (0.228) | (0.249) |
| N | 1973 | 5919000 | 5919000 | 5919000 | 5913000 | 5919000 | 5913000 | 5913000 | 5919000 | 5919000 |
| P-value‡ | | 0.441 | 0.397 | 0.388 | 0.449 | 0.379 | 0.415 | 0.417 | 0.115 | 0.029 |
| *Panel B: Show-Up Rates* | | | | | | | | | | |
| Above poverty line, far | 34.123 | 33.165 | 27.376 | 24.021 | 33.350 | 32.070 | 31.875 | 31.211 | 28.104 | 23.036 |
| Above poverty line, close | 39.116 | 37.465 | 31.807 | 28.157 | 37.463 | 35.164 | 37.465 | 37.465 | 37.465 | 37.465 |
| Below poverty line, far | 54.237 | 69.910 | 71.484 | 71.719 | 69.895 | 37.367 | 68.882 | 67.969 | 67.456 | 64.047 |
| Below poverty line, close | 57.895 | 67.194 | 68.116 | 67.618 | 67.211 | 36.866 | 67.194 | 67.194 | 67.194 | 67.194 |
| *Panel C: Show-Up Rate Ratios* | | | | | | | | | | |
| Poor to rich ratio, far | 1.589 | 2.108 | 2.611 | 2.986 | 2.096 | 1.165 | 2.161 | 2.178 | 2.400 | 2.780 |
| | (0.215) | (0.213) | (0.278) | (0.335) | (0.206) | (0.201) | (0.218) | (0.224) | (0.261) | (0.348) |
| Poor to rich ratio, close | 1.480 | 1.793 | 2.142 | 2.401 | 1.794 | 1.048 | 1.793 | 1.793 | 1.793 | 1.793 |
| | (0.177) | (0.178) | (0.228) | (0.255) | (0.185) | (0.186) | (0.182) | (0.184) | (0.191) | (0.185) |
| Difference of ratios | 0.109 | 0.314 | 0.470 | 0.584 | 0.302 | 0.117 | 0.368 | 0.384 | 0.607* | 0.987** |
| P-value | (0.278) | (0.277) | (0.359) | (0.428) | (0.276) | (0.273) | (0.285) | (0.291) | (0.330) | (0.390) |
| | | 0.602 | 0.428 | 0.352 | 0.624 | 0.985 | 0.517 | 0.495 | 0.249 | 0.067 |

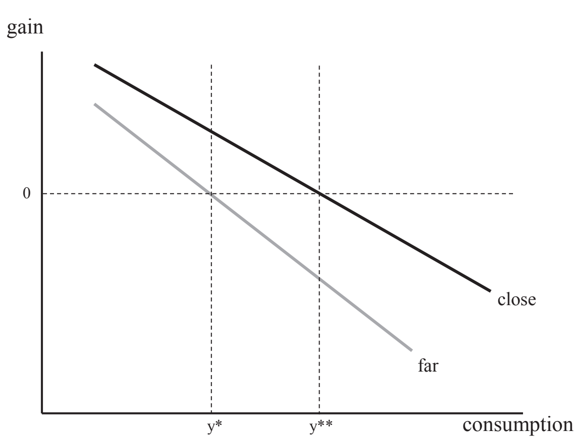
Notes: In order to run logits on predicted show up rates, we create 3000 copies of the data. The copies of each individual are assigned to show up or not in proportion to his predicted probability of showing up. Bootstrapped standard errors, clustered by village, are in parentheses. To compute the standard errors, for each bootstrap iteration we sample 2,000 households, clustered at the village level, to make the sample equivalent to that in Column 1. We perform 1,000 bootstrap iterations. The p-value in Panel A is the test of whether the coefficient on $[Close * LogPCE]$ is equal to the equivalent coefficient in Column 1. The p-value in Panel C is the test of whether the difference in ratios is equal to the difference in ratios in Column 1. *** p<0.01, ** p<0.05, * p<0.1 Significance levels not shown on first two rows of Panel C.

Figure 1. Probability of Obtaining Benefits vs. Log Per Capita Consumption

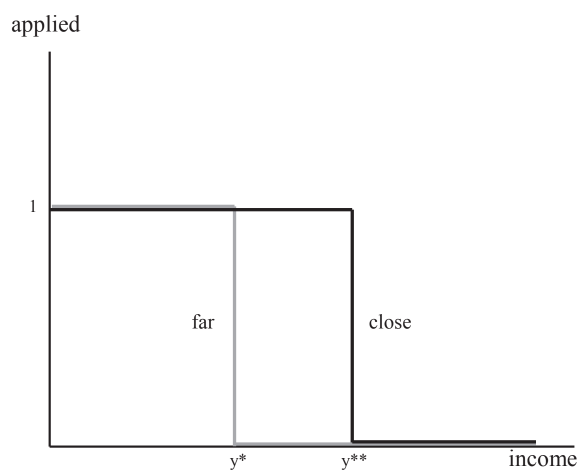



Notes: This figure shows the predicted probability of receiving the benefit, conditional on applying, from a probit model of receiving a benefit as a function of Log PCE. We include urban/rural interacted with district fixed effects in the probit, since the PMT cutoff for inclusion varies slightly for each urban/rural times district cell. These predicted values are the $\mu(y_i)$ that we use in the model.

Figure 2. Illustration of Utility Gain with No Errors

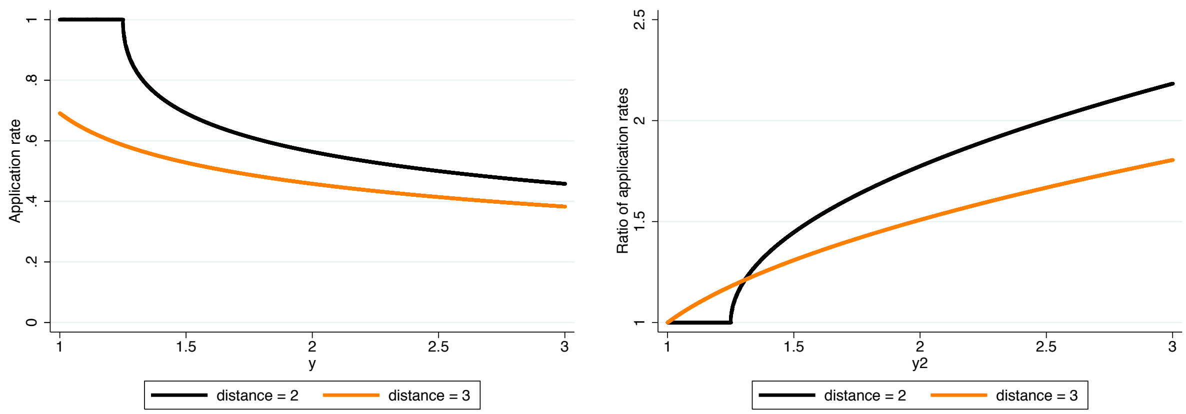



(a) Gain vs. Consumption for Close and Far Sub-Treatments




(b) Targeting Improves as Length of Ordeal Increases

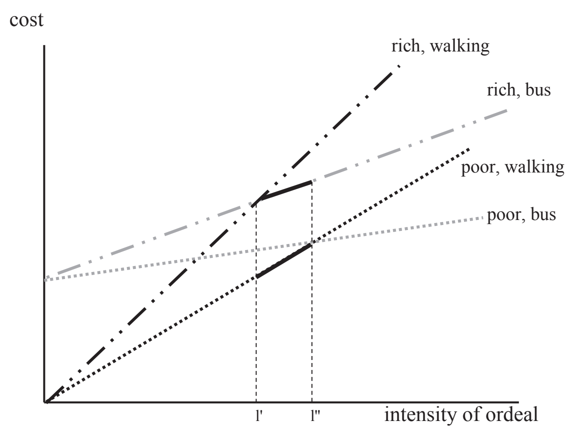
FIGURE 3. Illustration of Utility Gain with Log-logistic Errors




(A) Show Up Rates with Log-logistic Errors

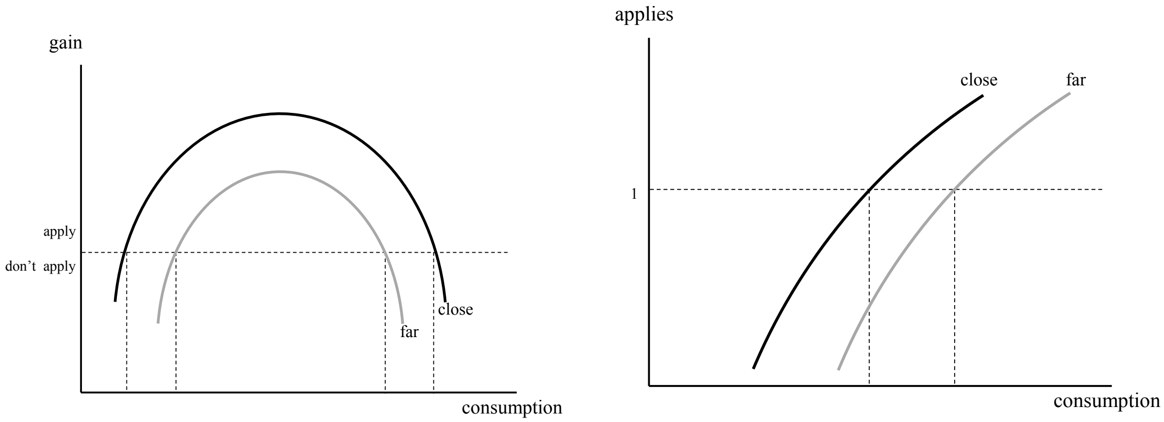
(B) Ratio of Show Up Rates of Rich ($y_2$) Compared to Poor ($y_1 = 1$)

FIGURE 4. Non-Linearities in Travel Costs




Notes: Increasing ordeal within l' to l", marginal cost for rich is lower than marginal cost for the poor.

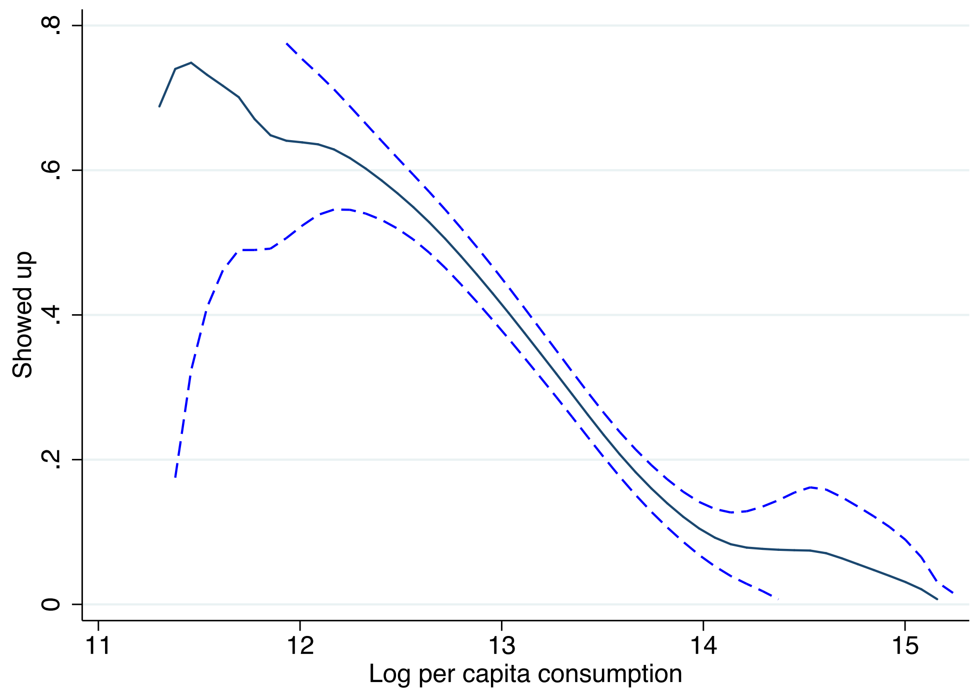
Figure 5. Illustration of Utility Gain with Concave Utility




(A) Gain vs. Consumption for Close and Far Sub-Treatments

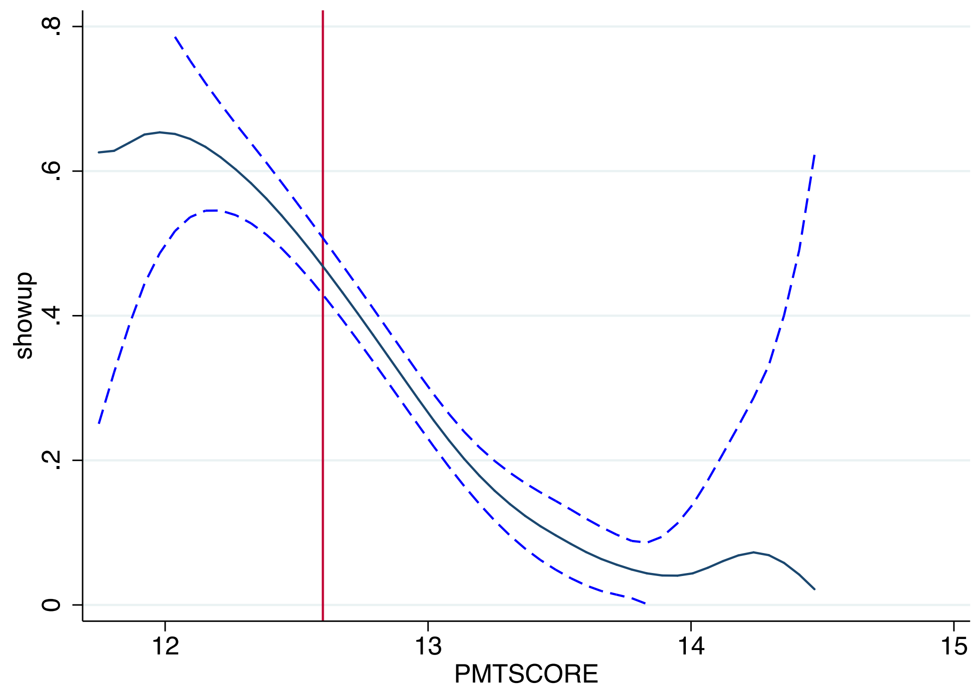
(B) Targeting can Worsen as Length of Ordeal Increases

Figure 6. Show Up Rates Versus Log Per Capita Consumption

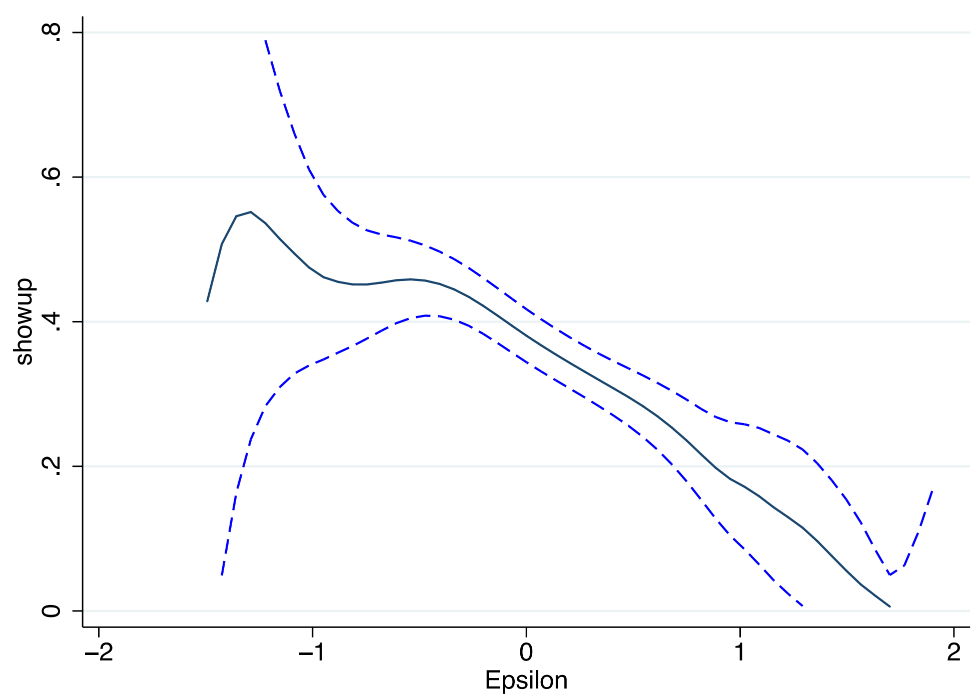



Notes: Figure provides a non-parametric fan regression of the probability of applying for PKH against baseline log per capita consumption in the 200 self-targeting villages. Bootstrapped standard error bounds, clustered at the village level, are shown in dashes.

Figure 7. Show Up Rates Versus Observable and Unobservable Components of Log Per Capita Consumption




(a) Show Up as a Function of Observable Consumption ($X_i'\beta$)

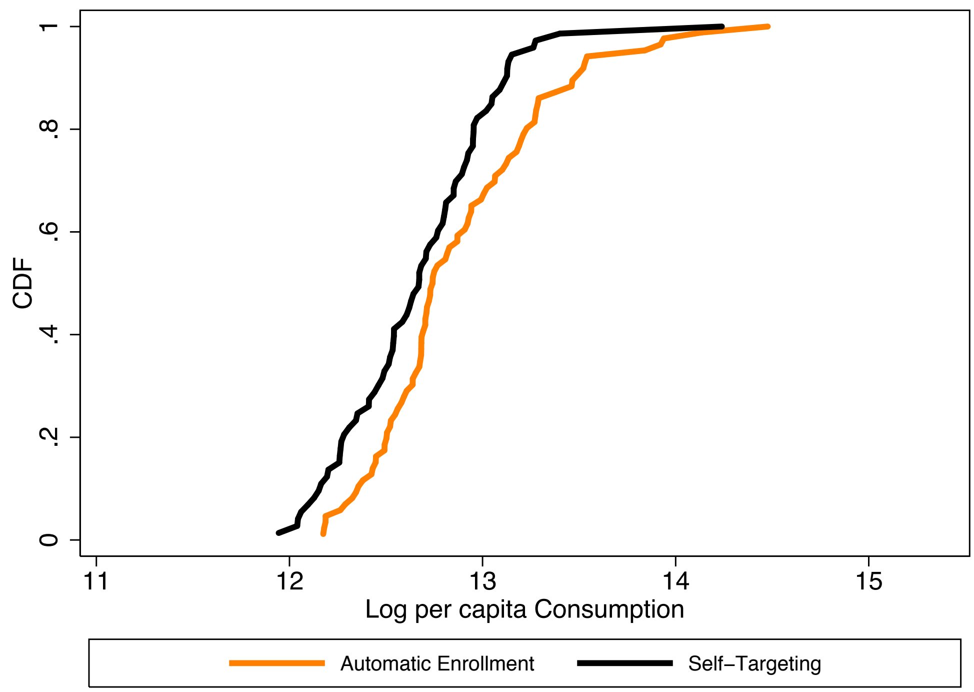



(b) Show Up as a Function of Unobservable Consumption ($\varepsilon_i$)

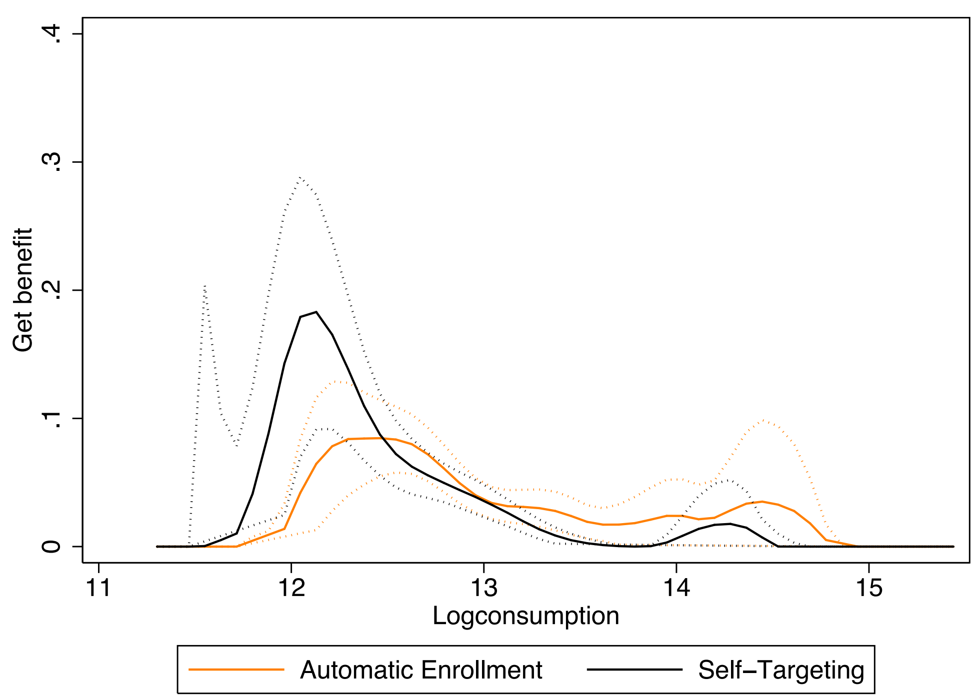


Notes: Figures provide non-parametric fan regressions of the probability of applying for PKH against components of baseline log per capita consumption in the 200 self-targeting villages. Bootstrapped standard error bounds, clustered at the village level, are shown in dashes.

Figure 8. Experimental Comparison of Self Targeting and Automatic Enrollment Treatments




(a) CDF of Log Per Capita Consumption of Beneficiaries

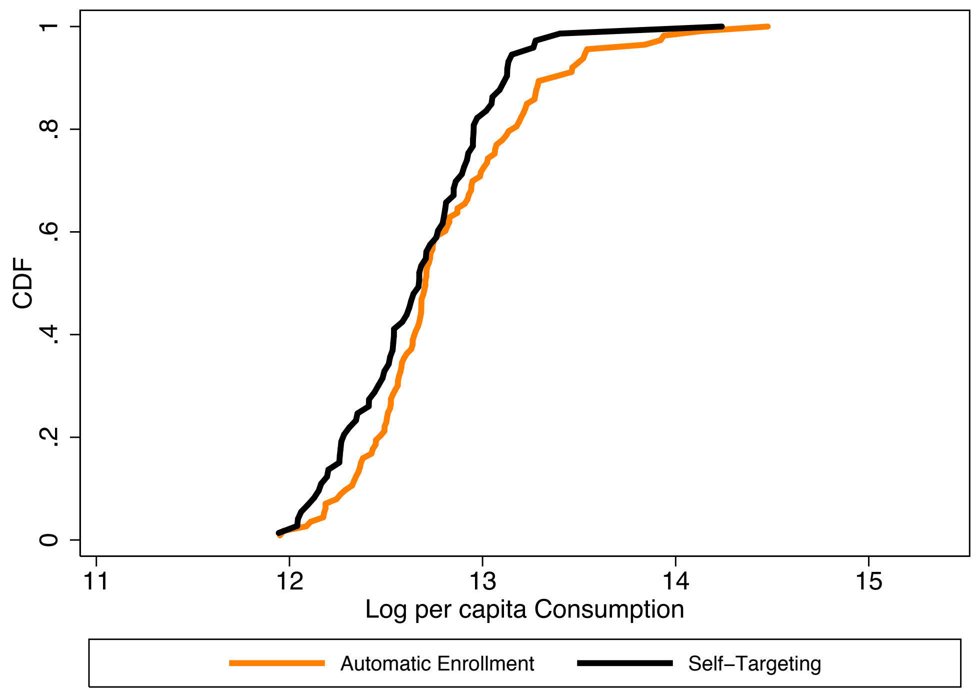



(b) Receiving Benefit as a Function of Log Per Capita Consumption

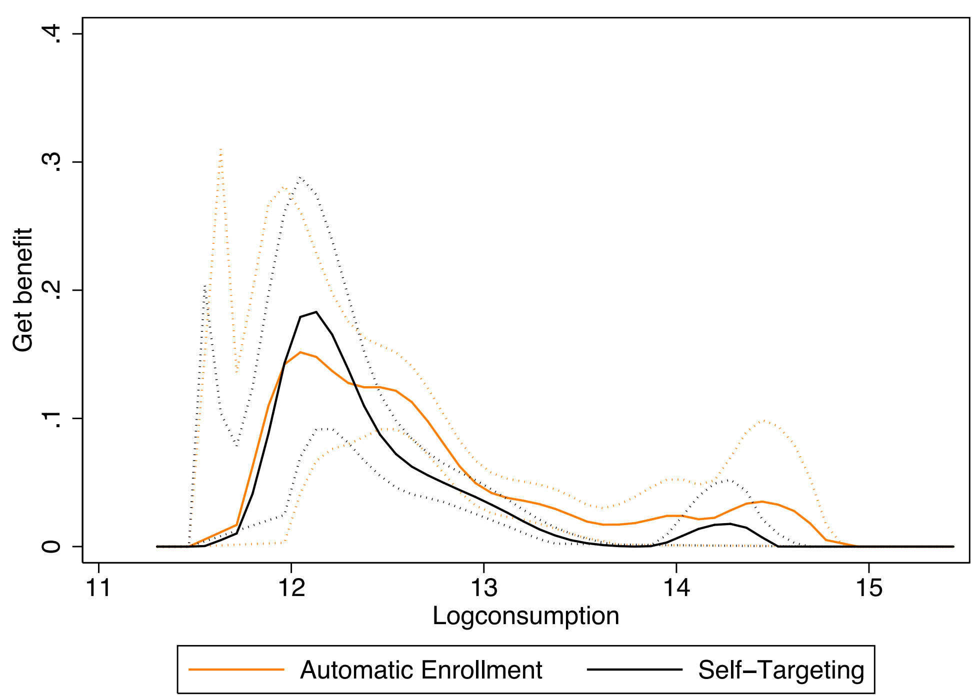


Notes: Panel A shows a CDF of log per capita consumption of beneficiaries. Kolmogorov-Smirnov test of equality yields a p-value of 0.10. Panel B presents a non-parametric Fan regression of benefit receipt on log per capita consumption. Bootstrapped standard errors, clustered at the village level, are shown in dashes.

Figure 9. Comparison of Self-Selection and Hypothetical Universal Automatic Enrollment




(a) CDF of Consumption of Beneficiaries

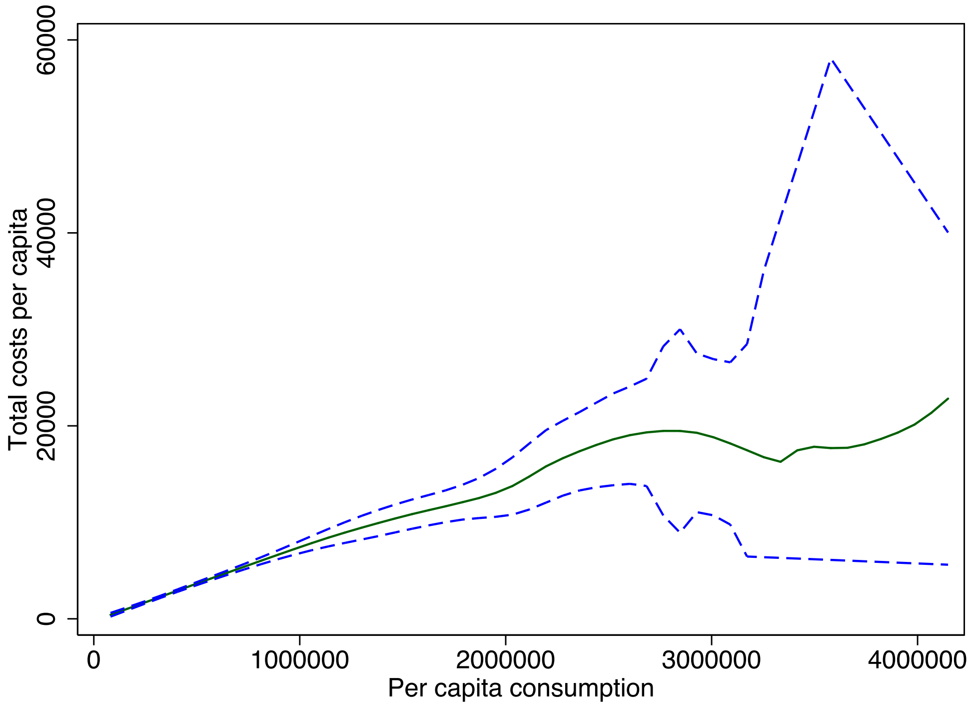



(b) Getting Benefit as a Function of Log Per Capita Consumption



Notes: Panel A shows a CDF of log per-capita consumption of beneficiaries. Kolmogorov-Smirnov test of equality yields a p-value of 0.29. Panel B presents a non-parametric Fan regression of benefit receipt on log per capita consumption. Bootstrapped standard errors, clustered at the village level, are shown in dashes.

FIGURE 10. Cost of Applying by Consumption




Notes: Figure shows a non-parametric Fan regression of total costs incurred in applying for PKH against per capita consumption. Bootstrapped standard errors, clustered at the village level, are shown in dashes. Costs assume one individual per household goes to sign-up location, even for households in opportunity cost sub-treatment.

Figure 11. Model Fit and Counterfactuals